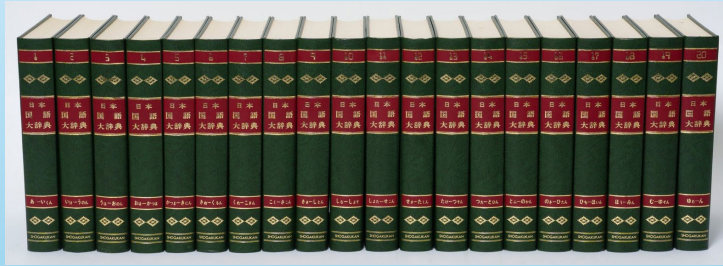
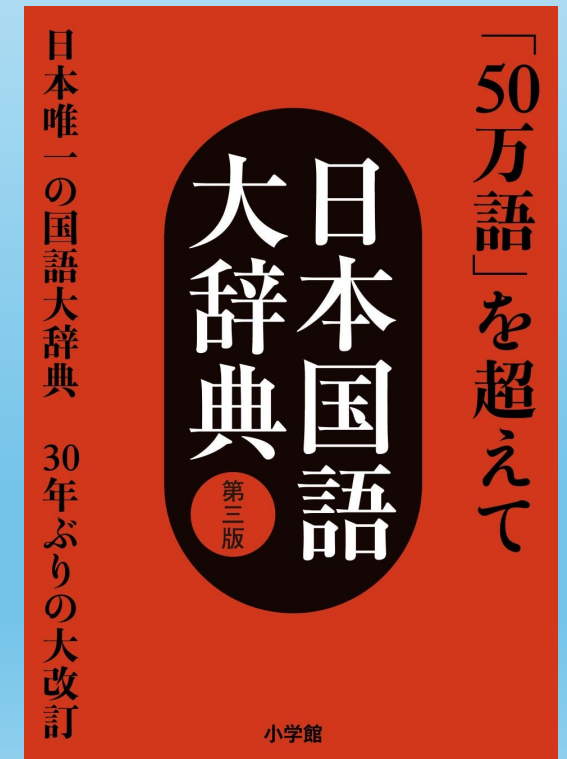
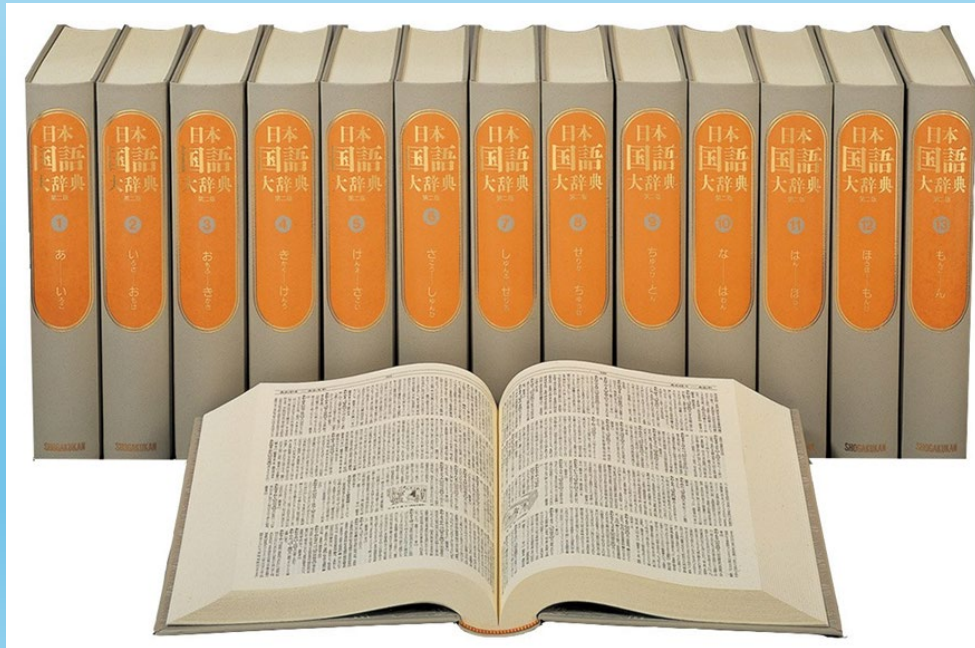


