

書評

荻野綱男 (2014) 『ウェブ検索による日本語研究』

朝倉書店

金水 敏 (大阪大学大学院)

ウェブ検索はもはや日常生活にとってなくてはならないツールとなりつつあるが、言語研究者にとっても魅力的なデータを提供してくれる。書評者も、ちょっとした調査(語形のゆれや文法の異例など)をしたいときに、ウェブの検索エンジンを使うことがあるが、出てくるページの基準やヒット数の意味など、よく分からないために、結果を十分生かし切ることができないという不満があった。そんな状況に対して、心強い指針を与えてくれるのが本書である。

「まえがき」によれば、本書は著者が勤務校で担当している授業のための教科書であるが、解説を書き加えることによって、ウェブで日本語研究を行いたい方のためにも役立つようにしたとのことである。目次を掲げておく。

目次

第1部 入門編

第1章 WWW の性格と日本語研究への応用 第2章 WWW 検索のヒット件数
第3章 検索エンジン 第4章 WWW 中の表現の間違い 第5章 WWW 検索の方法論

第2部 基礎編

第1章 複合語の認識とフレーズ検索 第2章 AND 検索 第3章 OR 検索 第4章 NOT 検索 (マイナス検索) 第5章 ワイルドカード検索 第6章 活用形の検索 第7章 ブログ検索 第8章 WWW の記事の誤り 第9章 WWW でのヒット件数の意味

第3部 問題解決編

第1章 「いらっしゃる」の意味 第2章 [ei] の表記のゆれ 第3章 外来語の「キ」と「ク」の違い 第4章 外来語の「ア」と「ヤ」 第5章 「dis られる」の意味 第6章 漢字語の読み方 第7章 WWW による問題解決の課題例

第4部 研究編

第1章 WWW 検索による方言語形の全国分布調査 第2章 WWW 検索と方言辞典の記載内容の確認 第3章 ブログに見る日本語の男女差 第4章 形容動詞連体形の「な／の」の選択 第5章 外来語の語形のゆれ (1) — 「チック」と「ティック」— 第6章 外来語の語形のゆれ (2) — 「バ」と「ヴァ」—

第5部 応用編

第1章 レポートの課題 第2章 「亡くなる」は動物に使わないのか 第3章 ニセ方言の使用状況 第4章 犬に「ちゃん」は付けない 第5章 「こんな」類と

「こういう」類の位置 名詞との共起 第 6 章 ゼッケンからナンバーカードへの変遷 第 7 章 レポート執筆の意味

文献索引

第 1 部, 第 2 部の内容については見ての通りであるが, 第 3 部は, 著者が授業中に受講者に課した宿題とその結果をまとめたものである。第 4 部は, 著者が書いた論文などを収録したもので, 「一読しておく」と学生の卒業論文などにも役に立つのではないかとのことである。第 5 部は, 著者が指導していた学生が提出したレポートを, 著者が全面的に書き直したものである。レポートの書き方としても参考になる。

第 1 部, 第 2 部の内容は, 私のような, ウェブ検索初心者にとって大変貴重な情報の宝庫である。例えば, 3.2.3 「日本語研究用検索エンジンは何がいいか」という, ずばり知りたいテーマが掲げられている。WWW 用の検索エンジンは, なるべく広い範囲の記事を検索できるようになってほしい。この点では Google に軍配があがる。しかし, Google には検索件数の不安定さという致命的な問題がある。そこであれこれテストを重ねた結果, 「ヒット件数のゆれを基準に考える場合は, 現状でも, goo がベストである。」(24 頁) という結論に至った。こういう知識は本当にありがたい。

あと, 案外勘違いしがちなのは, 「検索エンジンで検索対象となるのは「単語」ないし「単語列」であり, 「文字」ないし「文字列」ではない」という事実である (15 頁)。つまり, 単語より小さな単位では検索ができない。また, 活用語は, 学校文法に従って分解される。「通り抜けられない」は「通り／抜け／られ／ない」と分解され, それぞれの「単語」の AND 検索と同じ意味になる。フレーズ検索 (“ ”で単語列を囲む), OR 検索, NOT 検索 (マイナス検索) 等, ワイルドカード検索の, 日常ではあまり使わない技法も, 日本語研究にとっては心強いツールとなるだろう。

第 3 部以降のケーススタディは, 自分自身の研究に役立つのみならず, 授業でのレポート課題や卒論指導などで明日からでも応用可能な, 本当に涙が出るくらいうれしいコンテンツである。特に第 5 部は, 私にもなじみ深い先行研究の結論について, 学生にウェブ検索で検証させたものを, 著者が再検証しているもので, 興味深かった。例えば第 3 章「ニセ方言の使用状況」では, 田中ゆかり (2011) 『「方言コスプレ」の時代—ニセ関西弁から龍馬語まで—』(岩波書店) で提示されたニセ方言を goo ブログ検索で調査するというものである。ニセ関西弁から「～やん」「～やろ」「なんでやねん」を, ニセ東北弁からは「～だべ」「～っぺか」「～んだども」を, ニセ九州弁からは「～でごわす」を, ニセ中国方言からは「～じゃけん」を代表として調べている。全ヒット件数を確認するとともに, そのうちの先頭 100 件を 1 例ずつ見ていき, 方言としての使用か否かを確認した。次に, 方言としての使用の中から本方言, ジモ方言 (地元で使われているものの普段は使わない方言) を除外してニセ方言の例数を数えた。その結果, 「～やん」「～やろ」「～じゃけん」はニセ方言としての使用率が低いのにに対し, 「なんでやねん」「だべ」「～でごわす」はニセ方言率が高かった。田中の原典と併せて読めば, 意味深い結果である。

このようなレポートを著者は 10 年ほど学生に課しているとのことであるが, その意義として, WWW がそのような研究結果の検証に使えることがわかることと, 論文の読み

方が変わるということを上げている。特に 2 番目の点は書評者にとって重要と感じられた。論文を「先生」が書いた「正しい」ものとして受け取るのではなく、学部生であっても、批判的な態度で再検証を試みることができる。その経験は、学生が自分で論文を書く際に、直接的に生きてくる貴重な体験となるように思われる。「WWW 検索は、まさにこのように先行研究を検証し、批判し、さらに新しいアイデアを生み出すために有効なツールになるものである」(184 頁)とする著者の主張は深くうなずける。

(参考までに)

ISBN 978-4-254-51044-7 C 3081

B5 判 vii + 193 頁

(2015 年 9 月 24 日受付)