様式C-19

科学研究費補助金研究成果報告書

平成　23年　5月　20日現在

機関番号：12501
研究種目：若手研究（B）
研究期間：2009～2010
課題番号：21720159
研究課題名（和文）明治前期東京語における口語資料の電子化テキスト整備とその資料性の検証に関する研究
研究課題名（英文）Building of machine readable texts of the Tokyo dialect in the early Meiji period with an Evaluation of their Referential Value.
研究代表者
岡部　嘉幸（OKABE YOSHIYUKI）
千葉大学・文学部・准教授
研究者番号：80292738

研究成果の概要（和文）：本研究課題では、近世後期江戸語から近代東京語への流れの中で資料整備の遅れている明治前期東京語の口語資料について、資料として適当と思われる資料を選定し、電子テキスト化を行った。電子テキスト化に際しては、原本による校正を行い、テキストの信頼性を高めた。さらに、それらの資料の資料性を確かめるための文法的指標についても検討を行い、いくつかの指標を設定することで、当該資料の資料性を検証した。

研究成果の概要（英文）：In this study, I made some Tokyo dialect texts in the early Meiji period machine readable texts, because there are few texts which are suitable for a study of Tokyo dialect in the early Meiji period. Furthermore, I proofread these machine readable texts by the originals. Then I set up some grammatical merkmals to examine whether these texts were suitable for a study of Tokyo dialect in the early Meiji period. On the basis of these merkmals, I evaluated the referential value of these texts.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
2009年度	2,400,000	720,000	3,120,000
2010年度	900,000	270,000	1,170,000
年度			
年度			
年度			
総　計	3,300,000	990,000	4,290,000

研究分野：日本語学
科研費の分科・細目：言語学・日本語学
キーワード：明治時代語、東京語、江戸語、電子化テキスト、資料性

１．研究開始当初の背景

本研究課題研究代表者（以下、研究代表者）は、岡部2000「江戸語における終止形承接のソウダについて」（国語と国文学』77-9）、岡部2002「江戸語におけるソウダとヨウダ」（『国語と国文学』79-10）、岡部2004「近世江戸語におけるラシイについて」（『近代語研究 12集』、武蔵野書院）等の成果を踏まえ、「近世後期から明治期にかけての江戸語・東京語のモダリティに関する研究」（平成19-20年度科学研究費補助金（若手研究B）、課題番号 19720104、研究代表者：岡部嘉幸）を行った。この研究において、研究代表者は、江戸後期から明治・大正期にかけてのモダリティ形式（ここでは広義の推量・推定の助動詞および当為判断の助動詞をこう呼ぶ）の意味・用法変化の記述とその動因の分析を行っているが、これと並行して、江戸語資料および東京語資料の整備も行っている。資料整備の面でいえば、江戸語の資料については質・

量ともにある程度のものが整備できた。また、明治後期・大正期東京語の資料としては、大量の小説類や雑誌『太陽』のコーパスが公刊されている（なお、申請者が研究分担者として参加した「コーパス言語学の方法に基づく言文一致現象の解析」（平成 18-19 年度科学研究費補助金（基盤研究Ｃ）、研究代表者：田中牧郎）は、このコーパスに基づく研究である）。しかし、明治前期の口語資料については、整備が遅れているのが現状であり、江戸語から東京語への文法的変化を記述しようとする際の一つの障害となっている。現在、資料としてよく利用されるのは、仮名垣魯文『西洋道中膝栗毛』（明治 3（1870）年－9（1876）年）、同『安愚楽鍋』（明治 4（1871）年）などであるが、同一作家による作品であり、作家による何らかの偏りが存在する恐れがある。また、量的にも十分とはいえず、魯文以外の「明治期戯作類」（江戸の人情本・滑稽本の系譜を引く、会話体の小説類を私にこう呼ぶ）や河竹黙阿弥の歌舞伎台本、三遊亭円朝らの落語速記本などの資料を拡充し、この時期の資料に作者・ジャンルの多様性を持たせる必要がある。
ところで、ある資料が、信頼できる明治前期東京語の資料たりうるのかを確かめるには、すでに信頼性がある程度確かめられている同時代資料から抽出された明治前期東京語の文法的特徴を指標とし、その指標が当該資料に現れるか否かを、実際にその資料を利用して検証する必要がある。明治前期東京語の文法的特徴は、松村明 1998『増補　江戸語東京語の研究』（東京堂出版）をはじめとする多くの先行研究によって指摘されているが、そのうちのどれが指標として有効なのかは検討の余地がある。さらに、その検討をふまえ、文法的観点からいえば、どの資料が明治前期東京語資料として適当なのかを検討する必要もある。