国語辞典におけるコーパス利用について

～『日本国語大辞典』を中心に～



2025.9.9 神永 暁



本日の話の内容

- **辞書の用例について**

辞書になぜ用例を載せるのか

(『日本国語大辞典』を中心に)

- **『日本国語大辞典』 第3版の
コーパス活用について**

(第3版 2032年デジタル版 2034年書籍版の刊行を予定)

『日本国語大辞典』 改訂の歴史

大日本国語辞典	<u>1915</u>	修訂版	<u>1939</u>
日本国語大辞典			<u>1972-1973</u>
日本国語大辞典 第二版			<u>2000-2002</u>
日本国語大辞典 第三版			<u>2032</u> ???

初版は活版印刷だったが、第二版では初版のデータをOCRで読み取り、データ処理が可能になった。組版も最新技術を駆使した。
第三版では…



前身の「大日本国語辞典」以来、20-30年ごとに改訂が行われてきた

用例主義の継承と発展

『日本国語大辞典』では**用例主義**を採用。

「**国語大辞典の生命は、まずその引例文にある**」（初版「発刊の辞」）

膨大な文献資料から語の使用例（用例）を採集し、それを元に立項と解説（語釈）の執筆を行った。

初版や第二版と同じく、第三版でも**用例主義**を継承する。

一般的な国語辞典

...基本的には作例*1を掲出。実例はあっても少数。

日本国語大辞典

...基本的には必ず実例を掲出する。

*1辞書編集者が作った例
（コーパスで探した例を参考にする
ことも多い）

●古典資料の再検討

●近現代資料の検討

●デジタル化資料の活用

膨大な用例をどのように採集するか、
国語研コーパスなどデジタル化され
た資料をどう活用するか

『日本国語大辞典』 第2版の用例

- ・ 実際に小説、詩歌、随筆などの作品や、古辞書、歴史資料などで使われた例。
- ・ 約 3 万点の資料から100万例を収録。
- ・ 最古の例である聖徳太子の時代のものから、昭和の末年のものまでを、典拠とともに示す。

『日本国語大辞典』における用例採取の目安

- **いちばん古いと思われるもの**：いわゆる初出例だが、最古の例ではない
- **見出し語の意味をくみとりやすいもの**：とりあえずその語の存在証明として引用している用例もある
- **主要作品の用例**：「主要作品」とは何かという問題はあるが
- **その語の文字づかいの特殊なもの**：たとえば「がんじょう（＝堅固でじょうぶな意）」の場合、「五調」「頑丈」「岩畳」「岩乗」「巖茸」「強盛」などの表記の用例
- **その他の意味で適切なもの**：派生義の用例

