

デジタル人文学と雑誌 *Punch*

松木 孝幸^{†1} Lowe, Robert J.^{†2} 平野 真理^{†3}

(令和1年11月28日査読受理日)

Digital Humanities and *Punch*, or The London Charivari

Matsuki, Takayuki^{†1} Lowe, Robert J.^{†2} Hirano, Mari^{†3}

(Accepted for publication 28 November, 2019)

要約

この論文では、イギリスで150年間発行された週刊風刺雑誌 *Punch* の全データをテキスト化し、服飾史、英語教育、心理学の3分野における専門単語の関連性や単語の時代変遷等の統計的解析を実施し、その結果の分析を行った。服飾史においてはKHcoderと呼ばれる統計解析ソフトを用いた解析方法を提示し、共起ネットワーク図、対応分析図の出力、及び文書×抽出語の出力を行い、更にTF-IDF値を服飾史に良く表れる単語について計算を行った。これらの結果から結論づけられたことは、単語間の関係性を理解するためには元データを更にフィルターを掛けて絞り込むことが必要であると判明した。また、英語教育の分野では、150年分の変遷図を見ることにより、英語の使用が世界的に年を追って増加していくこと、及び英語教育が英国の特別な機関と連携していることを示す。心理学においては、300の性格語の使用頻度変化を検討することにより、時代によって社会から注目される人の性格のあり方が反映されることを確認した。

Abstract

In this paper, we generate text data extracted from the weekly magazine, *Punch*, or The London Charivari, which was published in England from 1841 through 1992. We perform an analysis of the history of clothing, English education, and psychology and determine the relationship between terminologies, change of terms of frequencies, etc. For the history of clothing, we use the software KH Coder to analyze the data, which generates figures of co-occurrence networks and correspondence analysis, and 'Export Document-Word Matrix'. Further, we calculate the values of TFIDF for frequently appearing terms. According to these analyses, we need to filter the data further by selecting only sentences which include terms from the dictionary for the history of clothing. In English education, by looking at the data on frequently appearing terms over the last 150-years, we show how the international use of English has increased over time and how the notion of English teaching was closely connected to particular British institutions. In terms of psychology, by examining changes in the frequency of use of 300 personality words, it was confirmed that the personality of people attracting attention from society was reflected by that particular period.

キーワード：デジタル人文学, *Punch*, 服飾史, 英語教育, 性格, ビッグデータ

Key words: digital humanities, *Punch*, history of clothing, English education, personality, big data

1. はじめに

日本語のポンチ絵の語源であるイギリスの週刊雑誌パンチ(正式名称は *Punch, or the London Charivari*)は、1841年から150年の長きにわたって発行され、1992年に刊行を終えた。従来、この雑誌を用いた19世紀や20世紀のイギリス文化についての研究方法としては、注目する年代の中心となる用語・考えを中心に熟読し、読み込んだ後に解説、あるいは以前の結果との比較対象で新しい知見を披露している。例えば、ファッションに特化した内容 [1]、パンチ誌全体に目を向けた内容 [2]、また、イラストを描いたジョン・リーチにスポットを当てた内容 [3]、等の書籍が出版されている。これらの書籍は、何年もかけてパンチを読み込んだ研究者により書かれている。

しかし、雑誌パンチに埋め込まれた内容を適切に抜き出

して詳細にかつもっと簡便に研究をしたいという欲求が沸き起こる。いわゆるテキストマイニングによる研究である。これを可能にするためにはパンチ誌の電子データ化が必須であり、それを実現して提供してくれたのがセンゲージ社による *Punch* のデータである [4]。センゲージにより提供されているデータは、*Punch* 誌の各ページのJPEG画像とその画像をOCRで文字化したXMLファイルであり、各ファイルが更に5種類に分類されている。

1980年代から現代にかけてデジタル技術を人文学に適用して研究する分野が現れた。デジタル人文学あるいはヒューマニティーズ・コンピューティングとも呼ばれる [5]。この論文では、*Punch* の電子データを対象にデジタル技術を用いてどのような分析が出来るかの探索的研究を試みる。まずは服飾分野を例に取り、テキストマイニングに強力な手助けとなる KHCoder ソフトを用いたデータの統計分析

†1 東京家政大学環境教育学科

†2 東京家政大学英語コミュニケーション学科

†3 東京家政大学心理カウンセリング学科

手法を説明する。こうした分析は、従来の熟読・読み込みによる解釈・研究を否定するものではなく、それを補完するものと考えている。

Punch の全データを取得する以前の本学での電子化の試みとしては、ヘミングウェイの 15 篇の短編小説集および同時代の比較対象としたフィッツジェラルドの小説を電子化し、統計量 TF-IDF による語彙分析、内容分析を用いた解析を試み、作品の概要把握や作品解釈のヒントとなる語が上位に表れるという知見が得られた [6]。また、フランス王妃マリー・アントワネットの衣装を納入していたエロフ夫人のフランス語の商業日誌『エロフ夫人の日誌』の 1787 年から 1793 年分を電子化し、服飾関連単語の解析を行い、その結果のデータベース化および各年の推移の傾向観察などで一定の成果を得ている [7]。しかし、ヘミングウェイと『エロフ夫人の日誌』のどちらも出現単語や出現頻度のデータベース化とその傾向の観察にとどまっており、そのデータをフルに利用した統計的な分析までには至っていない。そこで根津 [8] は、19 世紀の雑誌 *Punch* の手作業による電子化を数年試み、*Punch* に多い花文字による文字化けについてもできる限り訂正を行った。それにより得られた 6 年分の電子データを用いて、服飾史に関する統計解析を行い、十分な成果が得られた。

本論文では、電子化された *Punch* 誌 [4] から得られるビッグデータを適切に加工して得られた電子データを用いて、先行研究 [8] において 6 年分で行われた分析を 23 年分に置き換え、さらに KHCoder によるテキストマイニングを行うことでどのような解析が可能であるかを吟味検討する。また、同じ *Punch* の 150 年分のデータについて、他の分野である、英語教育、心理学の視点から解析を行うことで何が得られるかを考察する。

