

国立国語研究所学術情報リポジトリ

『明六雑誌』の一人称代名詞

メタデータ	言語: jpn 出版者: 公開日: 2020-03-18 キーワード (Ja): キーワード (En): 作成者: メールアドレス: 所属:
URL	https://doi.org/10.15084/00002772

『明六雑誌』の一人称代名詞

近藤 明日子（国立国語研究所コーパス開発センター）¹

1. はじめに

日本の近代の言語資料のコーパス化とそれをういた近代語研究は今後一層の発展の期待される分野である。コーパス言語学的手法による近代語研究には、形態論情報の付与されたコーパスの開発が必須であるが、近代の文語論説文を対象とした形態素解析辞書「近代文語 UniDic」の開発により、その環境整備は飛躍的に進んだ。現在は、実際にその技術を用いた近代語の形態論情報付きコーパスの開発が始まっており、「近代語コーパス」プロジェクトにおいても、明治初期に刊行された『明六雑誌』の形態論情報付きコーパスである『明六雑誌コーパス』を開発した。

本稿は、この『明六雑誌コーパス』を用いて、そこに出現する一人称代名詞の分析を行うものである。用例の抽出や分析では、形態論情報をはじめとするコーパスに付与された情報を用い、コーパスの特長を活かした研究となることを目指す。そして、『明六雑誌』というほとんどが論説文よりなる資料を用いることで、当時の書き言葉的要素の強い資料における一人称代名詞の使用実態の一端を明らかにしたい。

2. 『明六雑誌コーパス』の概要

『明六雑誌コーパス』は、明治7(1874)年から明治8(1875)年にかけて刊行された、明六社の機関誌である『明六雑誌』の全文コーパスである。明六社は当時の洋学者によって結成された学術団体であり、そこで行われた演説や討論を広く一般に発表する媒体として『明六雑誌』は刊行された。よって、そこに掲載された記事はほとんどすべてが、ある物事について論じ解説する論説文となっている。

この『明六雑誌』に基づく『明六雑誌コーパス』は、本文テキストに書誌・文書構造・形態論・文字等に関する情報を付与する設計となっている。付与される情報のなかで特に注目されるのは形態論情報であろう。なぜなら、これまで形態論情報の付与された近代語のコーパスはほとんど例がなく、近藤・小木曾・加藤(2010)の『高等小学読本』コーパスといったものがわずかに存在するだけだからである。本コーパスの形態論情報は、『高等小学読本』コーパス同様、近代の文語論説文(明治普通文)を対象とする形態素解析辞書「近代文語 UniDic」を用いて本文を形態素解析した後、人手修正を加えたものが付与される。それにより、語の単位として揺れない斉一な単位である「短単位」(小椋・小磯・富士池・他、2011)を採用し、表記の揺れや語形の変異にかかわらない見出し語を付与した、日本語研究に適した構造を持つ情報となっている。

¹ kondo@ninjal.ac.jp

本コーパスに付与された形態論情報をはじめとする情報に基づき、コーパスの規模を概観すると、全 43 号に掲載された記事の総数は 155 記事、著者（翻訳者含む）は異なりで 16 名、延べ語数は約 18 万 3 千語（記号類を除く）となる²。

表 1 は、著者別に記事数を示したものである。これを見ると、記事数の多い上位 3 名（津田真道・西周・阪谷素）によって著された記事が計 74 記事と、全記事数の約半分を占めていることがわかる。本コーパスの分析から導き出される実態が、当時の論説文の一般的なありようではなく、特定の著者による個別的なありようである可能性があることになり、本コーパスを言語資料として扱う際には、そのことを十分に念頭に置いておく必要があるであろう。

さらに、記事の地の文の文体について見ると、全 155 記事のうち、文語文体の記事が 150 記事、口語文体の記事が 4 記事、文語口語混合文体の記事が 1 記事となっており、ほとんどが文語文体の記事で占められ、口語文体の記事はごくわずかしかない。文体の面でもデータに偏りがあることにも留意する必要がある。

