

# 『時事新報』初期の社説の著者推定

安形輝(亜細亜大学)\*

上田修一(慶應義塾大学)

\*agata@asia-u.ac.jp

## 1. 『時事新報』社説とその著者

### 1.1 初期の『時事新報』

『時事新報』は、明治15年(1882)に慶應義塾出版社から創刊された日刊紙である。慶應義塾出版社は、三田の慶應義塾の内の古い日本家屋にあり、印刷から発送までがこの建物内で行われていた。慶應義塾出版社は、明治16年(1883)10月に日本橋に移転し、翌年7月には、時事新報社と改称した。

福澤諭吉(1835 - 1905)は、明治政府の大隈重信、伊藤博文、井上馨らからの要請で政府系の新聞を発行しようとしていた<sup>1)</sup>が、「明治14年の政変」の結果、独自に中立派の新聞を発行することとした。「明治14年の政変」とは、明治14年(1881)10月、「薩長藩閥政府の専制的な体制を固め、天皇制立憲国家への道を確定した政変」であり、「その具体的内容としては、(1)開拓使官有物払下げの中止、(2)10年後の国会開設の公約、(3)参議大隈重信一派の追放などがあ」った<sup>2)</sup>。

大隈重信とともに下野した中上川彦次郎(1854-1901)を中心に、若手の慶應義塾卒業生が執筆、編集する『時事新報』の創刊号は明治15年3月1日に刊行された。タブロイド判4ページからなり、第一面は、一段分の題字と四段の記事部分からなり、一段は一行25字、二面から四面までは五段、一行23字だった。活字印刷で、創刊当初は輪転機ではなく、手引き印刷機を用いていた。記事は、官令、諸達伺公報、叙任賞[勤]、時

事新報、雑報、漫言、電報、外国雑報、広告などの見出しのもとに掲載されている。雑報がニュース記事であり、漫言は現在のコラムに近い。「時事新報」欄は社説で無署名である。時事新報は第一面から二面にかけて毎日掲載されており、長さは不定である。一号では終わらない社説もあり、中上川彦次郎執筆と推定される「日本帝國の海軍」は、8号にわたって分載されている。

### 1.2 『時事新報』社説の筆者

『福澤諭吉全集』には福澤諭吉の『時事新報』社説とされるものが収録されており、これらは石河幹明が選定したものとなっている。

近年、井田進也の指摘<sup>3)</sup>があり、『福澤諭吉全集』に収録された社説の執筆者が福澤諭吉であるかどうかについて疑問がもたれはじめた。井田は、全集の中で『時事新報』論説だけで九巻にもなるのは、人間わざとは思われない分量であり、「かりに本人が目を通しただけの作品が混入しているとしたら、やはり特殊な全集ということになりはしまいか」と述べている。

井田は、『中江兆民全集』の編纂に携わった際、自筆原稿が残っていない兆民の論説を新聞掲載論説そのものから何らかの指標をみつけて認定する作業を行った経験をもとに、『時事新報』社説の著者推定を目的として、福澤執筆社説の認定基準を作った。

この認定基準は署名記事における語彙や表現といった文体からの特徴を集め、無署名の社説と特徴を比較することで著者を推定するものである。さらに、特徴から添削の

程度を推定し、カテゴリーAからEまでの5段階評価を行うものである。

平山は井田の手法を井田メソッドとして応用し、井田の著者推定を検証するとともに、全集に含まれた他の論説に対しても著者推定を試みている<sup>4)</sup>。

### 1.3 本研究の目的

福澤諭吉全集に含まれる無署名社説は前述のように編者である石河が選定したものであるが、以下のような理由から福澤諭吉以外の筆者の社説が混入している可能性が大きい。

- (1) 外的証拠から福澤諭吉以外が執筆したと判明しているものがある
- (2) 入社以前の社説の執筆者について石河は判定できないはず

これらの疑問を解決するために著者推定が行われ、既往研究では井田メソッドが用いられてきた。しかし、井田メソッドによる判定は人手で行うため、判定の根拠となる語彙は一部の特定のものに限定されてしまう。他の記者による原稿の修正、植字時点でのミスおよび修正、校正を想定したときに、一部の表現に依存するこの手法は、客観性に乏しいといえる。そこで本研究では機械的な著者推定手法を用い、客観性が高く再現性のある形で『時事新報』社説に対する著者推定を試みる。

なお、『時事新報』社説の著者推定についての従来に関心は、ある特定の社説を福澤諭吉が書いたかどうか集中することが多い。また、この課題には次節1.4で述べる問題が存在するため、難度の高い課題となっている。そのため、ここでは福澤諭吉の著作であるか否かの判定を中心とした著者推定実験を行う。また、今回は石河入社以前の『時事新報』の初期の社説を実験対象とする。