本論文は以下の構成を取る。まず次の章では、*Punch* 誌 [4] より提供されたデータをどのように加工して、今回の解析に使用できるデータを作成したかを述べる。3 章では、得られたデータを用いて、テキストマイニング解析用ソフトである KHCoder を用いて、服飾史の解析を試みる [9,10,11]。そのために、KHCoder が出力できる対応分析図、共起ネットワーク図、及びそれとは別個に 19 世紀後半に頻繁に使われていた単語についての TF-IDF 値の計算結果を示す。続く第 4 章と第 5 章では、人文学の異なる分野において、本データ分析がどのように応用可能かを検討する。まず第 4 章では、英語教育で重要となる単語集を作成し、150 年の変化を示すグラフを観察し、何が読み取れるかの解説を試みる。続く第 5 章では、心理学の観点から、人の性格を表す単語の出現頻度の経年変化を観察し、社会変化と人間理解の関連がどのように読み取れるかについて考察する。最後の章は、まとめと結論に当てられる。

2. データ作成について

センゲージ会社により提供された生データは、序章で書いたように各ページの JPEG 画像と OCR ソフトにより電子化されたデータを更に XML 化したファイル群である。テキストファイルは、1 週ごとにまとめられており、152 年で 7,833 個のファイルとなる。生データ以外には、図 1 で示されるホームページで閲覧できるようにしてあり、簡単な検索も出来る仕様である。以下の英語と心理の分野の研究のために、このページに備わっている「Term Frequency (頻出語)」のメニューを利用し、152 年間の各単語の出現回数を PNG 図と CSV ファイルの両方で出力できる機能を利用した。

Punch のデータで特徴的なことは、風刺絵が必ず描かれていることであり、この雑誌の研究者は必ずと言って良いほど、この絵から話を始める。しかし、我々の目的はテキストマイニングにあるので、XML ファイルから本文が埋め込まれている部分を抽出し、更に、週刊誌を年ごとのデータにまとめるという作業を行わなければならない。また、KHCoder でデータを使用するためには、各年毎の文章の先頭部分に「<H1>1841</H1>」のような年度をタグで挟んだ文の挿入が必要となる。我々は C#言語を用いて数百行のコードで書き、7,833 個の XML ファイルから 152 個の各年データを生成し、更に各年の先頭に年度タグを挿入して、これらを 1 つのファイルにまとめて最終的なデータを得た。得られたデータには、主に花文字に起因する誤字・記号・2 バイトコードが含まれているが、大勢には影響が無いとして解析を進めることとした。我々の経験では、これらの修正には多大な時間とマンパワーが必要となり、今回のような研究では実現不可能なことで成るからである。



図 1 *Punch* の閲覧ページ

3. 服飾史の解析方法

この章では、服飾史そのものの解析は行わずに、解析方法にどのようなものがあるかを述べる。テキストマイニング解析用ソフトである KHCoder はとても多彩な解析方法を提供する。ここでは、それら全ての機能を使わずに、主な

方法のみの結果を提示する。KHcoder を使用した服飾史の解析を行うが、根津による研究 [8]を基にして分析する。この研究 [8]では 1870 年-1872 年 (Vol. 58-Vol. 63), 1890 年-1892 年 (Vol. 98-Vol. 103) の計 6 年分の雑誌を対象としたが、この論文では、19 世紀末の 1870~1892 年を対象年とする。この年度を対象とした理由は、1870-90 年頃にイギリスで服装改良運動がおり、またイギリス文芸界に起こった美学運動が服飾にも影響を及ぼしていたことについて探求するためである。従って、Punch から生成したデータのうち、23 年分のデータを取り出し、解析対象とした。この解析では、各ドキュメントというのは各年度を意味する。

3.1 対応分析

抽出した単語の対応分析分布図を図 2 に表示する。この図の意味するところは、文書種類（年度別）と個々の単語の出現回数を縦・横軸に対応分析したものである。対応分析とは、クロス集計表（分割表）の行項目と列項目の相関が最大になるように行列を並び替えることで、出現パターンの似通ったもの同士が近接するように処理を行う方法である。コレスポネンス分析ともいわれる。また、処理されたデータを 2 次元の散布図として可視化することができ、より感覚的な分析が可能となる。分析においては、データに年代を HTML タグで分けたテキストデータを使用した。

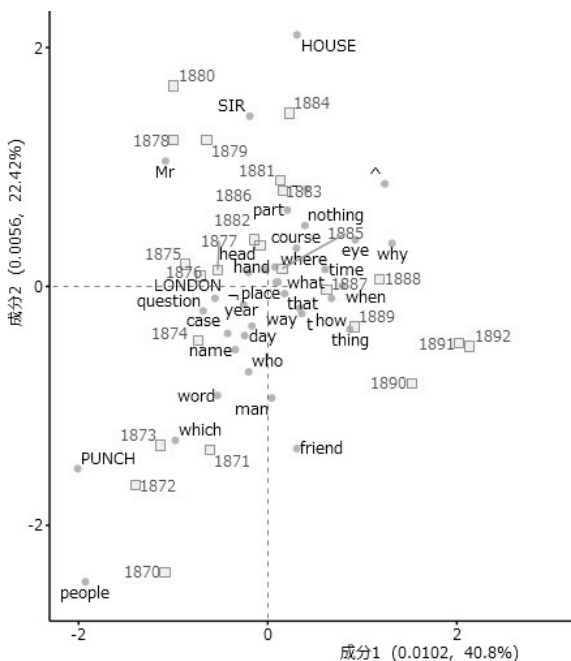


図 2 1870~1892 年の対応分析結果

この図をより意味のある図とするためには、コンコーダンスを用いて、我々の使用した服飾単語辞典に現れる単語のみを含む文章を抽出したデータを作成し、分析し直す必要がある。実際、そのように処理したデータを文献 [7] で使用している。他の解析についてもこれと同じデータを使

