

『日本語歴史コーパス』の増補に伴う第1章の記述に関する注意

『コーパスで学ぶ日本語学 日本語の歴史』の資料である『日本語歴史コーパス』が、2020年3月末に増補されました。本書のp.3の表1.1に示している、「『日本語歴史コーパス』の構成（2019年7月現在）」を、増補された資料を含めて改訂すると、次のようになります。赤字部分が、増補部分にあたります。

『日本語歴史コーパス』の構成（2020年4月現在）			
シリーズ	資料の成立年	主な資料	分量
奈良時代編Ⅰ万葉集	759年	万葉集	約10万語
奈良時代編Ⅱ宣命	797年	續日本紀宣命	約2万語
平安時代編	900-1110年	竹取物語、源氏物語、大鏡など	約86万語
和歌集編	905-1205年	古今和歌集、新古今和歌集など	約27万語
鎌倉時代編Ⅰ説話・隨筆	1100-1336年	今昔物語集、徒然草など	約95万語
鎌倉時代編Ⅱ日記・紀行	1223-1306年	海道記、とはずがたりなど	約23万語
室町時代編Ⅰ狂言	1642年	虎明狂言集	約23万語
室町時代編Ⅱキリストン資料	1592-1593年	天草版平家物語、同伊曾保物語など	約14万語
江戸時代編Ⅰ洒落本	1757-1836年	聖遊廓、花街寿々女など	約20万語
江戸時代編Ⅱ人情本	1821-1864年	比翼連理花酒志満台など	約41万語
江戸時代編Ⅲ近松淨瑠璃	1703-1722年	曾根崎心中、心中宵庚申など	約25万語
明治大正編Ⅰ雑誌	1874-1925年	明六雑誌、太陽、婦人俱楽部など	約1253万語
明治大正編Ⅱ教科書	1904-1947年	国定読本Ⅰ期～VI期など	約71万語
明治大正編Ⅲ明治初期口語資料	1869-1879年	安愚樂鍋、百一新論など	約20万語

この増補によって、第1章の「4. 表記の変遷の調査の例—「話す」の表記—」(p.8~12)に記してある通りに、このコーパスを検索すると、本書の記載とは異なる結果が得られ、「話す」の表記の正しい変遷が、たどれなくなりました。その理由は、増補の結果ヒットすることになった「江戸時代編Ⅲ近松淨瑠璃」の「話す」が、原本の表記を改めたものであることによります。近松淨瑠璃は、『新編日本古典文学全集』(小学館)を底本としていますが、この底本は、P.11のコラム「表記調査とコーパスの本文」に記すように、原本の表記を改めるところがあり、表記調査の対象にするには不適切だからです。

本書の記述の通りの検索結果を得るには、p.8に記す検索を行う際に、「中納言」の画面で、「検索対象の選択」をクリックし、「全てにチェックを入れる」を選択し、さらに、近松淨瑠璃として挙げてある作品名のチェックを、すべて外してください。そうすることで、本書の記述と同じ検索結果を得ることができます。