表 1 著者別記事数

著者	記事数
津田真道	29
西周	25
阪谷素	20
杉亨二	13
森有礼	12
西村茂樹	11
中村正直	11
神田孝平	9
加藤弘之	8
箕作麟祥	5
柏原孝章	4
福沢諭吉	3
清水卯三郎	2
箕作秋坪	1
津田仙	1
柴田昌吉	1
合計	155

3. 分析対象とする語の抽出とその度数の概観

以下、この『明六雑誌コーパス』に出現する一人称代名詞の分析を行う。近代語の人称代名詞の研究は、これまで話し言葉の性質の強い資料（小説の会話部分、落語速記、口語文典など）を中心に行われてきた。よって、論説文といった書き言葉の性質の強い資料における実態は未だ明らかになっていない部分も多い。本稿の分析によりその実態の一端を明らかにしたい。

分析のためには、まず一人称代名詞の抽出が必要となるが、抽出作業は次にあげる手順でおこなった。

本コーパスの形態論情報を用い、品詞が代名詞となっている見出し語を抽出する。国語辞典等を参照し、 から一人称代名詞の可能性のある見出し語を選別する。一人称代名詞と関わりの深い見出し語として本コーパスでは連体詞となっている「わが」「おのが」を に追加する。

までの作業で得られた見出し語に属する用例について、文脈を確認し、実際に一人称代名詞として用いられているものを選別し分析対象とする。さらに、関連の深い用法として、人称にかかわらず対象それ自身を指す、いわゆる反射指示代名詞として用いられている用例も分析対象とした。

² 本稿に示すコーパスに基づく数値は 2011 年 12 月時点のデータに基づくものであり、公開中のコーパスに拠るものとは一部異なる場合がある。

この手順により、異なりで 15 語、延べで 1202 語の一人称代名詞および反射指示代名詞が抽出された。語ごとに記事の文体別の度数と表記の種類を表したものが表 2 である³。

表 2 表記の種類と記事文体別度数

語	表記の種類	度数			
		文語記事	口語記事	混在記事	合計
わが	我(474)、我ガ(67)、吾(19)、吾ガ(13)	538	27	8	573
よ	余(189)、予(6)	195	0	0	195
われ	我(127)、吾(20)、我レ(13)、予レ(1)	159	1	1	161
おのれ	己(101)、己レ(31)	129	3	0	132
ごじん	吾人(49)	49	0	0	49
ぼく	僕(17)	17	0	0	17
ごはい	吾輩(15)	15	0	0	15
わがはい	我輩(14)	14	0	0	14
せっしゃ	拙者(12)	0	12	0	12
ごせい	吾儕(10)	10	0	0	10
よはい	余輩(10)	10	0	0	10
それがし	某(4)、某シ(3)	4	0	3	7
わたくし	私(4)	0	4	0	4
よせい	余儕(2)	2	0	0	2
ぼくはい	僕輩(1)	1	0	0	1
合計		1143	47	12	1202

これを見ると、度数の多い上位 5 語「わが」「よ」「われ」「おのれ」「ごじん」の度数を合計すると 1110 語と全体の 90%以上を占め、これら 5 語が『明六雑誌』で主たる語であったことがわかる。

また、記事の文体別の度数を見ると、「せっしゃ」「わたくし」の 2 語はすべての用例が口語記事中出现しており、これらの語の話し言葉の性質の強さがうかがえる。しかしながら、2. で述べたように『明六雑誌』の口語記事はごくわずかであり、その少量のデータに基づいて、語と文体との対応関係を分析し、当時の口語文体の論説文における一人称代名詞および反射指示代名詞の実態について論じるには限界がある。よって、以後は文語記事中出现する語に限って分析を進めることとする。

4. 語と後続助詞との対応関係

³ 表 2 にあげられた表記の中には、読みの特定が困難なものもある。例えば、「我」「吾」一字の表記は、「わが」と読むのか「われ」と読むのか(それともそれ以外で読むのか)、はっきりしない場合がある。また、「吾輩」二字の表記は「ごはい」と読むのか「わがはい」と読むのか、断言することは難しい。そこで、「我」「吾」表記は、文脈から判断して「わが」「われ」いずれかに割り振り、それ以外の漢字表記はそれぞれ種類の読みで倒して度数を数えた。よって、例えば「吾輩」表記はすべて「ごはい」と見なし、「わがはい」として数えることはしなかった。また、「己」表記は、助詞「が」が後続する場合「おの」と読むことも多分に考えられるが、「己レガ」という「おのれ」+「が」とほぼ確定できる表記があったため、すべて「おのれ」と見なし、「おの」として数えることはしなかった。