用する必要があると思われる。今回は、そのようなデータ抽出を行うまでには至らず、次期課題としておく。

3.2 共起ネットワーク

共起ネットワークでは、出現パターンが類似している語、つまり強い共起関係にある語を線で結び、語のネットワークを描くことができる。共起の強弱の計算には Jaccard 係数が用いられている。Jaccard 係数とは、2 つの集合があったとき、共通要素の割合からその類似度を表す数値表現の 1 つである。Jaccard 係数は以下の数式で表される。

$$\text{Jaccard 係数} = \frac{X \cap Y}{X \cup Y}$$

分母 XUY は 2 集団の要素の総数、分子 X∩Y は 2 集団の同じ要素の数である。この数値が高いほど類似している、つまり共起関係にあると言える。

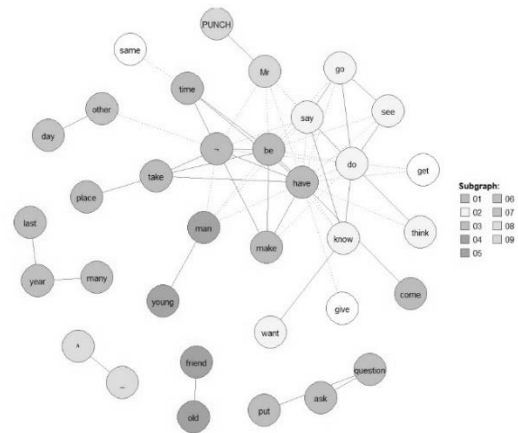


図 3 1870~1892 年の共起ネットワーク

共起ネットワークを描く基データとして、1870 年から 1992 年までの 23 年分の全データを使用したために、服飾史に関連する単語は大多数を占める冠詞や動詞に隠れてしまい、表に出て共起関係を示すことが出来なかった。我々が用いた根津の服飾辞書は、[2] に示された 3 つの文献から抽出された 39 単語からなる。共起関係を示すためには、服飾史に関連する辞書内の単語を含む文章のみをコンコーダンスを用いて取り出して共起ネットワークを描くことができるが、これは次の論文の課題とする。

3.3 文書×抽出語と TF-IDF 値

	D	E	F	G	H	I	J
1	name	length_c	length_w	man	time	day	thing
2	1870	2180363	563640		968	574	555
3	1871	2041459	532355		874	593	549
4	1872	2093978	539714		816	639	595
5	1873	1948252	504942		770	576	535
6	1874	1900680	494637		724	522	495
7	1875	1940612	504661		724	527	466
8	1876	2175117	565700		769	575	529
9	1877	2344684	610621		835	741	580
10	1878	2326278	607015		801	668	531
11	1879	2255691	591905		731	702	518
12	1880	2276701	588574		685	625	507
13	1881	2338984	613158		937	730	594
14	1882	2257232	589388		835	775	564

図 4 文書×抽出語

KHcoder でドキュメント（年）毎に抽出語を一覧にする機能が文書×抽出語である。その一覧の一部を図 4 に示す。

これを用いて、いわゆる TF-IDF 値を求めることが出来る。上記の図 4 では、服飾に関係の無い語が最頻出語として表示されているが、これとは別に服飾語辞典の一覧に現れる語の頻出度を求めた表があり、その表と各年度の総単語数を用いることにより、下記の式を使い TF-IDF 値を求めることが出来る。

$$(1) tf_i df_{i,j} = tf_{i,j} \times idf_i$$

$$(2) tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}}$$

$$(3) idf_i = \log \frac{|D|}{|\{d:d \ni t_i\}|}$$

(2) 式中の TF とは、Term Frequency の略で、文書 j 中の単語 i の出現頻度を表す。出現する頻度が高い単語ほど、その文書の特徴づけるといふ指標である。 $n_{i,j}$ は文書 j 中の単語 i の出現回数、 $\sum_k n_{k,j}$ は文書 j に出現した単語の総数である。(3) 式中の IDF とは、Inverse Document Frequency の略で、全文書中の単語 i を含む文書の数を表す DF の逆数の対数を取った量である。出現する文書数が少ないほど、その単語は希少性があると言える。 $|D|$ は総全文書数、 $|\{d:d \ni t_i\}|$ は単語 t_i が出現する文書数である。この計算方法では、全文書中に単語 t_i が存在した場合、IDF 値が 0 となる。全文書中に 1 つも単語が現れないときには、 $|\{d:d \ni t_i\}|$ が 0 となり、無限大が生じてしまうために、この場合の TF-IDF 値は考慮しない。また、この論文の中では、1 文書というのは 1 年間の雑誌全体を意味する。服飾関連辞書には 39 単語を採用した。それらの TF-IDF 値の一部を図 5 に示す。

	basque	bell-boy jacket	blouse	bodice	bolero
1870	4.71	DIV/0	0.00	3.26	0.00
1871	0.00	DIV/0	2.98	0.00	0.00
1872	0.00	DIV/0	0.97	0.00	0.00
1873	0.00	DIV/0	1.04	0.00	5.24
1874	0.00	DIV/0	1.07	0.00	5.39
1875	5.31	DIV/0	0.00	0.00	0.00
1876	0.00	DIV/0	0.00	0.00	0.00
1877	4.38	DIV/0	1.73	0.00	0.00
1878	4.41	DIV/0	0.00	0.00	0.00

図 5 服飾関連単語の TF-IDF 値

図 5 に示されている DIV/0 は、0 で割る演算が起きたと云う意味である。また、1 文書（1 年間）当たりの総単語数は、図 4 から見て取れるように 50~60 万単語であるので、(2) 式から分かるように図 5 の表の値に 10^{-6} を掛けた値が実際の TF-IDF 値となる。この一覧を見ると、現代で使われるような単語（bell-boy jacket, camisole, corselet 等）の TF-IDF 値は 23 年間に渡って現れないために、DIV/0、すなわち不定となっている。

4. 英語教育の解析

続いて、他分野への応用可能性を検討するために、*Punch* 誌に見られる英語および言語教育に関連する特定の用語の出現頻度に焦点を置いて議論する。解析の結果、特に興味深いパターンが 2 つ見受けられた。国際英語および世界英語に関連する単語類と、英語教育への特定の組織を連想させる単語類である。それぞれを以下で考察する。

4.1 パターン 1: 英語の普及と多様性

特筆すべき第一のパターンは、「国際英語 (international English)」「国際言語 (international language)」「リングワ・フランカ (lingua franca)」のような英単語に見られるように「international」あるいは「English」が含まれる単語に関するものである。また、特定の地域で発達した英語である「インド英語 (Indian English)」や「シンガポール英語 (Singaporean English)」などの単語についても言及したい。データからは、英語が国際的な言語として認知されて数十年経つにも関わらず、また独自の形式と機能性を持った地域的多様性のある英語が存在するにも関わらず、これらの事実があまり認知されていないことが示唆されている。

