

国立国語研究所学術情報リポジトリ

〈著書紹介〉 前川喜久雄 編/監修
辻井潤一・投野由紀夫・徳永健伸・丸山岳彦・山崎
誠・小木曾智信・中村壮範・山口昌也
著『コーパス入門』

メタデータ	言語: jpn 出版者: 公開日: 2015-10-30 キーワード (Ja): キーワード (En): 作成者: 前川, 喜久雄 メールアドレス: 所属:
URL	https://doi.org/10.15084/00000746

前川喜久雄 編/監修
辻井潤一・投野由紀夫・徳永健伸・丸山岳彦・山崎誠・小木曾智信・
中村壮範・山口昌也 著
『コーパス入門』
講座日本語コーパス 1
2013年7月 朝倉書店 A5判 196ページ 3,400円+税



前川 喜久雄

ここ数年、コーパスを用いた日本語研究の入門書が頻繁に出版されるようになった。これは世界的な潮流でもあるが、我が国の場合、日本語文法・社会言語学・日本語教育学などの入門書でありながら、コーパスとその利用法に1章をあてているものが少なくない。このような流行を生み出した、とまではいわずとも促進要因となったのが『現代日本語書き言葉均衡コーパス』の公開であることは衆目の一致するところだろう。

BCCWJと呼ばれる1億語規模のこのコーパスは日本語初の均衡コーパスであり、2006年に当時は独立行政法人であった国立国語研究所の研究開発部門言語資源研究グループによる開発が始まり、文部科学省科学研究費の特定領域研究「代表性を有する大規模日本語書き言葉コーパスの構築：21世紀の日本語研究の基盤整備」（略称「日本語コーパス」、2006～2010年度）の支援を得て2011年に公開を果たしたものである。

上記の特定領域研究の研究費は総額7.8億円にのぼり、いわゆる文科系のプロジェクトとしては従来最大規模のものであった。コーパスの開発だけでなく、コーパスを利用した様々な応用研究の可能性を探ることもプロジェクトの重要な目標であった。

今般、朝倉書店による刊行が始まった『講座日本語コーパス』はこのプロジェクトの成果を8巻の書籍に記録しようとするもので、本書がその第1巻にあたる。本書は6つの章と2編の付録から構成されている。章のタイトルと執筆者を紹介しておこう。氏名に[†]が付されているのは上述の特定領域研究のメンバー、*が付されているのは国立国語研究所の関係者である。

- 第1章「コーパスの存在意義」(*[†]前川)
- 第2章「コーパスと計算言語学」(辻井潤一)
- 第3章「言語教育とコーパス」([†]投野由紀夫)
- 第4章「言語処理とコーパス」([†]徳永健伸)
- 第5章「日本語コーパスの発展」(*[†]丸山岳彦)
- 第6章「語彙調査の系譜とコーパス」(*[†]山崎誠)
- 付録A「コーパス検索ツール(1)『中納言』の使い方」(*[†]小木曾智信・*中村壮範)
- 付録B「コーパス検索ツール(2)全文検索システム『ひまわり』」(*[†]山口昌也)

本書には従来の類書にはみられない3つの特徴がある。まずタイトルが示すようにコーパ

スの入門書であって、コーパス言語学の入門書ではないこと。第二に、コーパス利用のテクニカルな問題よりもむしろコーパスの存在が要請される所以を論じていること。第三に、付録の形でコーパス分析のテクニックにも配慮していることである。

講座の監修者として、このような1冊を講座の冒頭に据えることが必要と考えたのは、コーパスの流行とはうらはらに、言語研究者中にコーパスの存在意義に対する無理解が横行していると感じるからである。コーパスは何のためにあるかと聞かれて用例収集のためとしか答えられない言語学者や国語学者には是非本書を読んでいただきたい。

前川 喜久雄 (まえかわ・きくお)

国立国語研究所言語資源研究系教授，系長，コーパス開発センター長。博士（学術）（東京工業大学）。

主な著書・論文： *A frequency dictionary of Japanese*（共著，Routledge, 2013）， *Prominence marking in the Japanese intonation system*（共著， *The Oxford handbook of Japanese linguistics*. Oxford University Press, 2008）， *Coarticulatory reinterpretation of allophonic variation: Corpus-based analysis of /z/ in spontaneous Japanese* (*Journal of Phonetics* 38 (3), 2010).

社会活動：一橋大学大学院連携教授， Editorial board member of *Phonetica*.