

令和 6 年 6 月 18 日現在

機関番号：62618

研究種目：若手研究

研究期間：2019～2023

課題番号：19K13211

研究課題名（和文）コーパスを用いた近世読本のルビと漢字表記の研究

研究課題名（英文）Study on ruby and kanji notation of Yomihon of the late Edo period by constructing a corpus

研究代表者

片山 久留美 (Katayama, Kurumi)

大学共同利用機関法人人間文化研究機構国立国語研究所・研究系・プロジェクト非常勤研究員

研究者番号：10803778

交付決定額（研究期間全体）：（直接経費） 2,800,000円

研究成果の概要（和文）：本研究では、江戸時代後期に流行した小説の一種である読本の表記上の特徴について明らかにすることを試みた。時代や作者の異なる読本の作品を複数選定してテキストデータ入力を行い、品詞や語種などの形態論情報付きのコーパスを構築した。このコーパスデータを用いて、読本全体の漢字使用の特徴や、ルビが表す語と本行の漢字が表す語との関係を品詞や語種などの観点から整理して明らかにした。また、ルビと本行の漢字とが特殊な対応をすることがある読本の特徴を生かしたコーパス構築の手法について検討を行った。

研究成果の学術的意義や社会的意義

従来、近代以前の資料においては使用する底本の問題などによりコーパスを用いた表記研究が行われにくい傾向があった。本研究では新たなコーパスデータ構築および国立国語研究所『日本語歴史コーパス江戸時代編』のデータの活用を行い、コーパスを用いた表記研究の一事例を提示することができた。また、これまでコーパス化が行われてこなかった近世読本のコーパス構築の端緒を開いたことで、今後の近世語研究・コーパス言語学の発展に資するものとなったと考える。

研究成果の概要（英文）：This study aims to elucidate the orthographic characteristics of "yomihon", a type of novel that was popular in the late Edo period. I selected multiple works of yomihon from different periods and authors, input their text data, and constructed a corpus with morphological information such as parts of speech and word types. Using this corpus data, I clarified the overall characteristics of kanji usage in yomihon and the relationship between the words represented by ruby and the words represented by kanji in the main text from the perspectives of parts of speech and word types. Additionally, I examined methods for constructing a corpus that takes advantage of the unique correspondence between ruby and kanji in yomihon.

研究分野：日本語学

キーワード：コーパス 表記 近世語 ルビ 日本語史

1. 研究開始当初の背景

江戸時代から明治初期にかけて、表記スタイルと文体・読者層には密接な関わりがあるとされる。田島(2017)では、漢字カタカナ交じり文が知識層を対象とした漢文訓読体で書かれ、漢字ひらがな交じり文は庶民を対象とした和文調で書かれているとしている。両者の違いは、漢字カタカナ交じり文では必要な箇所だけにふりがなを付すパラルビ表記を採用し、漢字ひらがな交じり文ではすべての語にルビが付される総ルビ表記であるという点にも及ぶ。屋名池(2009)ではパラルビ表記は漢文訓読体の、総ルビ表記は平安時代以来のひらがな文の末裔であるとし、両者は異なる表記システムであると述べている。このように、江戸時代から明治期にかけて文体と表記体系は不可分な関係を持っていることが指摘されてきた。

近世読本は、中国の白話小説に典拠を持ち、漢文脈の影響を強く受ける一方、通俗の小説として和語も多用する和漢混淆体で書かれている。そして、表記の面では同時代の他の戯作類よりも多くの漢字を含む漢字ひらがな交じり表記で書かれている。読本の表記に関する研究・文体に関する研究はそれぞれ多くの蓄積があるが、表記と文体とを関連づけた研究は管見の限り多くなかった。本研究では、文体的な和漢混淆の程度と表記における漢字使用の量を合わせて検討することで、読本の表記体系が和漢混淆文の歴史の中で、また明治期以降現代まで続いていく感じひらがな交じり表記成立の変遷の中でどのように位置づけられるのかを明らかにすることを目指したものである。

2. 研究の目的

本研究の目的は、読本の表記を文体および語彙的な特徴と関連づけてとらえ、読本の表記の特徴を明らかにすることである。これまで読本の表記についての研究は白話語彙に付されたルビなど表記体系の一部を扱ったものが多かったが、本研究では一部の語を取り上げるのではなく表記体系全体をとらえようとするところに特色がある。漢語が多用されているから漢字が多く使われているのか、和語は仮名で書かれているのか、白話語彙はどのような語を表記するのに用いられているのか。本研究では語ごとに品詞や語種などの形態論情報を付与したコーパスを作成し、本行の漢字表記語とルビとの対応関係を考察する。

3. 研究の方法

本研究では、近世読本の複数作品について形態論情報付きのコーパス構築を行った。具体的には、曲亭馬琴『南総里見八犬伝』『椿説弓張月』『月氷奇縁』、山東京伝『忠臣水滸伝』の4作品についてテキストデータを入力し、国立国語研究所『日本語歴史コーパス(CHJ)』の仕様を参考に、テキストデータに対するXMLタグ付与および形態素解析を行った。