現代において英語が国際的な言語であることは広く認められており、クリスタル [13] によれば、全世界への英語の普及に影響を与えた地理的および歴史的な要素は多数ある。英国およびアメリカの植民地政策は、英語圏の国家に統治された地域において英語が重要な役割を担うようになるという役割を果たした。世界的な貿易に大きく関わった英国の産業主要国としての地位も、その国際的な権力によって貿易において使用される言語に大きな影響を及ぼしたため、英語の世界的利用の拡がりをもたらした。その後のアメリカの経済的影響力も英語の世界的普及に貢献し、現在では英語はグローバルビジネスの事実上の公用語とみなされている。また、ハリウッドからのマスメディアの普及という文化的な流れもこの言語の国際的な普及に影響を及ぼしたと言える。これら全ての要素が合わさり、英語は広く「グローバル言語」 [13]、「国際 (インターナショナル) 言語」 [14] もしくは「lingua franca」 [15] [16] [17] とみなされ、このように呼ばれる結果となった。

Punch 誌に見られる単語の統計的特徴は、これらの論説が経時的に受け入れられていく度合いを表していると考えられた。まずは「国際英語 (international English)」 (図 6) というフレーズから見ていく。*Punch* 誌においてこのフレーズの使用回数は、時間と共に微増していったことが示されているものの、大幅な増加は見られず、定期的に使用されているわけでもない。英語に限定しているわけではないが密接に関連する「国際言語 (international language)」や「リングワ・フランカ (lingua franca)」 (図 7, 8) というフレーズについても、使用頻度はほとんど増加していない。これ

は、英語が世界的な普及を見せたという大きな流れがあったにも関わらず、そのことに関する論議が *Punch* 誌の論説に組み込まれた様子は見られず、学術領域に限定されていた可能性を示しているといえるだろう。*Punch* は、英語の世界的な普及を認識すべきはずの風刺雑誌であるため、これは驚くべきことである。

英語の多様性に目を移しても、*Punch* 誌においてそれはほぼ認識されず、表現もされていない。既述の通り、英語は英国の植民地政策を部分的なきっかけとして世界的に広まった。そのため、インドや香港、シンガポールといった英国の植民地では、英語は自然に高い影響力を持つこととなった。植民地の義務教育制度に英語が含まれることに関しては、英国人の間で幾らかの懸念も見受けられたが、英語教育がこれらの植民地の文治への鍵と見られていたことは事実であり、そのことに対する需要は大きかったと言える [18]。これらの旧植民地では、独自の語彙や発音、および文法を持つ英語の特定の派生形や方言が時間と共に発達した。主要な派生言語としては、インド英語 [19]、香港英語 [20]、シンガポール英語 [21] が挙げられる。これらの派生言語の発達は、現在では広く認識されているが、*Punch* 誌データの統計において「シンガポール英語 (Singaporean English)」や「ナイジェリア英語 (Nigerian English)」といった用語はほとんど見られていない。しかし、図9に見られるとおり、「インド英語 (Indian English)」という用語の使用は時間とともにいくらか増えている。英語の多様性の増加を反映して、英国やアメリカのスタンダードな英語以外にも、こうした派生言語が存在するという理解が深まっていることを表しているといえよう。ただし統計を見ると、インド英語については徐々に認識され使用されているという兆候がある程度見受けられるものの、それ以外の多様化された英語については認識されていないことが示されている。

これらの統計は、*Punch* 誌が出版されている期間において英語が国際的な言語として普及し多様化を見せたにも関わらず、この事実がほとんど認識されておらず、また紙面においてまれにしか言及されないことを示しているといえるだろう。

4.2 パターン2：英語教育とブリティッシュ・カウンシル

統計から明らかになった興味深いパターンの二つ目は、英語の教育に関連する用語の使用と、英国と関係のある特定の文化的および教育的機関の存在である。この項では「ブリティッシュ・カウンシル (British Council)」、「英語の教え (English teaching)」および「英語を教える (teaching English)」という用語の使用の増加に焦点を当てる。ブリティッシュ・カウンシルは「当初は地中海におけるファシストのプロパガンダへの対抗策」として、1934年に設立された [22] (p. 240)。この組織の存在は海外での英語教育の促進と、その教育を通じた英国文化の知識と

理解の奨励を目的としていた。ブリティッシュ・カウンシルは、色々な意味で、現実の帝国が衰退し始める中で英国の利益を促進しようとする「ソフトパワー」の一種であったと捉えることができる。初期においては、そうした文化的な取り組みを中心に行っており、そのことはカウンシルの名称の原型である「他国との関係のためのブリティッシュ・カウンシル」に明白に表れている [23] が、現代においては英語教育のイメージが最も強くなっている。ブリティッシュ・カウンシルは一般的に英語教育の分野における権威的かつ影響力の強い組織として認識されており、世界中にあるオフィスにおいて、その国における英語教育を改善する様々な履修課程プログラムに携わっている。

統計内では、「ブリティッシュ・カウンシル (British Council)」、「英語の教え (English teaching)」および「英語を教える (English teaching)」というフレーズにおいて相関的な増加傾向が見られた。このことはいくつかの理由で重要である。第一に、ブリティッシュ・カウンシルが *Punch* のような風刺雑誌において注目を集めていたことである。雑誌で取り扱われていたということは、20世紀における名声によって、カウンシルの海外での活動が注目され、精査の対象となっていたことを表している可能性がある。次に、「ブリティッシュ・カウンシル (British Council)」、「英語の教え (English teaching)」および「英語を教える (English teaching)」というフレーズがほぼ平行して増加していたことは、これらの概念の強い結びつきを示していると考えられる。すなわちブリティッシュ・カウンシルが英語教育と関連付けられて認識されており、その出現頻度は英語教育の普及とともに増加していることが推測された。ただし、この3つのフレーズは同時期に増加していたわけではない。「British Council」の増加が見られるのは1950年代、「teaching English」や「English teaching」は1960年代で増加している。すなわちシンプルな相関関係ではなく、ブリティッシュ・カウンシルが英語教育を推奨したことにより、徐々に「teaching English」や「English teaching」のフレーズが使われるようになっていったことが推察される (図10, 11, 12)。とはいえこれら3つの単語はそれぞれにとっても類似したパターンで増加しており、関連性が示されている。これはブリティッシュ・カウンシルの言語教育との連想の度合いを示しているため重要であるだろう。このことは、ブリティッシュ・カウンシルの英語教育の奨励の成功を表しているのかもしれない。名目上は世界における英国文化の宣伝を目的とした半官半民組織であるが、この統計はカウンシルが英語の教育と学習に最も密接に関連する (及び歴史的に関連している) ことを示している。さらに、ブリティッシュ・カウンシルが人気のある雑誌において広く言及されているという事実は、それが英帝国崩壊後のソフトパワー普及媒体として成功していることを表している。