### 1.4 『時事新報』社説著者推定の問題

・起草者、執筆者、添削者などの存在

掲載までに、著者以外の多くの関係者が関与するため、著者の文体が直接的に社説に反映されないことがあると考えられる。

・社説の形式と分量

文の長さ(句点によって判別できる)や読点の頻度は著者推定の有力な手掛かりの一つであるが、明治期の新聞として『時事新報』では句読点が打たれていない。また、後述の既往研究では著者推定に関して90%以上の精度を出すには20,000文字以上のデータ長が必要であったが、社説は2,000文字前後の長さのものが多く、

・福澤以外の真筆文献の少なさ

福澤諭吉以外の著者候補者については『時事新報』社説についての草稿など、真筆であるという外部的な証拠がほとんど現存しない。そのため、社説からだけでは十分な真筆文献を収集することが難しい。

## 2. 著者推定手法

実験に用いる著者推定手法としては、安形による圧縮改善率からの著者推定手法<sup>5)</sup>を用いる。これはBenedettoらの圧縮プログラムを応用した類似データ同定手法<sup>6)</sup>を修正したものである。

### 2.1 圧縮プログラムを用いたデータ同定

圧縮プログラムを応用した類似データの同定の基本的な考え方は、非常に単純なものである。二つのデータがあったときに、データ同士が類似していればしているほど、共通する冗長な部分が多くなると考えられる。そこで、ある二つのデータを連結する(二つのデータを単純に並置し、一つのファイルとする)ときに、圧縮プログラムがその連結データをより高い圧縮率で圧縮できるほど、つまり、生成される圧縮ファイルのサイズが小さ

くなればなるほど、その二つのデータは類似しているとみなす手法である。

圧縮改善係数による著者推定手法を近代日本文学者92作品の著者推定実験に用いた既往研究<sup>7)</sup>では最も性能が高い圧縮プログラムとの組み合わせでは99%以上の非常に高い著者推定精度が得られている。また、書簡や小説など形式の違いに左右されない手法であることが明らかとなった。

## 2.2 圧縮改善係数からの著者推定手法

この手法は以下の式に示される圧縮改善係数を用いる。

$$\text{圧縮改善係数} = 2 \cdot \frac{LZ_X + LZ_{A_i}}{L_{X+A_i}} - \frac{LZ_{X+A_i} + LZ_{A_i+X}}{L_{X+A_i}}$$

ここで、L はファイルサイズを示し、LX は基準データ X のファイルサイズを、LX+Ai は基準データ X と比較データ Ai を連結したファイルサイズを表している。LZ は圧縮ファイルのサイズを示しており、LZX は X を圧縮した場合のファイルサイズを、LZX+Ai は基準データ X を先に、比較データ Ai を後として連結した場合の圧縮ファイルサイズを、LZAi+X は逆に連結した場合の圧縮ファイルサイズをそれぞれ表している。

圧縮改善係数はデータを連結したときの圧縮されやすさがデータ単体と比較してどの程度改善されたかを示しており、この値が高ければ高いほど、類似度が高いことを意

味している。そのため、あるデータに対する類似度順の出力は、基準データと各比較データのすべての組み合わせについて圧縮改善係数を算出し、値が高いものから順に比較データを並べる手順となる。

この手法はどのような圧縮プログラムとも組み合わせることができるが、ここでは既往研究<sup>7)</sup>で性能が高かった 7-zip を PPMD に設定したものを組み合わせることとした。

## 3. 著者推定実験

『時事新報』の初期の社説を対象とした著者推定実験は大きく二つの段階に分け行っている。第一に真筆の判明している文献を対象とした著者推定実験を行い、このデータに対してどの程度の精度で著者推定が可能かを検証する。第二に真の著者が問題となっている「脱亜論」などの特定の社説を対象として著者推定実験を行う。

### 3.1 著者候補者の選定

『時事新報』で社説を執筆していた社説記者の在籍期間を示したのが図1である。この図において石川入社以前に在籍していたのは福澤諭吉、中上川彦次郎、高橋義雄、渡辺治、波多野承五郎である。