『日本国語大辞典』 第二版での用例採集方法：二通りの方式

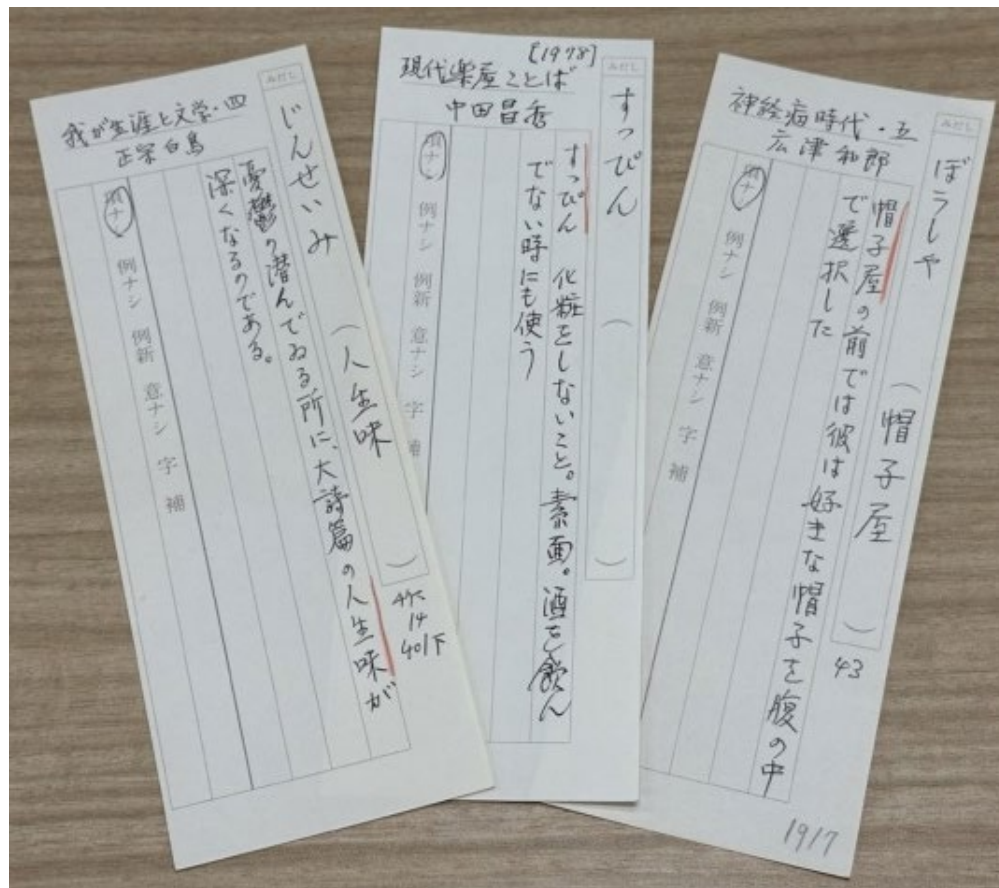
①スカウト式用例採集

- 資料を選び、採集者の判断で用例を採取する。
- 採取した用例はカード化して、それを一つ一つ入力する。

②特に重点的に用例を採取すると決めた資料については、編集部で独自に本文を入力。

- ただし当時はコーパス化という発想はほとんどなかった。
- 採取を決めた資料をコピーし、ここからここまでという用例部分を示すかぎ括弧「 」をつけ、さらに見出し語の候補であることを示すよみがなつけてそれら全体を入力する。
- いわゆるKWIC検索方式（キーワードを前後の文脈つきで表示する方式）。

●

[illegible]

襦 わぎ に氣なし (官外骨) や は 腋 わき 下 もと も も 老 らう 所 しよ とせよ	項ナシ 例ナシ 例新 意ナシ 字 (補)	腋 わき 毛 け は何の用を為すかといふに	みだし わぎ (腋毛) 89
--	-------------------------------------	---	-------------------------

1920

『日本国語大辞典』第二版での用例採集方法（続き）

- 採取した用例データと、OCRで読み取った初版の見出し語データとをマッチング処理し、マッチ項目（用例の追加候補）とアンマッチ項目（新規立項項目の候補）とに分けて処理する。
- どちらかといえば資料優先、用例優先。

第3版では資料優先で用例を採取するものもあるが、多くのデジタル化資料を使って、幅広く用例を採取する予定。

辞典に用例（実際の使用例）をつけることの効用

- ① その語が確かに使われていたという証拠になる
- ② その語の使われていた時代がわかる
- ③ その語の意味の理解を助ける
- ④ その語の使い方の手がかりが得られる
- ⑤ その語の読みに関する資料となる
- ⑥ その語の表記に関する資料となる