また、『明六雑誌』では濁音を表記する仮名に濁点が用いられていない場合があり、「わが」の「が」も「カ」と表記されることがあるが、それらはすべて「ガ」に校訂した上で表記ごとの度数を数えた。

次にあげる表3は、文語記事中出现する語ごとに代名詞としての用法別度数を示したものである。

表3 代名詞用法別度数

語	一人称	反射指示	合計
わが	456	82	538
よ	195	0	195
われ	107	52	159
おのれ	0	129	129
ごじん	49	0	49
ぼく	17	0	17
ごはい	15	0	15
わがはい	14	0	14
ごせい	10	0	10
よはい	10	0	10
それがし	4	0	4
よせい	2	0	2
ぼくはい	1	0	1
合計	880	263	1143

ここから、一人称用法を持つ語は12語、反射指示用法を持つ語は3語あることがわかる。同じ用法を持つ語が複数存在する場合、内部でさらに何らかの使い分けがなされていると考えられるが、そうした語の間の違いについて探るため、後続する助詞・助動詞ごとに度数を示した表4を用い、コレスポンデンス分析を行った。コレスポンデンス分析は、データ表の行や列に含まれる情報を少数の成分(次元)に圧縮し、それらの関係を散布図上に布置することで、行カテゴリー間の関係、列カテゴリー間の関係、および行カテゴリーと列カテゴリー間の関係を視覚的に捉えることができる分析手法で、コーパス言語学においても活用範囲が広いとされるものである(石川・前田・山崎(編), 2010, pp.245-249)。

表4 後続助詞別度数

語	が体	ナシ	の体	が用	に	を	は	の用	と	も	より	てふ	なり	など	合計
わが一	444	0	0	12	0	0	0	0	0	0	0	0	0	0	456
わが反	75	0	0	7	0	0	0	0	0	0	0	0	0	0	82
よ	24	100	5	26	1	3	22	3	3	7	0	0	1	0	195
われ一	0	58	8	0	24	5	5	3	2	0	1	0	1	0	107
われ反	0	22	9	0	7	6	0	1	4	0	1	2	0	0	52
おのれ	50	14	31	4	11	16	0	2	0	0	1	0	0	0	129
ごじん	0	28	13	0	0	2	0	6	0	0	0	0	0	0	49
ぼく	1	13	0	2	0	0	0	1	0	0	0	0	0	0	17
ごはい	0	10	2	0	0	0	1	1	0	0	0	0	0	1	15
わがはい	0	10	1	0	0	0	0	3	0	0	0	0	0	0	14
ごせい	0	6	1	0	0	0	0	3	0	0	0	0	0	0	10
よはい	0	5	2	0	0	0	1	1	0	1	0	0	0	0	10
それがし	0	4	0	0	0	0	0	0	0	0	0	0	0	0	4
よせい	0	2	0	0	0	0	0	0	0	0	0	0	0	0	2
ぼくはい	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
合計	594	273	72	51	43	32	29	24	9	8	3	2	2	1	1143

表4では、一人称と反射指示の両方の用法を持つ「わが」「われ」については、用法ごとにカテゴリー化し、一人称用法のものを「わが一」「われ一」、反射指示用法のものを「わが反」「われ反」として示した。また、後続する助詞・助動詞のうち「が」「の」につい

では、後ろの体言にかかる連体用法をとるものと後ろの述語にかかる連用用法をとるものとを分けてカテゴリー化し、前者を「が_体」「の_体」、後者を「が_用」「の_用」として示した。助詞・助動詞の後続しないものについては「ナシ」としてカテゴリー化した。さらに、「わが」については、「わが」の「が」を後続する助詞と見なして度数をカウントした。