### 3.2 実験データの選定

上述の社説記者5人について福澤諭吉以外の4人については真筆が明らかな社説はほとんどない。そこで、社説ではなく雑誌な

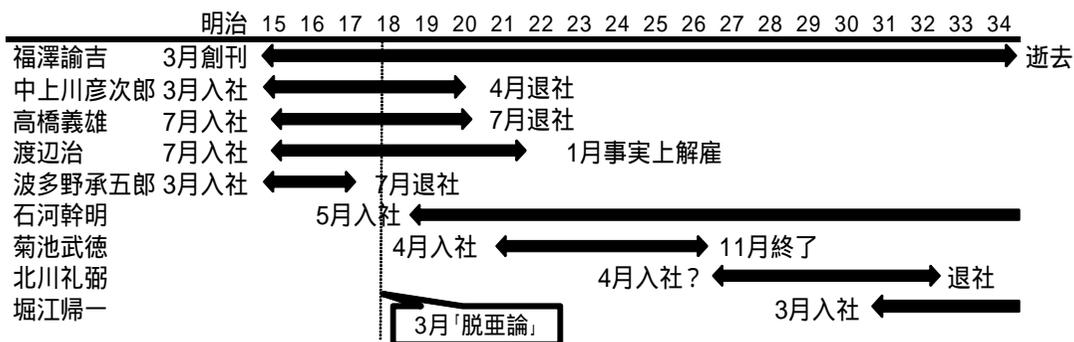


図1 『時事新報』社説記者の在籍期間

どの他の媒体に掲載された文献も含め、真筆の文献を選定した。(表1)

### 3.3 実験データの作成

著者推定実験を行うためにはテキストデータが必要であるが、『時事新報』社説のテキストデータは「脱亜論」のように有名なもの以外には存在しない。実験用テキストデータは図2のような手順で作成している。

### 3.3 実験結果

#### 3.3.1 真筆の明らかな文献に対する実験

表1の16文献に対する著者推定では、福澤か否かという形での評価では16文献中正解が15文献であった。不正解であった文献は高橋義雄の「国債論」だけである。福澤以外の各候補者についても著者推定を行う形で評価をした場合、16文献中10文献について正解であった。

#### 3.3.2 「脱亜論」の著者推定

「脱亜論」に対する著者推定においては波多野承五郎の4文献を除く12文献と「脱亜論」の合計13文献で実験を行うものとする。ここで波多野の文献を除く理由は、図1の社説記者の在籍期間からは『時事新報』に「脱亜論」が掲載された時に波多野承五郎は退社しているからである。

## 4. まとめ

本研究では既往研究と比較し、客観性の

表1 実験対象文献

著者	論題	情報源
福澤諭吉	政略	福澤諭吉全集
	医薬分業行はれ難し	福澤諭吉全集
渡辺治	倶楽部ノコト	交詢雑誌
	領事裁判治外法権を論ず	交詢雑誌
	日本国ノ船業	交詢雑誌
高橋義雄	国債論	交詢雑誌
	教育論	交詢雑誌
	人種改良の問・社員某の答	交詢雑誌
中上川彦次郎	遣清特派全権大使	福澤諭吉全集
	朝鮮事変の処分法	福澤諭吉全集
	人民教育ノ説	民間雑誌
	青砥左衛門ノ話	民間雑誌
波多野承五郎	独逸国の武勇	交詢雑誌
	僻地政党者の心得	交詢雑誌
	演説の禁止を解釈す	交詢雑誌
	水月鏡花	交詢雑誌

高い手法を用いて『時事新報』無署名社説に関する著者推定実験を行った。各社説の文字数が少ない点、必ずしも単独の著作でない点などから難度の高い課題であったが、少なくとも福澤か否かについては一定の精度での判定が可能であった。

#### 【引用文献】

- 1) 牧原憲夫. 民権と憲法. 岩波書店, 2006, 209p.(岩波新書 1043)
- 2) 田中彰. 明治14年の政変. 世界大百科事典 2006年改訂版. 平凡社, 2006.
- 3) 井田進也. 歴史とテキスト 西鶴から諭吉まで. 東京. 光芒社, 2001. 370p.
- 4) 平山洋. 福澤諭吉の真実. 東京. 文藝春秋. 2004, 244p.(文春新書 394)
- 5) 安形輝. 圧縮プログラムを応用した著者推定. Library and Information Science, no. 54, no.7, 2005, p.1-18.
- 6) Benedetto, Dario et al. Language trees and zipping. Physical Review Letters, vol.88, no.4, 2002, p.048702-1-048702-4.
- 7) 安形輝. 複数の圧縮プログラムを用いた著者推定実験. 2006年度日本図書館情報学会春季研究集会, 2006, 発表要綱 p.55-58

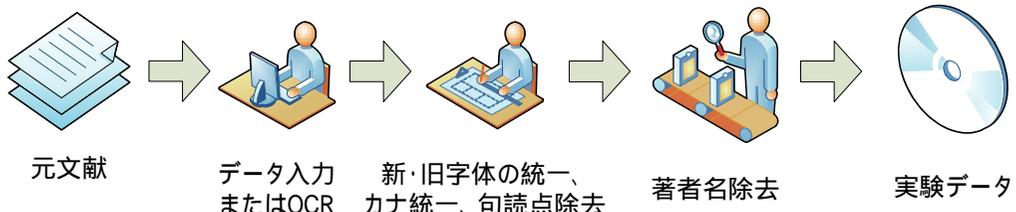


図2 実験データ作成手順