（松井栄一『出逢った日本語・50万語』（2002年）より）

「辞典に用例（実際の使用例）をつけることの効用」に追加したいこと

⑦いわゆる誤用といわれているの用法の広まりがわかる。

JapanKnowledge Personal コンテンツ | 登録情報の確認と

日本国語大辞典

親見出し まと【的】

まとを得（え・う）る

的確に要点をとらえる。要点をしっかりとおさえる。当を得る。的を射る。

＊白く塗りたる墓〔1970〕〈高橋和巳〉二「よし子の質問は実は的をえていた」

「的を得る」は現行の国語辞典も判断が分かれる

- 『明鏡国語辞典』 第3版

「当（とう）を得（え）る（＝理にかなう）」や「要領を得る」との混同から「的を得る」とするのは誤り。

- 『三省堂国語辞典』 第8版

「的を射る」「的を得る」とも、戦後に広まった言い方。「得る」は「要領を得る」などと同様、「うまくとらえる」の意味。同義語「正鵠（せいこく）を〈射る/得る〉」も二通りの言い方がある。

BCCWJで「的を得る」「的をえる」を検索すると…

〈合計：29例〉

- 「このような指摘もまったく的をえていないとはいいがたいのだが」
(1989年)
- 「これは誰が語ったものか忘れたが、なかなか的を得ている」 (1990年)
- 「勿論推測でしかないわけだけれど、まあ九十％は的を得ていると思うわ」 (1992年)
- 「筆者は、ポズナーに対するその他の批判の多くは的を得ていると考えており」 (1994年)

などなど

「的を射る」「的を得る」とも、戦後に広まった言い方

(『三省堂国語辞典』第8版)