コレスポネンス分析に用いたのは表4全体ではなく、網掛けを施した部分である。「わが」については、そもそも後続の助詞・助動詞という観点からの分析にはそぐわないため、分析対象から外し、また、外れ値の影響を考慮して、合計の度数が10未満のカテゴリーについては分析対象から外したものである。分析には、統計分析パッケージRのMASSライブラリーのcorresp関数を用いた。

分析結果から、もっとも寄与率の高い第1次元(47.62%)と第2次元(29.94%)の得点を2次元空間上に布置したものが図1・図2で、図1は後続助詞の得点の散布図、図2は語の得点の散布図である。

まず、図1の第1次元を見ると、正の方向に「が_体」「を」「の_体」「に」が布置され、負の方向に「は」「の_用」「ナシ」「が_用」が布置されている。負の方向に布置される助詞群が受ける語は、多くの場合、述語に対し動作主や経験者といった意味的役割を担う⁴。一方、正の方向に布置される助詞群が受ける語は、述語に対し動作主や経験者といった意味的役割を担うことは「が_体」「の_体」の場合はもちろんなく、「を」「に」の場合も多くはない。つまり、第1次元は動作主や経験者といった意味的役割を担うか否かに基づくものであることになる。これを図2と対応させてみると、他の語から大きく離れて正の方向に布