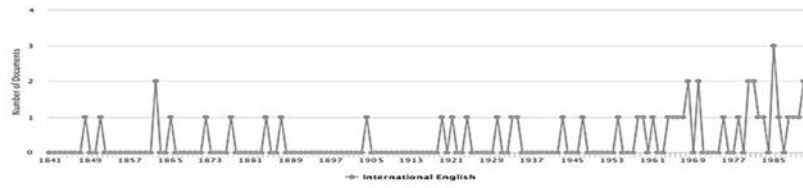


図 6 international English の出現頻度

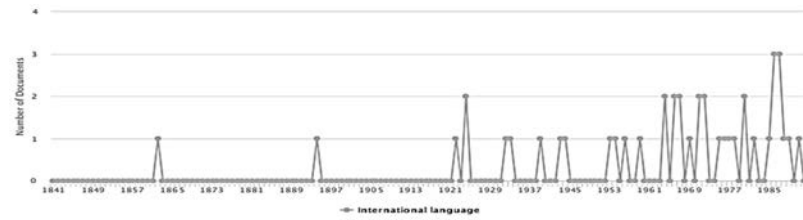


図 7 international Language の出現頻度

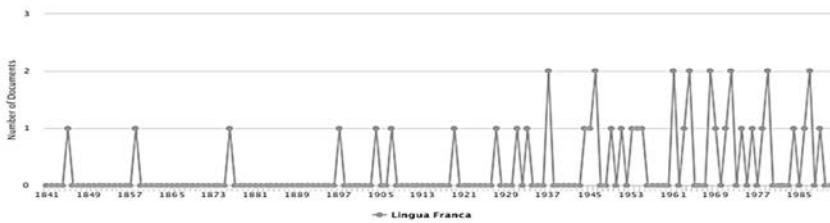


図 8 lingua franca の出現頻度



図 9 Indian English の出現頻度

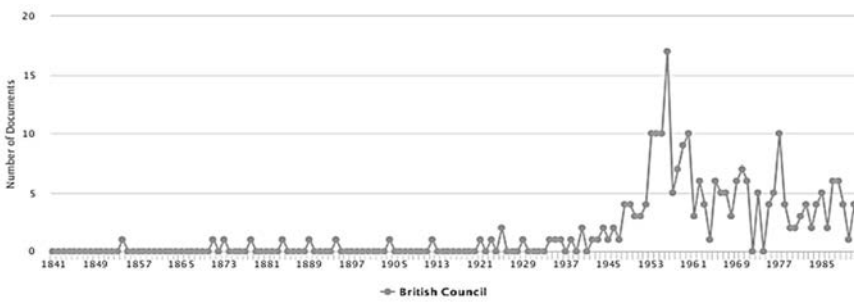


図 10 British Council の出現頻度

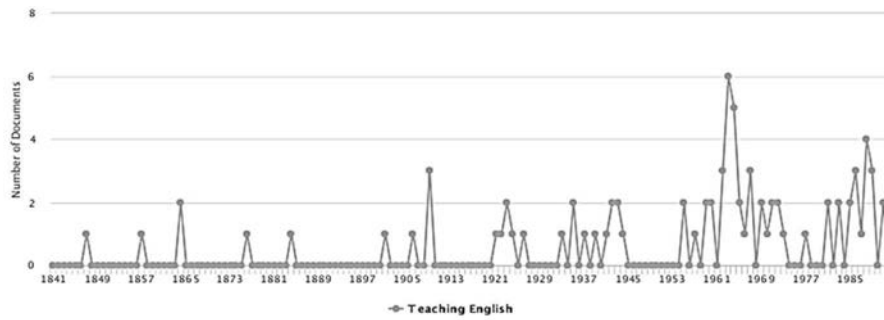


図 11 Teaching English の出現頻度

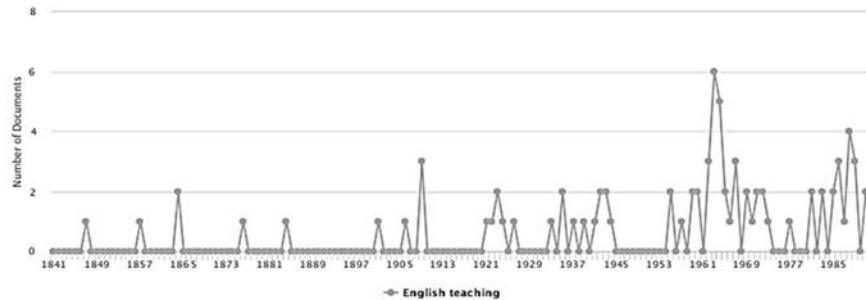


図 12 English teaching の出現頻度

S

4.3 結論

この *Punch* 誌のコーパスの統計分析は 2 つの興味深い点を論証している。第一に、学術文献以外では国際的言語としての英語の発展と普及は、ほとんど認識されていなかったことである。これは驚くべきことである。*Punch* が出版された期間は社会言語学的表展をした時期と重なり我々はこのことを認識していたはずである。しかしながら、*Punch* 誌の統計分析データはこのことを示していない。第二に、ブリティッシュ・カウンシルのような組織が、理論言語学および応用言語学の専門家の研究の外側で、英語の推奨と教育において世間の想像の中で連想されている、ということを示していることである。これらの結果は当然単なる相関であり、必ずしも因果関係を意味するとは捉えられない。しかし、これらの用語を歴史的及び社会政治学的な脈絡と合わせてみると、これらの用語の使用の傾向が、用語が関係付けられている歴史的発達を反映していることを理解することができる。これらの傾向が *Punch* のような非専門的出版物において観察できるという事実は、学会の外でのこれらの論説が一般大衆に対して広く行き渡っている本質を表しているため、特に興味深い。

