

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年5月31日現在

機関番号：62618

研究種目：基盤研究（C）

研究期間：2009～2011

課題番号：21520491

研究課題名（和文）「単語情報付きコーパス」を用いた近現代の語彙・語法史の研究

研究課題名（英文） A historical study of modern Japanese vocabulary and usage using "corpus with morphological annotations"

研究代表者 田中 牧郎（TANAKA MAKIRO）

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 言語資源研究系 准教授

研究者番号：90217076

研究成果の概要（和文）：

コーパス言語学の方法を近代語から現代語への史的研究に導入し、語彙史と語法史の研究に新しい領域を開拓した。まず、近代語テキストに対して、コンピューターを用いて自動的に単語に分割して読みや品詞情報を付与する形態素解析を行うことができるように、「近代文語 UniDic」を拡張した。次に、近代語の「単語情報付きコーパス」を作成し、近代語から現代語への語彙の変化や語法の変化について記述研究を展開した。

研究成果の概要（英文）：

We introduced the method of corpus linguistics into the historical study to contemporary Japanese from modern Japanese, and developed a new field to research of the history of vocabulary and usage. First, we extended "Kindai Bungo UniDic" which is the morphological analysis parser for modern Japanese texts. Next, we advanced descriptive researches about the history of vocabulary and usage, through constructing corpus with morphological annotations.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,400,000円	420,000円	1,820,000円
2010年度	1,000,000円	300,000円	1,300,000円
2011年度	900,000円	270,000円	1,170,000円
年度			
年度			
総計	3,300,000円	990,000円	4,290,000円

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：国語学 言語学 コーパス 近代語 現代語 語彙史 語法史

1. 研究開始当初の背景

コーパス言語学の方法を日本語学に適用する研究は増えてきたが、日本語史の分野では未開拓であった。これは、古典テキストに対して単語に区切り読みや品詞を付与する技術が未整備なことによると考えられた。そこで、近代語を例に、こうした技術を整備し、単語情報付きコーパスを構築し、近代から現

代への語彙史・語法史研究に貢献する例を示していく必要があった。

2. 研究の目的

近代を対象として「単語情報付きコーパス」を作成し、現代語のそれと対照するなどして、日本語史研究の分野に、コーパス言語学の方法を取り入れ、新しい研究領域を開拓

する。

3. 研究の方法

(1) 近代語テキストに対して、形態素解析(コンピューターで自動的に単語に区切り、読みや品詞等の情報を与えること)が実施できるような形態素解析辞書を整備する。

(2) 既存の電子テキストなどを利用して、近代語の「単語情報付きコーパス」を作成する。

(3) 近代語の単語情報付きコーパスを用いて、現代語の単語情報付きコーパスとも対照しながら、近代から現代への語彙史・語法史の記述を実施する。

4. 研究成果

(1) 近代語のテキストに対して、形態素解析に用いる辞書として開発した「近代文語 UniDic」を大幅に増補し、近代語テキストに対する形態素解析の精度を上げた。

(2) 国立国語研究所運営費交付金によるプロジェクトと共同で、近代語資料として著名な、明治初期の啓蒙雑誌『明六雑誌』の単語情報付きコーパスを作成した。このコーパスは、今後構築を進めていく近代語コーパスのモデルとして、2012年秋に公開する予定である。

(3) コーパスを用いた近現代の語彙史研究・語法史研究の分野に、新しい手法による研究成果を提供した。

① 語彙史研究

明治期に登場した新漢語の定着と淘汰についての研究を進め、和語との意味関係を形成しながら漢語が定着していく過程や淘汰されていく過程を、コーパスによる共起語分析を通して記述した。

② 語法史研究

明治初期の人称詞の用法の記述、および言文一致期に進行する文語法から口語法への交替の諸相の記述を、計量的方法で記述した。

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7件)

- ① 田中牧郎、新漢語定着の語彙的基盤—『太陽コーパス』の「実現」「表現」「出現」と「あらかず」「あられる」など—、日本語の学習と研究(日語学習と研究)(中国)、査読無、160、2012年、印刷中
- ② 小野正弘、文献学と語彙(第2部第7章)、これからの語彙論(斎藤倫明・石井正彦編、ひつじ書房)、査読無、2011年、pp.163-173
- ③ 小野正弘、国語史研究の可能性、文学・語学(全国大学国語国文学会)、査読無、

200号、2011年、pp.119-132

- ④ 小木曾智信、明治期国定『高等小学読本』の可能表現形式、成蹊國文、査読無 45、2011年 pp.72-84
- ⑤ 田中牧郎、近代漢語の定着—『太陽コーパス』に見る—、文学(岩波書店)、査読無、12巻3号、2011年、pp.136-153
- ⑥ 小木曾智信、明治大正期における補助動詞「去る」について、近代語研究(近代語学会)、査読無、15集、2010年、pp.444-465
- ⑦ 近藤明日子・小木曾智信・加藤文明子、『高等小学読本』の形態論情報付きコーパス、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～、査読有、2010年、pp.189-194

〔学会発表〕(計8件)

- ① 田中牧郎、新漢語定着の語彙的基盤—『太陽コーパス』を用いて—、漢字文化圏近代語研究会、2012年3月17日、浙江財経学院(中国)
- ② 近藤明日子、明治初期論説文における一人称代名詞の分析—『明六雑誌』コーパスを用いて—、第1回コーパス日本語学ワークショップ、2012年3月6日、国立国語研究所
- ③ 田中牧郎、日本語研究のためのコーパス—国立国語研究所のコーパスを中心に—、日本語学会(韓国)(招待講演)、2012年2月18日、慶北大学(韓国)
- ④ 田中牧郎、近代雑誌コーパスの作成、共立女子大学 公開集中講演会(招待講演)、2010年11月13日、共立女子大学
- ⑤ 田中牧郎、雑誌コーパスでとらえる明治・大正期の漢語の変動、国際学術研究会集 漢字漢語研究の新次元、2010年7月30日、国立国語研究所
- ⑥ 小野正弘、文の種類と数をかぞえる—近代語資料への一視点—、第274回近代語研究会、2010年6月26日、国立国語研究所
- ⑦ 小木曾智信、コーパスを利用した近代語研究—太陽コーパスと近代文語 UniDic—、東京外国語大学 GCOE プロジェクト講演会/語学研究所定例研究会、2009年12月2日、東京外国語大学
- ⑧ 近藤明日子・小木曾智信、形態素解析を用いた近代文語と現代語の語彙の比較、日本語学会 2009年度春季大会、2009年5月31日、武庫川女子大学

〔その他〕

ホームページ

「近代文語 UniDic」のページ：

<http://www2.ninjal.ac.jp/lrc/index.php?UniDic%2F%B6%E1%C2%E5%CA%B8%B8%ECUniDic>
<http://www2.ninjal.ac.jp/lrc/index.php?UniDic>

6. 研究組織

(1) 研究代表者

田中 牧郎 (TANAKA MAKIRO)

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・言語資源研究系・准教授
研究者番号：90217076

(2) 研究分担者

小木曾 智信 (OGISO TOSHINOBU)

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・言語資源研究系・准教授
研究者番号：20337489

小野 正弘 (ONO MASAHIRO)

明治大学・文学部・教授

研究者番号：90177270

近藤 明日子 (KONDO ASUKO)

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・コーパス開発センター・
プロジェクト奨励研究員

研究者番号：30425722

橋本 和佳 (HASHIMOTO WAKA)

同志社大学・文学部・講師

研究者番号：40411053

(3) 連携研究者

小椋 秀樹 (OGURA HIDEKI)

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・言語資源研究系・准教授
研究者番号：00321547