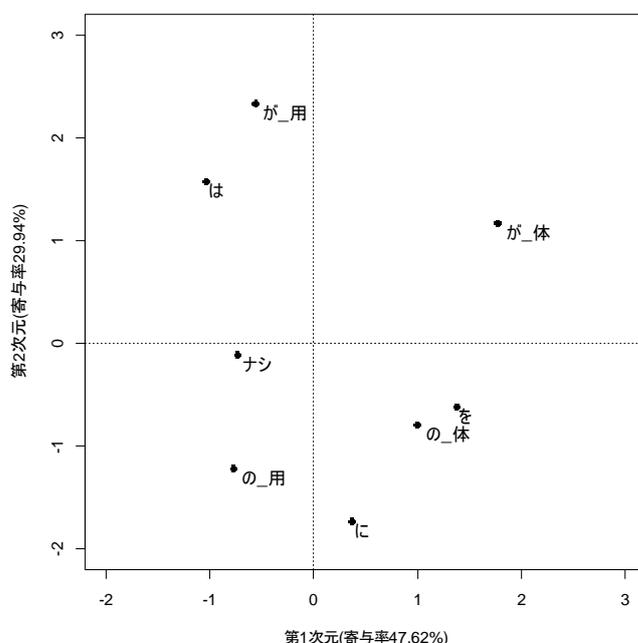


図1 後続助詞の散布図

⁴ 「は」の場合は相当する格助詞に置き換えた場合の意味的役割について言う。

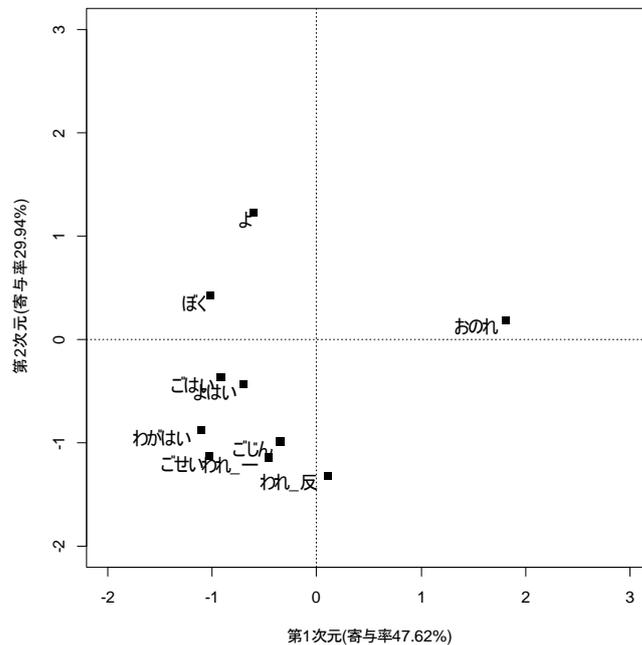


図2 語の散布図

置されている「おのれ」は、動作主や経験者といった意味的役割を担うことが少ないといった点で特徴付けられることになる。

次に、図1の第2次元を見ると、正の方向に「が_用」「は」「が_体」が布置され、負の方向に「に」「の_用」「の_体」「を」が布置されている。この軸の解釈は難しいところがあるが、正の方向に助詞「が」が、負の方向に助詞「の」が集まっている点には留意される。「が」「の」は人を表す体言をうける場合、待遇表現上の区別が認められ、「が」の用いられる場合はその人物に対する親愛・軽蔑・憎悪・卑下等の感情を伴い、「の」が用いられる場合には敬意あるいは心理的距離があると言われている。とすると、第2次元は待遇の程度に基づくものであることが考えられる。図2と対応させて見ると、正の方向に布置される「よ」「ぼく」は、負の方向に布置される「われ_反」「われ_一」「ごせい」「ごじん」「わがはい」「よはい」「ごはい」と比較して、相対的に待遇の程度が低いことになる。

以上のように、後続助詞と語との間には明らかな対応関係があり、それにより語は大きく次の3つのグループに分けることができると考えられる。

- A おのれ
- B よ・ぼく
- C われ_一・われ_反・ごじん・ごはい・わがはい・ごせい・よはい

このグループ分けと語の代名詞としての用法との関係を考えてみると、まずAグループの「おのれ」は反射指示の用法を専らとする点で他の語と区別される。Bグループの「よ」「ぼく」は一人称で、かつ書き手自身のみを指す単数用法を専らとする点で他の語と区別される。Cグループの内、「われ」を除いた「ごじん」「ごはい」「わがはい」「ごせい」「よはい」は

一人称で、かつ書き手だけでなく書き手を含めた複数の人を指す複数用法を取り得る点で他の語と区別される。本コーパスでの用例を見ると、「われ」を除く C グループの中で最も度数の多い「ごじん」は専ら複数用法をとり、「ごはい」「わがはい」「ごせい」「よはい」も複数用法が認められる。

このように、後続助詞との対応関係に基づく語のグループは、代名詞としての用法に基づく語の分類とほぼ一致することがわかる。

なお「われ」は、代名詞としての用法から見ると、一人称・反射指示両方の用法を持つ点で他の語とは区別されるが、後続助詞という観点からは「ごじん」等と同じ C グループに属する結果となった。「われ」については別の観点によるさらなる分析が必要であると言える。

5. 連体用法における語と被修飾体言との対応関係

次に 4. で分析の対象外とした「わが」について見てゆく。表 4 に示したように、「わが」は一人称・反射指示のどちらの用法でも連体用法をとることが多い。そこで、連体用法をとる「わが」および「の_体」「が_体」を伴う他の語について、被修飾体言との対応関係について検討し、語の間の違いについて見ていく。

表 5 は、各語が連体用法をとる場合の被修飾体言を示したものである。() 内は各体言の度数を示す。体言の種類が多い場合は、代表的な体言のみを示し以下は省略した(「...」で表記)。また「如し」にかかるものもここに含めて示してある。

表 5 連体用法における被修飾体言

語	が_体	の_体
わが_一	国(195)、帝国(40)、人民(17)、大日本帝国(13)、国内・地球・政府(7)、邦人(6)、国民・民・国産・社(5)、日本帝国・日本・心(4)、アジア・法律・今上天皇陛下・性・東州・東方(3)...	
わが_反	国(5)、身・父(4)、為・同生同人・同人・日本・自由・父母・用・物品・子・本体(2)...	
よ	ロジック・考・言(3)、胸臆・頭脳・所見(2)...	喜び・憶説・論・意・如し(1)
われ_一		有・文章・障子ガラス・義務・民... (1)
われ_反		如し・三法・下・精神・国・父... (1)
おのれ	力(5)、為・身体(3)、用・自由・三宝・意・身・一身・利・鋭利・労(2)...	意(3)、欲・如し(2)、迷信・子・力・胸中・責・権利・国... (1)
ごじん		為・性・心裏(2)、進歩・生活・感覚・天性... (1)
ぼく	論(1)	
ごはい		雲仍(2)
わがはい		目(1)
ごせい		如し(1)
よはい		首唱・鄙見(1)