２．研究の目的
本研究課題は、以上のような背景から、以下のような目的をもつ。

（１）　明治前期東京語の口語資料として適当と思われる資料を選定し、電子テキスト化するとともに、そのテキストに校訂を加え、信頼性があり、かつ研究者が手軽に使えるテキストとして整備すること。
（２）　明治前期東京語の文法的特徴の中から選定資料の資料性を検証するのにふさわしい特徴が何であるかを検討すること。
（３）　（２）の検討をふまえ、選定資料の明治前期東京語資料としての資料性を検証すること。

３．研究の方法
　本研究課題は、明治前期東京語の資料整備という目的と、その資料の資料性の検討という目的をもっている。したがって、研究計画も、資料整備に関するものと、資料性の検討に関するものとにわけられる。
　資料整備の面に関しては、まず、研究代表者が資料の選定を行う。選定に際しては、先行研究を参照するとともに、国立国会図書館「近代デジタルライブラリー」を利用し、その本文を参照しながら、選定を行う。ついで、必要な各種タグを設定した上で、データ入力業者に入力を依頼し、それらの資料を、タグ付きの電子化テキストとしてデータ化を行う。さらに、電子データ化されたテキストについては、研究代表者や専門的知識をもった協力者（大学院生等）が原本との校訂作業を行う。
　資料性の検討の面に関しては、研究代表者が、資料性を検証するための文法的指標を、これまでの研究成果や先行研究を参照しながら、設定する。次に、それらの指標が個々の資料にどのような形で現れるかについて調査を行う。調査に際しては、構築した電子化テキストをパソコンで検索することにより、作業の効率化を図る。調査の結果を受けて、設定した指標のうちどの指標がより強力な指標となりうるのかを分析する。さらに、その分析に基づいて、それぞれの資料の明治前期口語資料としての資料性を判定するとともに、文法的観点からみた資料の位置付けを明らかにする。

４．研究成果
　本研究課題の研究成果は、以下のとおりである。

（１）明治前期東京語資料の整備と電子テキスト化

　明治前期東京語口語資料として適切と思われる資料を、国立国会図書館「近代デジタルライブラリー（http://kindai.ndl.go.jp/）を利用し、その本文を参照しながら選定し、明治期戯作を中心とする10作品の電子テキスト化を行った。電子テキスト化に際しては、原本ページ数タグ、タイトルタグ、割り注タグ、発話者タグ、狂歌・和歌・川柳タグ、手紙文タグなどの各種タグを付し、コンピュータによる検索および今後作成を予定しているXML文書化の利便性を図った。電子テキスト化を行った作品は以下のとおりである。

1.『東京娘風俗』初編（松村春輔編、由己社、明治15年）、2.『花街膝栗毛』初編（大久保常吉、九春社、明治16年）、3.『春色花暦』第１編（梅亭金鵞、秩山堂、明治16年）、4.『滑稽新話　書生肝粒誌』（菊

亭静、今古堂、明治16年）、5.『不二の白雪』第１編～３編（素娥自然、秩山堂、明治16年）、6.『明治流行嘘八百』第１、２編（菊亭静、績文舎、明治16年）、7.『滑稽笑談　清仏船栗毛』第１編～５編（伊東専三ほか、明治17-18年）、8.『貞節雪廼梅』（鶯亭昌安、日月堂、明治18年）、9.『妹と背かゞみ』第１号～１３号（春の舎おぼろ（坪内雄蔵）、会心書屋、明治18-19年）、10.『返咲園廼花』（鶯亭金升、鶴声社、明治19年）

