

国立国語研究所学術情報リポジトリ

〈著書紹介〉 近藤泰弘, 田中牧郎, 小木曾智信
編 『コーパスと日本語史研究』

メタデータ	言語: jpn 出版者: 公開日: 2016-03-16 キーワード (Ja): キーワード (En): 作成者: 小木曾, 智信 メールアドレス: 所属:
URL	https://doi.org/10.15084/00000832

近藤泰弘, 田中牧郎, 小木曾智信 編

『コーパスと日本語史研究』

ひつじ研究叢書(言語編)第127巻

2015年10月 ひつじ書房 A5判 293ページ 6,800円+税

小木曾 智信



1. 本書について

本書は2009年から行われてきた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」(プロジェクトリーダー:近藤泰弘,2012年より田中牧郎)と関連するプロジェクトの研究成果をまとめたもので,ひつじ研究叢書〈言語編〉第127巻として刊行された。同プロジェクトのメンバーのほか,お互いに連携して研究を行ってきたオックスフォード大学の「上代日本語コーパス」プロジェクト,国立情報学研究所の「ロボットは東大に入れるか」プロジェクトのメンバーを含む総勢14人により執筆されている。「通時コーパスの設計」プロジェクトは,当初計画していたコーパスの設計に留まらず,これに基づいたコーパスの試作にまで進み,2013年には『日本語歴史コーパス』としてその一部「平安時代編」の公開が開始された(プロジェクトとコーパス構築の経緯については,近藤2012,田中2014を参照)。本書はこの『日本語歴史コーパス』の概説にはじまり,コーパスを活用した日本語史研究の実践例,所外の関連プロジェクトの成果紹介,コーパスを活用した日本語史研究の文献リストと解題,さらにコーパスの形態論情報の解説までを含む。したがって本書は研究論文集であるとともにコーパス活用の概説書としての性格も有している。前近代のコーパスを活用した日本語史研究の専門書として初めてのものである。

2. 本書の構成

本書は10編の研究論文と付録とから構成されている。研究論文は以下の通りである。

『日本語歴史コーパス』と日本語史研究 近藤泰弘

通時コーパスによる言語の研究 山元啓史

平安期の〈名詞句+係助詞〉の格 その実態から見た係助詞の性質と副助詞との関連性
山田昌裕

受諾場面における形容詞使用の実態 中古語「よし」「やすし」の場合 高山善行

中古和文における接続表現について 岡崎友子

中古和文における文体別の特徴語 小木曾智信

『今昔物語集』に見る文体による語の対立 本朝仏法部と本朝世俗部の語彙比較
田中牧郎

人工頭脳プロジェクト「ロボットは東大に入れるか」国語試験古文問題解答に向けて

横野光

Why Romanize a Corpus of (Old) Japanese? Bjarke Frellesvig

The Oxford Corpus of Old Japanese Stephen Wright Horn and Kerri L. Russell

付録は、次の2編である。

コーパス日本語史研究目録 間淵洋子・鴻野知暁

『日本語歴史コーパス 平安時代編』の形態論情報 富士池優美

3. 本書の内容

10編の研究論文は、通時コーパスの問題点や方法論を扱う2編と、中古語の広い意味での文法を扱う3編と語彙の位相差を扱う2編、そして関連プロジェクトを扱う3編の順に配されている。冒頭の近藤論文は、国立国語研究所の『日本語歴史コーパス』の技術的背景や利用方法を紹介したうえで、コーパスを用いた日本語史研究を行うにあたっての問題点やノウハウについて論じている。この論文が本書で扱われているコーパスを活用した日本語史研究全体への導入となっている。続く山元論文は、通時コーパスを利用した言語研究を行う際の理論的な諸問題を論じたものである。

このあとは、『日本語歴史コーパス』を実際に活用した研究例である。山田論文は中古語の格と係助詞・副助詞との関係を記述している。高山論文は、受諾場面を例として、コーパスを活用した語用論的分析を行ったものである。岡崎論文は、中古語の接続詞を分類しコーパスによる使用状況を記述している。ここまでは広い意味での文法研究にコーパスを取り入れたものである。続く、小木曾論文はコーパスを用いて会話・地の文・和歌それぞれの特徴語を記述したものである。田中論文は、『今昔物語集』の語彙の位相差をコーパスに基づいて分析している。

さらに国語研究所以外で行われてきた関連する研究プロジェクトに関する論文が続く。横野論文は、「ロボットは東大に入れるか」プロジェクトで取り組んでいる自然言語処理技術を応用した古文の入試問題解答について論じている。Frellesvig論文は、氏が代表となって構築したオックスフォード「上代日本語コーパス」がローマ字の本文となっていることについてその意義・必要性を論じている。Horn & Russell論文は、同コーパスのアノテーションやデータの構成について、コーパス開発に直接携わってきた2氏が解説するものである。

続いて、通時コーパスに関連する付録2編を収録している。間淵・鴻野による一つ目は、コーパスや電子化資料を用いて行われてきた日本語史研究の文献リストとそのうちの一部の文献の解題からなる目録である。従来、時代別・分野別の研究文献リストが作られることはあってもコンピュータを用いた手法に着目してまとめられたものはなかった。コーパス・電子化資料を活用した日本語史研究のこれまでを一望できて有益である。富士池による二つ目の付録は、『日本語歴史コーパス』（平安時代編）に付与された形態論情報（単語の区切り方、品詞付与や見出し語化）の基準をまとめたものである。コーパスを活用するには、この形態論情報を十分に理解しておく必要がある。コーパス構築作業のために作られた大部の規程集から要所を抜き出し、利用にあたって押さえておくべき点をコンパクトにまとめている。

以上のように、本書は全体として、日本語史研究にコーパスを用いる際の方法論・具体例・参考文献・参考情報を一冊にまとめたものとなっており、コーパスを用いた日本語史研究の現在を知ることのできる書となっている。

●参考文献●

近藤泰弘(2012)「日本語通時コーパスの設計について」『国語研プロジェクトレビュー』3(2): 84-92.

田中牧郎(2014)「『日本語歴史コーパス』の構築」『日本語学』2014年11月臨時増刊号33(14): 56-67.

小木曾 智信 (おぎそ・としのぶ)

国立国語研究所言語資源研究系准教授。修士(文学)、博士(工学)。東京大学大学院人文社会系研究科博士課程単位取得満期退学。奈良先端科学技術大学院大学情報科学研究科博士課程修了。明海大学専任講師、独立行政法人国立国語研究所研究員を経て2009年10月より現職。コーパス開発センター、研究情報資料センター兼任。国語研究所で開発された『太陽コーパス』『現代日本語書き言葉均衡コーパス』『日本語歴史コーパス』の構築に携わる。

主な著書・論文:『雑誌『太陽』による確立期現代語の研究—『太陽コーパス』研究論文集—』(共著, 博文館新社, 2005), 『講座日本語コーパス2 書き言葉コーパス—設計と構築—』(共著, 朝倉書店, 2013).