ここから、語と被修飾体言との関係を見てゆく。

まず「わが」については、特に一人称用法の「わが」は、被修飾体言が「わが」の「わ」としての「所属先」という関係になる場合が多いということが言える。「わが」以外の語

では、被修飾体言は各語にとっての「所有物・所属物」という関係をとることが多いのとは対照的である。典型的なのは最も度数の多い「わが国」で、「わ」の所属する国」の意となる。「わが帝国」「わが地球」「わが社」「わがアジア」「わが東州」等も同様である。さらに、体言が「所属先の所有物・所属物」、特に「所属する国の所有物・所属物」という関係になることもある。例えば「わが人民」とは「わ」の所属する国に所属する人民」の意で用いられている（「わ」の統治する人民」「わ」の所有する人民」の意ではない）。「わが政府」「わが邦人」「わが国民」「わが民」「わが法律」等も同様の関係にある。このように、一人称用法の「わが」は、被修飾体言の関係が「所属先」「所属先の所有物・所属物」となる点で特徴付けられる。なお、反射指示用法の「わが」については、その被修飾体言が「所属先」「所属先の所有物・所属物」の関係となる割合は一人称用法のものほど高くない、「身」「父」「自由」「物品」等の「所有物・所属物」の関係となる場合も比較的多くなっている。

「わが」以外の語は、先に述べたように、被修飾体言が「所有物・所属物」の関係になることが多い。その中で、「よ」は被修飾体言が「ロジック」「考」「言」といった「所有する考え・意見」を意味する語で多く占められる点で特徴付けられる。「ぼく」「よはい」も被修飾体言に「論」や「首唱」「鄙見」をとり、「よ」と同様の傾向があるものと見られる。

以上のように、連体用法において語と被修飾体言との間にはいくつかの対応関係が見いだされることがわかった。

6. 主な語の特徴

以上の分析結果に基づき、主要な語についてそれぞれの特徴をまとめる。取り上げる語は文語記事での度数の多い上位5語「わが」「よ」「われ」「おのれ」「ごじん」である。

まず「わが」であるが、連体用法をとる主たる語であり、(1)(2)のように被修飾体言が「所属先」「所属先の所有物・所属物」の関係となる点で特徴的である。

(1) 夫レ我ガ國ノ文字先王始メ之ヲ漢土ニ取テ之ヲ用ウ(1号「洋字ヲ以テ国語ヲ書スルノ論」西周)⁵

(2) 目今諸省ニ於テ許多ノ洋人ヲ雇テ其學術ヲ傳取スル如ク彼尤善尤新ノ法教師ヲ雇テ公然我人民ヲ教導セシメバ奈何(3号「開化ヲ進ル方法ヲ論ズ」津田真道)

次に、「よ」であるが、一人称単数の用法をとる主たる語である。述語に対し動作主や経験者といった意味的役割を担い、(3)のような著者の個人的な体験を語る文脈でも用いられるが、論説文という文章の性質上、(4)(5)のように著者の意見や主張を述べる文脈で用いられることが多い。連体用法をとる場合も同様で、(6)のように著者の意見や主張を意味する語が被修飾体言となる。著者個人を指し示す語ゆえに、卑下の感情を伴う助詞「が」のほうが「の」よりも後続しやすい。

(3) 余會テ歐洲ニ遊テ煉火石造ノ家屋ヲ見ル(4号「煉火石造ノ説」西周)

⁵ 本文の引用に際しては、末尾の()内に号数・記事題名・著者名を示す。

- (4) 故ニ余敢テ謂フ我邦人倫ノ大本未ダ立ズト(8号「妻妾論ノ一」森有礼)
- (5) 余ハ思フニ政府ハ猶精神ノ如ク人民ハ猶軀骸ノ如クナリ(2号「学者職分論ノ評」津田真道)
- (6) 余ガ考ニハ狗ヲ連ルヨリモ兎ヲ輸入シテ錢ヲ取ラル、方逢ニ恐ル可シト思フ位ノコトナリ(26号「内地旅行西先生ノ説ヲ駁ス」福沢諭吉)