さらに、原本を用いて、電子化テキストの校訂を行い、電子テキスト本文の信頼性を高めた。作成した電子化テキストは、関心のある研究者に配布するとともに、後日、web 上での公開を予定している。また、機会があれば、当該テキストの XML 文書化等も行いたい。

（２）資料性検証のための文法的指標候補の選定

先行研究などを参照し、（１）の電子化テキストの明治前期東京語資料としての資料性を検討するための文法的指標候補として、

（ア）打消しの助動詞の過去表現の形（たとえば、「行かなかった」と「行かなんだ」など）
（イ）動詞丁寧形の打消し過去表現の形（例えば、「行きませんでした」「行きませなんだ」など）
（ウ）可能表現形式（例えば、「行ける」「行かれる」「行けない」「行かれない」など）
（エ）「副詞「必ず」の否定表現との共起」現象（例えば、「必ずこのことを忘れるな」や「必ず行くまい」のような言い方が存在するか否か）

を設定し、明治前期東京語資料として先行研究において用いられていることの多い仮名垣魯文『西洋道中膝栗毛』（明治 3－9 年）、河竹黙阿弥『富士額男女繁山（女書生繁）』（歌舞伎台帳、明治 10 年）における使用状況を調査し、それぞれの項目の指標性を検討した。

項目（ア）については、『西洋道中』では「なんだ」の用例が２例、「なかった」の用例が３例（「なんだ」の使用率 40％）、『富士額』では、「なんだ」の用例が４例、「なかった」の用例が２例（「なんだ」の使用率約 67％）であった。ここから、打ち消しの助動詞の過去形として「なんだ」が使用されるということを、指標として設定できると判断した。具体的な用例は以下のとおりである。

・北「おらアさつぱり氣がつかなんだ」（『西洋道中』
・小助「今まで心附かなんだが、見ればお前は目も腫れて、どうやら悪い様子だね。」（『富士額』）
・北「ヲイ～～そりやアまつたくおれが知らなかつたからのことだ」（『西洋道中』）
・惣助「さうさ晩飯を喰つたツけか、喰はなかつたけか、ドンを聞かねえから分らねえ。」（『富士額』）

項目（イ）については、『西洋道中』では動詞丁寧形の打ち消し過去表現の用例が存在しなかった。『富士額』では、「ませなんだ」の用例が２例であり、「ませんかった」「ませんでした」の例は存在しなかった。確定的なことを述べるにはさらなる検討が必要であるが、動詞丁寧形の打ち消し過去表現として「ませなんだ」を用いるということを、一応の文法的指標の候補とすることができそうであることがわかった。具体的な用例は以下のとおりである。

・お虎「今夜爰へ御一緒に、泊らうとは存じませなんだが、嘸あなたお嬉しうござりませうね。」（『富士額』）

項目（ウ）については、様々な動詞語彙について検討を行う必要があるが、現段階では「行く」という動詞語彙（「行ける」「行かれる」「行けない」「行かれない」）のみの調査結果を報告する。『西洋道中』では、移動行為としての「行く」の可能を表す場合はすべて「行かれる」「行かれない」であった（４例、うち肯定２例、否定２例）。これは『富士額』でも同様であった（３例、すべて否定形）。ここから、少なくとも移動行為「行く」の可能を表すときに、「行かれる」「行かれない」の形が使用されるということを文法的指標とすることができると判断した。具体的な用例は以下のとおりである。

・弥「コウ北や間の悪イときはまがわるいもんだぜこの姿じやアどこへとツて行かれはしねへ」（『西洋道中』）
・お繁「衣類手道具何やかや、山ほどあれど出て行くに、まさか持つても行かれまい。」

項目（エ）については、従来論じられることが少なかったが、注目される現象だと思われる。『西洋道中』では、「肯定表現と共起するもの」が４例、「否定表現と共起するもの」が２例（否定との共起率約 33％）。『富士額』では「肯定表現と共起するもの」が７例、「否定表現と共起するもの」が８例（否定との共起率約 53％）であった。ここから、副詞「必ず」が否定表現と共起することを指標として