JapanKnowledge Personal

コンテンツ | 登録情報の確認と変更 | 使い方／サポ

日本国語大辞典

親見出し まと【的】

まを射(い)る

(1) 矢や弾を的に向かって放つ。また、放った矢、弾が的に命中する。

＊撰集抄〔1250頃〕九・三「何々とあらなんと案じゐたるは、とかくの弓に亀毛の矢をはげ、空化のまをいんとするにたがはず」

＊日葡辞書〔1603～04〕「Matouo iru (マトヲイル)〈訳〉的に向かって射る」

(2) 的確に要点をとらえる。

＊島崎藤村〔1946～56〕〈平野謙〉新生「このようなトルストイの姿勢が的(マト)を射ているかどうかは」

＊宇宙人〔1964〕〈倉橋由美子〉「Lが発揮する直観力は驚くほどの的を射る」

現時点で、「的を射る」の「的確に要点をとらえる」という意味の用例は。平野謙『島崎藤村』とほぼ同時時代の戦後の用例しか見つかっていない。

第3版で、戦前の用例がデジタル化資料などから見つけれられるかが課題。

採取した用例はそのまま引用するわけではない

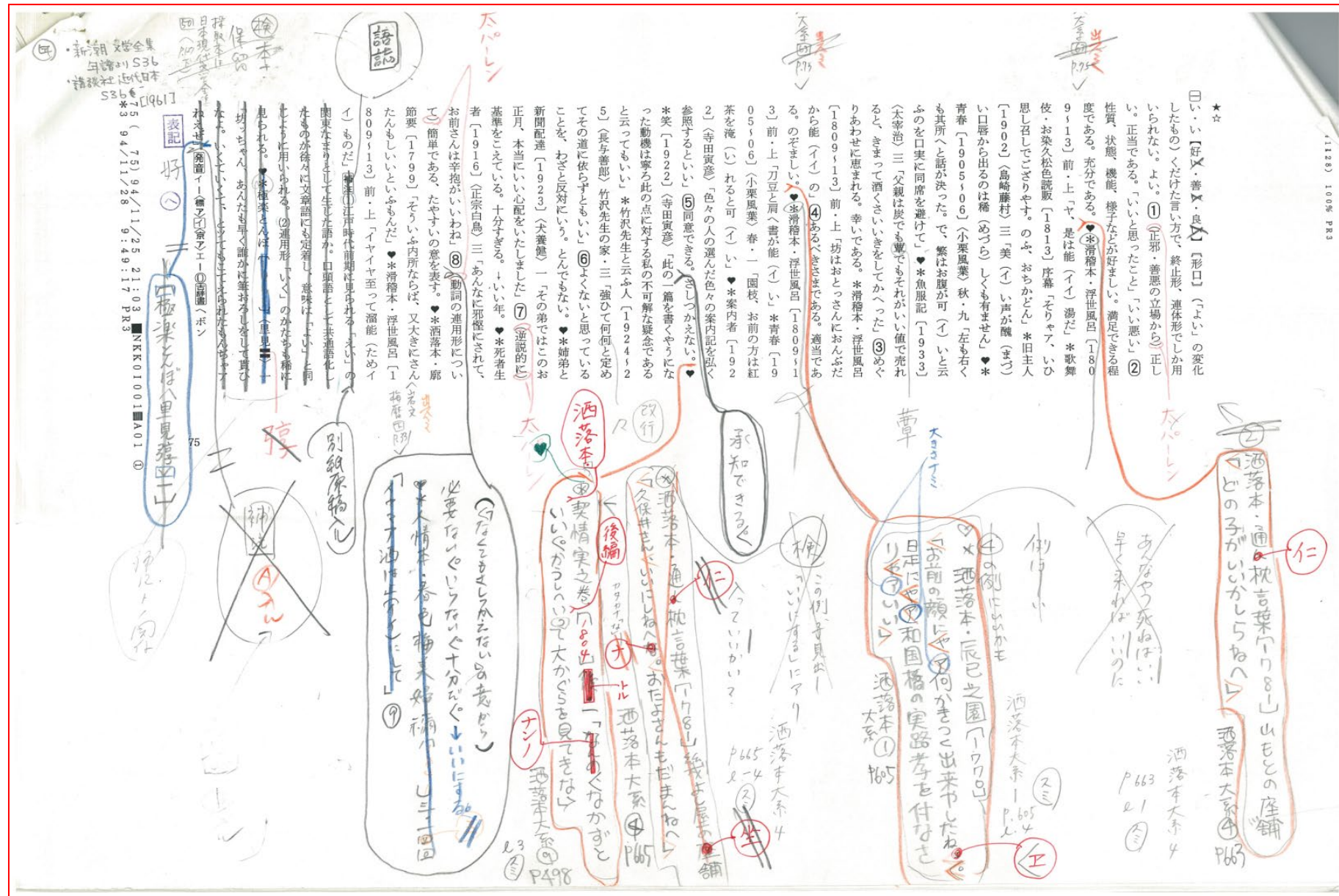
「出典検討」

- 前提：用例と語釈は不即不離の関係にある。
- 「出典検討」とは、原稿の執筆者が採用した例を、決められたテキストで当たりなおす作業。
- 文字・かなづかいを点検し適当な長さに引く。
- 決められた注記、即ちジャンル名・小題・番号・作者名等を付記する。
- 出てくる用例は時代を問わず、ジャンルを問わずであるが、その意味を読みとり、語釈と合わせ考え、語釈が適切か判断する。