次に「おのれ」であるが、反射指示用法をとる主たる語である。(7)(8)のように連体用法をとることが多く、述語に対して動作主・経験者といった意味的役割を担うことは少ない。

- (7) 是皆個々人々日夜孜々汲々己ガ勞ヲ厭ハズ己ガ力ヲ盡シテ之ヲ求ムベキ者ニシテ(38号「人世三寶説(一)」西周)
- (8) 今日ニ至リテハ諸邦ノ君主タトヒ聰明衆ニ超タリトモ己ノ意ヲ以テ命令ヲ下スコトナシ(12号「西学一斑(前号ノ続)」中村正直)

次に「ごじん」であるが、一人称複数の用法をとる主たる語である。著者の個人的な意見について述べる文脈で用いられやすい「よ」とは異なり、(9)(10)のように、より一般性のある説や論を述べる文脈に用いられることが多い。また、著者自身のみならず他の人も含めて指し示す語ゆえに卑下の意味を伴う助詞「が」が後続することはない。

- (9) 想像ハ瞑目思想ノ間吾人觀見スル所ノ形象事歴ニシテ頗ル屢氣樓ト相類似ス(13号「想像論」津田真道)
- (10) 若夫レ吾人ノ性中情欲ヲ缺ク時ハ人類何ニ由テ生々蕃植スルコトヲ得ンヤ(34号「情欲論」津田真道)

最後に「われ」であるが、一人称・反射指示の両用法をとる語であり、一人称用法の「われ」は「よ」「ごじん」との違いを、反射指示用法の「われ」は「おのれ」との違いを明らかにしたいところである。

一人称用法の「われ」は、後続助詞との対応関係から「ごじん」と同じグループに属し、さらに(11)のように複数用法と思われる用例が見いだされ点でも「ごじん」と共通する。

- (11) 米利ノ戰艦一旦江戸海ニ侵入シ請求スルニ通信ノ約ヲ以ス是ニ於テ我之ヲ託シ始テ彼ニ日本來往ノ便ヲ得シム(7号「独立国権義」森有礼)

反射指示用法の「われ」は、(12)のように述語に対し動作主や経験者といった意味的役割を担うことが少なくなく、その点が「おのれ」とは異なる。

- (12) 自由ヲ伸シ羈絆ヲ脱シ租税ハ吾之ヲ増減スベシ官吏ハ吾之ヲ進退スベシ是人民ノ利ナリ(39号「政府与人民異利害論(六月一日演説)」西村茂樹)

本稿の分析からは、このような「われ」と他の語との類似点・相違点が指摘できるが、ではさらに進んで「ごじん」との違いはどのような点にあるのか、「おのれ」との違いは何によってもたらされるのかといったことについては、明らかにできなかった。今後の課題としたい。

7. おわりに

以上、『明六雑誌コーパス』を用いて分析を行い、当時の文語論説文における一人称代名

詞（および反射指示代名詞）について実態の解明を試みた。語と後続助詞、語と被修飾体言との間には明らかな対応関係が見いだされ、それにより一部ではあるが各語の特徴が明らかになった。今後は、別の観点からの分析を加え、一人称代名詞間の違いについてより詳細に考察したい。また、他のコーパスを用いて分析を行い、近代の一人称代名詞の通時的変遷についても考察する予定である。

文献

- 石川慎一郎、前田忠彦、山崎誠（編）（2010）『言語研究のための統計入門』、くろしお出版
小椋秀樹、小磯花絵、富士池優美、宮内佐夜香、小西光、原裕（2011）『特定領域研究「日本語コーパス」平成22年度研究成果報告書 『現代日本語書き言葉均衡コーパス』形態論情報規程集 第4版（上）（下）』
近藤明日子、小木曾智信、加藤文明子（2010）『『高等小学読本』の形態論情報付きコーパス』、情報処理学会シンポジウムシリーズ Vol.2010, No.15 人文科学とコンピュータシンポジウム論文集 人文工学の可能性～異分野融合による「実質化」の方法～、pp.189-194

関連 URL

近代文語 UniDic <http://www2.ninjal.ac.jp/lrc/index.php?UniDic>