5. 性格語の解析

続いて第三の分析分野として、心理学の視点から、*Punch* 誌において用いられている人の性格語の変遷について見てゆく。

人々の性格（パーソナリティ）傾向は、時代とともに変化していることが指摘されることがある。例えば近年では、「昔に比べて軟弱化している」「コミュニケーション力の低い人々が増えている」「キレやすい若者が多い」などの言説を耳にすることが多い。心理学における調査データにおいても、性格傾向が調査年によって変化していることは示されており、例えばアメリカにおいては近年になるほど自尊心や自己愛といった自己を肯定的に評価する性格特徴が上昇している一方で[23][24]、日本においては低下していったこと [25] などが報告されている。時代とともに人々の心のあり方が変化していったことは、実感としてもデータとしても示されているといえる。

しかしながらこうした変化は、人の本質が自然に変わったものとは考えにくい。人の心のあり方の変化は、その人が生きる社会に適応するために自らを発達させた結果であり、人が社会の中でどのように振る舞うかは、その社会でどのようなあり方が求められているかを反映している [26]。天沼[27]は、日本で起きた様々な災害や事件などに関するメディア記事において、以前は「頑張れ」というメッセージというのが多数であったのに対して、近年になるにつれて「頑張らない」ことに価値を置く表現が目立つことを指摘している。時代による人々の性格傾向の変化は、時代の中で注目され言語化される性格特徴が変化していくことに応じた適応の結果と考えるべきであろう。

そこで、*Punch* 誌の経年テキストデータを用いて、時代によって社会から注目される人の性格のあり方にどのよう

な変化がみられるかについて検討を行うこととした。具体的には、*Punch* 誌の記事の中で用いられている性格語の量的変化を検討した。変化を見る性格語としては、性格研究において広く用いられている性格語リストの中でも最も古いものを採用することが適当と考え、1960年代に作成された *The adjective check list manual* [28] の300語を用いた。

5.1 時代を反映する性格語

まず、300語の性格語のうち特定の時期に多く言及されるようになった語について見ていく(図13)。

今回の分析対象時期の初期である1920年において急増加が見られた性格語は“charming”であった。これはヨーロッパの芸術文化が盛んになりはじめた時代であり、ファッションを含めた文化のあり方において、軽く洗練されたものが価値を持っていたことを反映しているのかもしれない。

一方で1953年ごろから使用頻度が右上がりとなっているのが“sexy”である。第二次世界大戦後のイギリスにおいて、人物に対する評価として性的魅力が表現されるようになったことが読み取れる。杉村 [29]は、戦後においてイギリスの週刊誌「ウーマン」に描かれた女性像を概観し、「女性たちが活動の範囲を男性領域に広げていくにつれて、女性たちの性的魅力がより強調される」ことを考察している。*punch* で用いられていた“sexy”の語は、必ずしも女性に対する形容としてのみ用いられていたわけではないと考えられるものの、同様の傾向が示されていると言えるだろう。

その後、1960年に使用頻度が増加したのは“sophisticated”である。戦後の新しい時代にむけた人間観として、洗練されたあり方を目指すことが価値を持ったことがうかがえる。

また、ここまでは主に人の外見的特徴を示す性格語に特徴的变化が見られていたが、1960年ごろからは、内面的特徴を示す性格語にも変化が見られた。まず、このことに使用頻度が低下しているのが“anxious”である。不安や心配は、誰もが持ちうる自然な感情である。しかしながら、戦後のヨーロッパでは神経症における薬物療法が全盛となり、不安感情が病理として扱われるようになった。そのため、不安感情は通常の表現の中から抑圧されつつあったのかもしれない。

1985年ごろからは、“healthy”“mature”という語とともに、“frank”“friendly”という性格語の使用が増加していた。経済的に厳しい時代における当時のヨーロッパの人々のあり方として、健康で成熟した大人として、気取らずに、おおらかでオープンであることに価値を持っていた可能性が読み取れる。

以上を総じて、*Punch* で用いられた性格語の変遷から、戦前はかわいらしく(charming)であることが価値をもっていたのに対して戦後は性的魅力をもつ(sexy)ことが価値を持ち、不安感情(anxious)は抑圧され、洗練されたありかた(sophisticated)を求められ、その後の時代においては健康(healthy)で成熟した(mature)人間として、気取らずに(frank)でフレンドリー(friendly)なコミュニケーションをもつ、という人の心のあり方が社会の中で認められていたことが推察された。

5.2 求められる心の強さの変化

もう一点、社会の逆境の中での人の心のあり方の変遷を確認するために、“severe”と“tough”という語に注目したい。図14を見ると、過酷さや厳しさを意味する“severe”の語は全時代を通してコンスタントに出現しているのに対して、屈強であることを意味する“tough”の語は時代とともに徐々に増加していたことが示されている。このことは、人々に意識される厳しい状況や苦しさは時代を通してあまり変化がないにもかかわらず、そこで求められる心の強さは1980年代まで常に増加していったことを意味すると考えられる。すなわち、精神的に強靱であることの価値は第二次世界大戦を経て高まり続け、価値を持ったことが推察される。

なお、今回の対象である英国のデータとは異なるが、平野[29]は、1980年～2008年にアメリカ英語で書かれた書籍において使用された“toughness”の用語の変遷を概観し、“toughness”の用語が2001年のテロ以降低下し始めたことを報告している。その代わりに用いられるようになった用語は“resilience”であった。このように、社会において求められる心の強さのあり方は、戦争やテロといったその時代の大きなストレスイベントの種類によって影響を受けやすく、それがメディアの中で用いられる言葉にも反映されている可能性があると言えるだろう。