『日本国国語 大辞典』 初版原稿

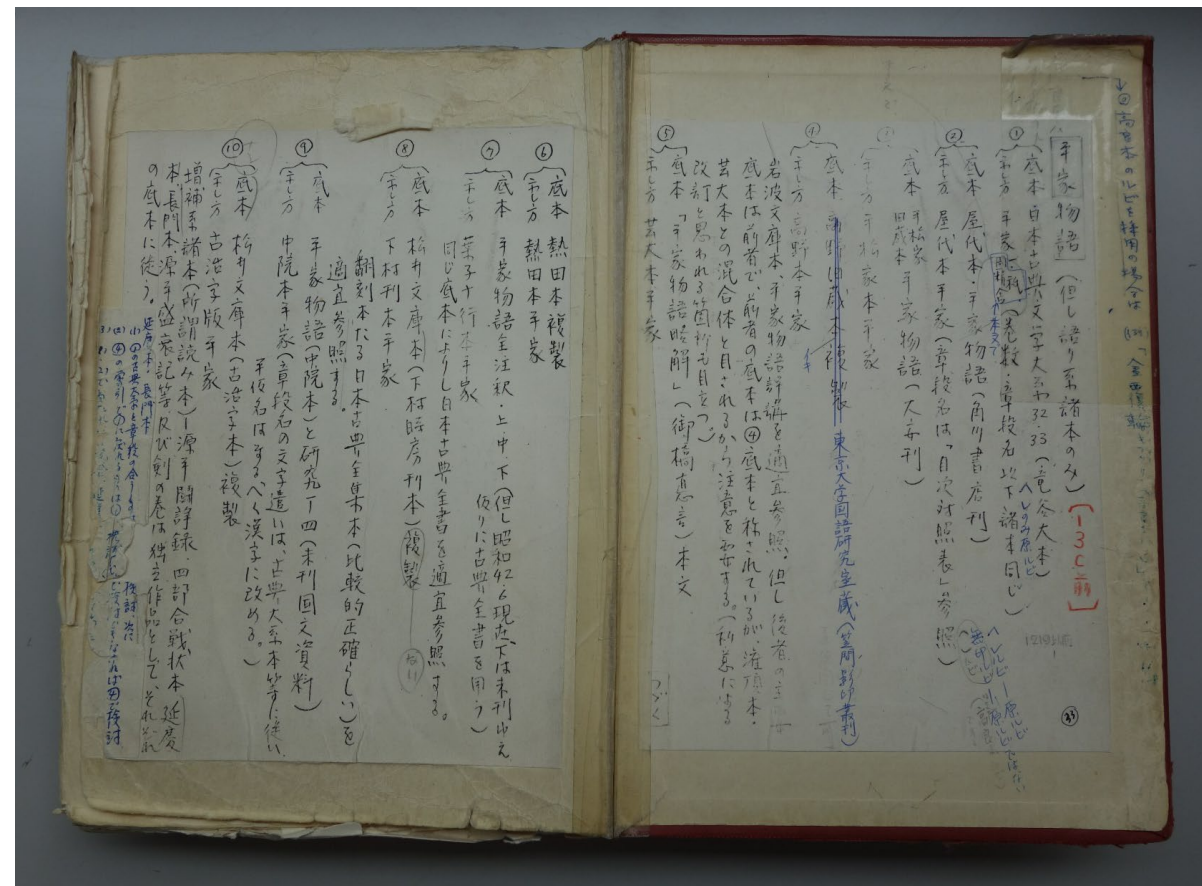
										あ い う え お									

『日本国国語大辞典』 第2版ゲラ



編集部には用例を採取した数万点の資料がある。
今後このような資料の収集はほとんど必要なくなる
かもしれない。

2025/9/9



『日国』の『平家物語』の底本。第1底本の岩波
古典文学大系本の見返し。第1底本で用例が取れ
なかった場合は、第2底本（屋代本）以下底本順
位を下げて出典検討をしていくという指示。

18

『日本国語大辞典』第2版の用例の課題と展望

- すべての時代の用例をまんべんなく採取しているわけではない。
- 見出し語を立てていながら、用例のないものがある。
- 「初出」例を載せるように心がけているが、実際にはさらに古い例が見つかることもある。
- 第二版までは紙版だったので載せられる分量の制約があったことから、基礎的な語でも近代例が手薄な語もある。
- 漫画からの用例をどのように扱うか。

それらを解消するためにコーパス等、外部リポジトリを活用する

『日本国語大辞典』では、BCCWJをどのように活用できるか

- 1, 頻度数によって見出し語選定の参考にできる。
- 2, 記述する意味の説明の参考になる（特に多義語の場合）。
- 3, 新立項項目の参考になる。
- 4, 表記に揺れのある語の見出し表記を決める根拠が得られる。
- 5, 初出にはならないにしても、例ナシ項目（多くはないが存在する）の用例を探することができる。
- 6, ある語と接続しやすい語の特定が可能になる
（典型的な使用例。コロケーション）。