設定できると判断した。具体的な用例は以下のとおりである。

・通「イエサそりやアかならずいふべからず損は徳の基ダ」(『西洋道中』
・繁「いえなに、それと存じたら、必ず父は逢ひますまい。」(『富士額』)
・その後出入の節はかならず部屋の戸に錠をおろす（『西洋道中』）
・お繁「つい昔が出てならない、牛で一杯上げたいから、休暇に必ず出ておいでよ。」(『富士額』)

以上（ア）～（エ）の項目のうち、(イ）については、今後の検討を要するものと考えられるが、(ア)（ウ）（エ）については、文法的指標として認められるものと考える。ただし、これらの指標のうち、どの指標がより強力な指標となるのかについては、未だ明らかではなく、さらなる検討が必要である。

（３）電子化テキストの資料性の検証

（２）で設定した（ア）～（エ）の指標を用いて、(１）の電子化テキストの資料性を検証した。ただし、全体として、本研究課題で作成した個々の電子化テキストの言語量が少ないこと、また、設定した文法的指標の数が少ないことがあり、当該電子化テキストの十分な資料性の検討ができたとは言い難い。今後とも、文法的指標やその他の指標としてどのような現象がありうるかを検討し、さらなる資料性の検討を行うことが今後の課題である。ここでは、現時点での、中間報告的な検証結果を記すこととしたい。それぞれの項目の数字は確認された用例数を示す。また、作品名の後の括弧内の数字は、文法的指標を満たしている数を示す。

①『東京娘風俗』（1）
（ア）該当例なし
（イ）「ませんでした」1
（ウ）「行かれる」1
（エ）「肯定」1

②『花街膝栗毛』（0）
（ア）該当例なし
（イ）該当例なし
（ウ）該当例なし
（エ）該当例なし

③『春色花暦』（1）
（ア）「なんだ」2
（イ）「ませんだった」1
（ウ）該当例なし
（エ）該当例なし

④『滑稽新話　書生肝粒誌』（1）
（ア）「なんだ」3
（イ）該当例なし
（ウ）該当例なし
（エ）「肯定」3

⑤『不二の白雪』（2）
（ア）該当例なし
（イ）「ませんでした」1
（ウ）「行かれない」2
（エ）「肯定」1「否定」2

⑥『明治流行嘘八百』（1）
（ア）該当例なし
（イ）「ませんでした」2
（ウ）該当例なし
（エ）「肯定」11「否定」1

⑦『滑稽笑談　清仏船栗毛』（2）
（ア）「なんだ」5「なかった」4
（イ）「ませんでした」2
（ウ）「行かれない」2
（エ）「肯定」2

⑧『貞節雪廼梅』（0）
（ア）該当例なし
（イ）該当例なし
（ウ）該当例なし
（エ）「肯定」1

⑨『妹と背かゞみ』（2）
（ア）「なんだ」1「なかった」5
（イ）該当例なし
（ウ）該当例なし
（エ）「肯定」11「否定」4

⑩『返咲園廼花』（2）
（ア）「なんだ」4「なかった」1
（イ）「ませんだった」1
（ウ）該当例なし
（エ）「肯定」5「否定」1

すべての指標を満たす作品は存在しなかった。２つの指標を満たすものとしては、⑤『不二の白雪』、⑦『滑稽笑談　清仏船栗毛』、⑨『妹と背かゞみ』、⑩『返咲園廼花』があげられる。これらの作品を、現時点では、明治前期東京語資料としての資料性を満たすものとして認定しておきたい。ただし、最初にも述べたように、作品の言語量や文法的指標の数の不十分さは否定できず、十分な検証とはなっていない。この点は、今後の課題としたい。

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１件）

①岡部　嘉幸、「否定と共起する「必ず」について―近世後期江戸語を中心に―」、千葉大学『人文研究』第 40 号、千葉大学文学部、査読なし、2011 年、pp. 1-16

〔学会発表〕（計１件）

①岡部　嘉幸、「近世後期江戸語における副詞「必ず」について－否定と呼応する場合を中心に－」、国立国語研究所共同研究プロジェクト「通時コーパスの設計」研究発表会、2010 年 9 月 20 日、国立国語研究所

６．研究組織

(1)研究代表者
岡部　嘉幸（OKABE YOSHIYUK）
千葉大学・文学部・准教授
研究者番号：80292738