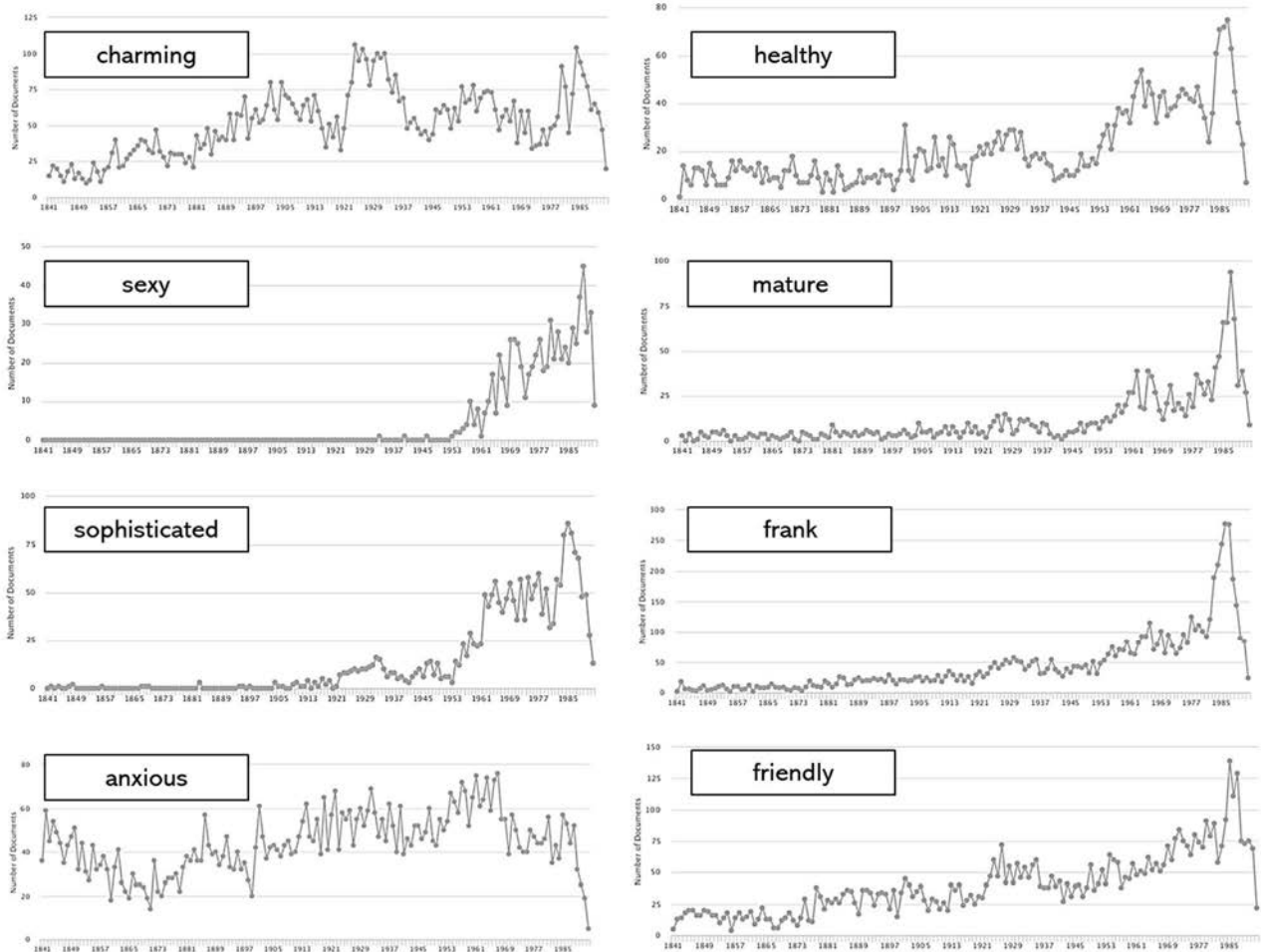


図 13 出現頻度の特徴的变化を持つ性格語 (1841 年～1985 年)

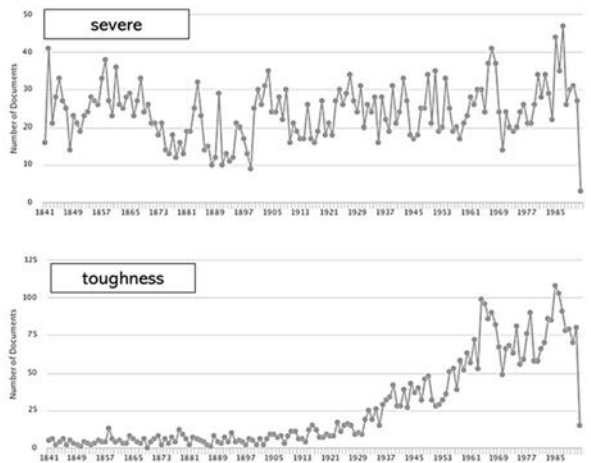


図 14 “severe”と”tough”の出現頻度の変遷

6. おわりに

この論文では、センゲージ社から提供された元データを処理して 152 年分のテキストデータを抜き出して統計処理を行い、服飾史と英語教育と心理学の 3 分野での分析を試みた。テキストの抽出のために C#言語による 300 行余りの

プログラミングが必要であった。また、KHCoder というソフトを使用した。この配下にデータベースソフトや図を描く R のコンパイラも用いられていた。KHCoder の基礎データを作成するときには 1 時間強の時間が必要であったが、これは配下のソフトを 64 ビットソフトにすればもう少し短時間で基礎データが得られたと思われる。

服飾史では、テキストマイニング専用ソフト KHcoder を用いて、23 年分のデータから各種グラフやデータを生成し、TF-IDF の計算値も得た。しかし、本格的な分析を行うには、根津 [8] の研究で行ったように、コンコーダンスソフトを用いて、服飾単語辞典内の単語を含む文章だけを抜き出して KHcoder を用いた解析を行う必要がある。

英語と言語教育の分野では頻出単語の年度変化を考察することにより、英語という言葉が世界にどのように使われているか、広まって行ったのかという疑問にある程度答えが得られた。

心理学の分野では、記事で用いられている性格語の経時間変化を比較することにより、時代によって社会から注目される人の性格のあり方がテキストマイニングによって読み取れることが確認された。