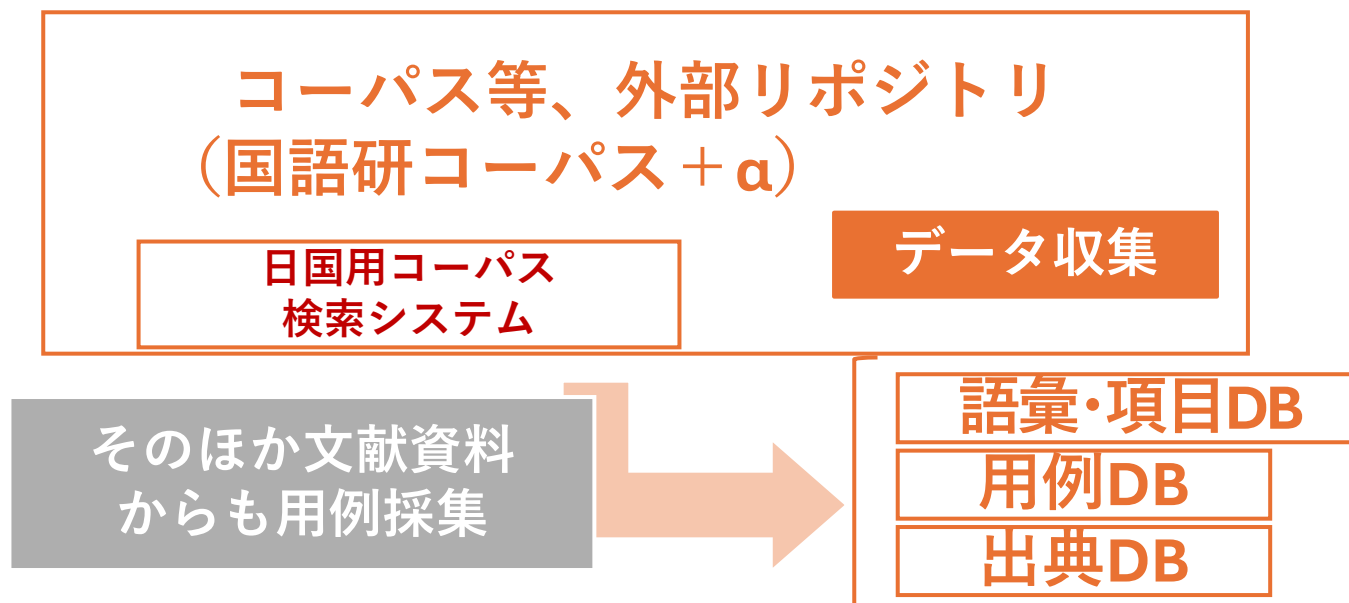
『日本国語大辞典』 第三版改訂の要諦

校正紙(ゲラ)から、データへ

書籍版中心から、電子版中心へ

ペーパーベースドからデジタルベースドに移行し、
編集作業も、提供の形態も、デジタル技術を駆使して進化させる。

編集作業の進化 = DX デジタル・トランスフォーメーション【digital transformation】



※1日本語歴史コーパス
※2現代日本語書き言葉均衡コーパス

CHJ※1、BCCWJ※2等さまざまなデジタル化資料を活用、および必要な資料をコーパス化
独自検索システムを作成、編集支援システムと連携

資料から用例とその出典情報を採集し、編集支援システムに取り込み

編集支援システム
データ編集

編集支援システム

編集支援
システム

データ編集

作業フローの中核

用例を元に語の意味を語釈、あるいは既存の語釈を、新しく採集した用例を元に修正

組版プログラム

エクスポートしたXMLをプログラム処理で自動組版

出典検討作業

校正システム

校正作業

知的情報のハブへ