作成したコーパスデータに基づいて作品全体の漢字使用の実態やルビと漢字表記語との対応関係を整理することとした。また、同時代の他の戯作類との表記上の特徴の違いを考察するため、CHJ「江戸時代編」のデータとの対照も行った。

4. 研究成果

(1) 読本コーパスの構築上の課題について

読本コーパスの構築にあたって、読本独自の課題とその解決策について検討を行った。

読本という資料の持つ特徴として、ルビが表す語と本行が表す語がまったくの別語である場合があること挙げられる。同様の問題は、CHJ「江戸時代編 洒落本」構築の際に形態論情報の多重化という解決策が見出されているが(村山ほか2017) 読本では特殊な対応をする箇所の数が多いこと、漢字表記語自体が白話語彙という従来のコーパスデータには見られない語彙を用いていることなどの新たな課題が存した。表記研究に耐えうるコーパスデータ構築を目指し、対応策について検討を重ね、その成果を学会発表の形で報告した(片山久留美「近世読本コーパスの構築と課題」「通時コーパス」シンポジウム2022、片山久留美「近世読本コーパスの設計と活用 表記研究での利用を目指して」言語資源ワークショップ2023 ほか)。

当初計画ではコーパスデータの外部公開を行う予定であったが、研究期間内に解決しきれなかった課題が残っているため、これらの課題を解消したうえで引き続きコーパスデータの外部公開を目指す。

(2) 読本の表記上の特徴についての検討

読本の表記上の特徴の一つとして、本行に現れる漢数字とそれに対応するルビとの関係について論文を執筆した(片山久留美(2022)「総ルビ表記の中の漢数字 『南総里見八犬伝』を中心に」『藝文研究』123号)。漢数字で表記された数詞において、接尾要素に漢語系統のルビが付されている場合には漢数字部分にルビがない場合が多く、和語系統の接尾要素の場合には漢数字も含めた全体にルビが付されていることなどを指摘した。数詞の表記は数詞以外の一般語とは異なる独自の表記システムを有している可能性があるため、まずその特徴を整理したものである。

読本コーパスデータの構築の進行とともに、そのデータを用いてルビと漢字表記との対応関

係の整理にも着手した。片山久留美「近世読本における複線表記のバリエーション」(「通時コーパス」シンポジウム 2024)では、大規模に用例収集することが可能な CHJ を活用した研究を行った。CHJ「江戸時代編 人情本」に現れるルビの付された漢字表記語を抽出し、ルビと本行の漢字表記が表す語の対応関係について、両者がほぼ同義で対応するもの、本行の漢字表記がルビの表す語の意味を限定するものなどといった語の意味の観点から整理を行った。こうした類型をもとに人情本と読本とを対照すると、人情本では特にルビの表す語と本行の漢字表記の表す語の意味の乖離が大きい例が見られやすいことを指摘した。読本においては、白話語彙など特殊な漢字表記語の使用が目立つものの、ルビと本行の語との意味の乖離という観点から見ると同義・類義のものが中心である可能性がある。ただし、こうした違いを人情本と読本というジャンル差として位置付けてよいかは、他の資料群も含めた慎重な検討が引き続き必要であると考えられる。

研究の進展に伴い、当初の目的として掲げた文体と表記との関係という観点よりも、上記で述べたようなルビと本行の漢字列との意味的な結びつきのバリエーションの検討が読本の表記、また総ルビ表記の本質を解明するために重要である可能性が明らかとなった。研究期間内の公刊はかなわなかったが、こうした観点から読本およびその他の近世・近代の総ルビ表記資料を対象とした漢字表記とルビとの対応関係の史的変遷を明らかにする博士論文を執筆中であり、本研究における成果を反映させる予定である。

【参考文献】田島優(2017)『「あて字」の日本語史』風媒社、村山実和子・中村壮範・小木曾智信(2017)「形態論情報の多重化による洒落本コーパスの質的拡張」『情報処理学会研究報告』Vol.2017-CH-114, No.8, 1-8)、屋名池誠(2009)「総ルビの時代」『文学』10巻6号 pp.117-130

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 片山久留美	4. 巻 123(1)
2. 論文標題 総ルビ表記の中の漢数字：『南総里見八犬伝』を中心に	5. 発行年 2022年
3. 雑誌名 藝文研究	6. 最初と最後の頁 178-193
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計4件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 片山久留美
2. 発表標題 近世読本コーパスの設計と活用 表記研究での利用を目指して
3. 学会等名 言語資源ワークショップ2023
4. 発表年 2023年

1. 発表者名 片山久留美
2. 発表標題 近世読本における複線表記のバリエーション
3. 学会等名 「通時コーパス」シンポジウム2024
4. 発表年 2024年

1. 発表者名 片山久留美
2. 発表標題 近世読本コーパス構築の課題 『忠臣水滸伝』を例に
3. 学会等名 「通時コーパス」シンポジウム2023
4. 発表年 2023年

1. 発表者名 片山久留美
2. 発表標題 近世読本コーパスの構築と課題
3. 学会等名 「通時コーパス」シンポジウム2022
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関