いずれの分野においても、単純な頻出単語の年度変化や生データのみの分析には限界があり、これらのデータを、

例えば、辞書に含まれる単語を含む文章のみを抜き出すと言うフィルターを加えたデータ作成が必要と考えられる。それには、人間がやるとやはり多大な時間が掛かるので、何らかの手段で自動的な抜き出しを可能にさせる工夫が必要である。今回は、152年分ものビッグデータが得られたことにより、このデータをどのようにして活用できるかを探る試みであった。最初の試みとしては、成功であったと感じる。

謝辞

我々は、データ作成にご協力頂いた、末吉璃子、大森恵理奈、丸山文葉の各氏に感謝する。また、熱心な議論・助力を頂いた根津夏菜子氏および名誉教授である能澤慧子氏にも感謝する。

参考文献

- 1) Christina Walkley, *The way to wear'em : 150 years of Punch on fashion*. P. Owen, London , 1985
- 2) 小池滋編, 『ヴィクトリアン・パンチ 図像資料で読む19世紀世界』, 1-7巻, 柏書房, 1995年
- 3) John Leech's pictures of life and character from the collection of "Mr. Punch", Bradbury, Agnew, & co., London, 1886
- 4) 学内公開されている雑誌 *Punch* へのリンク : <https://gdc.gale.com/gdc/artemis/?p=PNCH&u=tkaseiu>
- 5) 楊曉捷・小松和彦・荒木浩 編, 『デジタル人文学のすすめ』, 勉誠出版, 2013年 ; <https://ja.wikipedia.org/wiki/%E3%83%87%E3%82%B8%E3%82%BF%E3%83%AB%E3%83%BB%E3%83%92%E3%83%A5%E3%83%BC%E3%83%9E%E3%83%8B%E3%83%86%E3%82%A3%E3%83%BC%E3%82%BA>
- 6) 平井千津子, 松木孝幸, 新井哲男, 2015, 「統計処理の文学への応用—ヘミングウェイの場合」, 『東京家政大学研究紀要 第55集(2)』, pp.39-47.
- 7) 海和夏希, 2012, 「エロフ夫人の日記から見るマリー・アントワネットの時代の服飾」, 東京家政大学大学院家政学研究科被服造形学専攻・修士論文(未公刊).
- 8) 根津夏菜子, 2016, 「雑誌『*Punch*』の服飾史に関する統計的な分析」, 東京家政大学大学院家政学研究科被服造形学専攻・修士論文(未公刊).
- 9) Itoh, S. and Goto, N.: An Adaptive Noiseless Coding for Sources with Big Alphabet Size, *Trans. IEICE*, Vol.E74, No.9, pp.2495-2503 (1991).
- 10) Foley, J. D. et al.: *Computer Graphics - Principles and Practice*, System Programming Series, Addison-Wesley, Reading, Massachusetts, 2nd edition (1990).
- 11) 千葉則茂, 村岡一信: レイトレーシング CG 入門, *Information & Computing*, Vol.46, サイエンス社 (1990).
- 12) Cunnington, C. Willett: *English women's clothing in the nineteenth century*, London, 1969; Cunnington, C. Willett: *English women's clothing in the present century*, London, 1952; Cunnington, C. Willett and Phillis and Beard, Charles: *A dictionary of English costume 900-1900*, London, 1960.
- 13) Crystal, D. (1998). *English as a global language*. Cambridge: Cambridge University Press.
- 14) McKay, S. (2002). *Teaching English as an international language: Rethinking goals and approaches*. Oxford: Oxford University Press.
- 15) Jenkins, J. (2000). *The phonology of English as an international language: New models, new norms, new goals*. Oxford: Oxford University Press.
- 16) Kiczkowiak, M., & Lowe, R. J. (2018). *Teaching English as a lingua franca: The journey from EFL to ELF*. Stuttgart: DELTA Publishing.
- 17) Seidlhofer, B. (2011). *Understanding English as a Lingua Franca*. Oxford: Oxford University Press.
- 18) Pennycook, A. (1998). *English and the discourses of colonialism*. London: Routledge.
- 19) Sailaja, P. (2009). *Indian English*. Edinburgh: Edinburgh Univ. Press.
- 20) Setter, J., Wong, C. S. P., & Chan, B. H.-S. (2010). *Hong Kong English*. Edinburgh: Edinburgh University Press.
- 21) Deterding, D. (2010). *Singapore English* (Reprint). Edinburgh: Edinburgh University Press.
- 22) Howatt, A. P. R., & Widdowson, H. G. (2004). *A history of English language teaching* (2nd ed.). Oxford: Oxford University Press.
- 23) Phillipson, R. (1992). *Linguistic imperialism*. Oxford: Oxford University Press.
- 24) Twenge, J. M., & Campbell, W. K. (2001). Age and birth cohort differences in self-esteem: A cross-temporal meta-analysis. *Personality and Social Psychology Review*, 5(4), 321-344; Twenge, J. M., Konrath, S., Foster, J. D., Campbell, W. K., & Bushman, B. J. (2008). Egos inflating over time: A cross-temporal meta-analysis of the Narcissistic Personality Inventory. *Journal of Personality*, 76, 875-901.
- 25) 小塩真司・岡田 涼・茂垣まどか・並川 努・脇田貴文 (2014). 自尊感情平均値に及ぼす年齢と調査年の影響—Rosenbergの自尊感情尺度日本語版のメタ分析— *教育心理学研究* 62(4), 273-282.
- 26) Erikson, E. H. (1959). *Identity and the life cycle*. Selected papers.
- 27) 天沼 香 (2001). 時代相の変化とコア・パーソナリティ—「頑張る」日本人と「頑張らない」日本人— *東海女子大学紀要*, 21, 127-139.
- 28) 杉村使乃 (2014). 「グラマー」たちの第二次世界大戦: イギリスの週刊女性誌『ウーマン』表紙に見る女性表象 *人文社会科学研究所年報*, 12, 27-40.
- 29) 平野真理 (2019). 社員のレジリエンスで組織力を高める—メンバーの多様性がもたらす強靭さ— *りそな一丸*, 17(1), 7-10.