

1 はじめに

『日本語歴史コーパス 江戸時代編V注釈書』は、本居宣長の古今和歌集の注釈書『古今集遠鏡』1作品を収録している。『古今集遠鏡』コーパスは、はじめに『本居宣長全集』第三集のテキストをもとに Himawari 版 (Ver.0.8) が公開され、その後、市村太郎所蔵の文化十三年の刊記を持つ『古今集遠鏡』を底本とした XML データの再構築が行われ、Himawari 版 (Ver.1.0) が公開されている (市村 2023、市村ほか 2024)。再構築された XML データに形態論情報を付与し、検索アプリケーション「中納言」で検索可能な形としたものが本コーパスである。本文書では、「中納言」での公開にあたって底本に加えた校訂と、それらが「中納言」で検索・表示される際の各種情報について、その概要を述べる。

2 テキストの凡例

2.1 テキストの範囲

本コーパスのテキストは、文化十三年刊『古今集遠鏡』のうち、原則として全てのテキストを対象としている。ただし、横井千秋による序文、広告、刊記は対象範囲外とした。

2.2 一部漢字の置き換え

本コーパスでは、XML データの再構築にあたり、『本居宣長全集』で用いられていた一部の「正字體」を、形態素解析や「中納言」上での文字列検索を視野に入れ、現在の通行の字体に変換している (市村 2023「5. 漢字の使用に関する問題」)。

例えば、「樂」を「楽」と改めたほか、文化十三年版本では「歌」「哥」のほか「哥」の上部が「可」ではない字体がみられたが、データ上ではすべて「歌」に統一している。置き換えた漢字のリストは市村 (2023) を、『古今集遠鏡』での実際の表記は、「中納言」で参考リンクとして示される、底本と同じ版と見られる早稲田大学所蔵本を参照されたい。

2.3 カタカナ表記の置き換え

底本では、宣長訳が漢字カタカナ交じりで表記されているほか、一部振り仮名がカタカナで表記されている。本コーパスでは、宣長訳をひらがな漢字交じりに変換し、振り仮名内のカタカナは底本のままとした。変換前の文字列は「中納言」の検索結果の「原文文字列」列および「原文 KWIC」欄に表示される。

【例】

- (1) 〈原文〉 ○年内ニ春ガキタワイ コレデハ 同シ一年ノ内ヲ 去年ト云タモノデアラウカ ヤツハリ
コトシト云タモノデアラウカ

- (2) 〈入力〉 ○年内に春がきたわい これでは 同し一年の内を 去年と云たものであらうか やつぱり ことしと云たものであらうか (52-遠鏡 1793_00201、620)

2.4 濁点・半濁点の付与

濁音・半濁点が期待される箇所に濁点・半濁点が付されていない場合は、諸資料を参考に、必要な箇所は濁点を補った。濁点を補う前の文字列は、「原文文字列」に表示される。また、XML データの段階で濁点を補わなかったものの、形態論情報付与作業の際に濁点落ちと判断して語形登録・処理した箇所もある。

【例】

- (3) 〈原文〉 をちこちのたづきもしらぬ山中におほつかなくもよぶこどりかな
 〈入力〉 をちこちのたづきもしらぬ山中におほつかなくもよぶこどりかな (52-遠鏡 1793_00201、30030)
- (4) 〈原文〉 夜ノ\ソレデモヤツハリ袖ハナミダガ氷ルワイ
 〈入力〉 夜ノ\それでもやつぱり袖はなみだが氷るわい (52-遠鏡 1793_00212、50170)

2.5 踊り字の校訂

踊り字仮名 1 字分の踊り字 (ゝ、ゞ、ゝ、ゞ) は、想定される仮名に置き換えた。変換前の文字列は「原文文字列」および「原文 KWIC」に表示される。2 字分以上に相当するくの字点は置き換えの対象としない。

【例】

- (5) 〈原文〉 春霞たてるやいづこみよし野のよしのゝ山に雪はふりつゝ
 〈入力〉 春霞たてるやいづこみよし野のよしのの山に雪はふりつつ (52-遠鏡 1793_00201、2020)
- (6) 〈原文〉 アゝゝ鶯ノオモシロウ鳴イテキタ
 〈入力〉 あああ鶯のおもしろう鳴てみた (52-遠鏡 1793_00202、39160)
- (7) 〈原文・入力同じ〉 野べちかく家ゐしせればうぐひすのなくなる声は朝なノ\きく (52-遠鏡 1793_00201、17560)

2.6 傍記の扱い

底本における傍記は、次のような性質のもので構成されている。

- ・振り仮名 (図 1)
- ・送り仮名 (図 2)
- ・捨て仮名 (図 3)
- ・対訳 (図 4)
- ・発音注記 (図 5、6)
- ・内容・句の順番注記 (図 7)

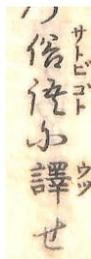


図 1

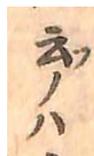


図 2

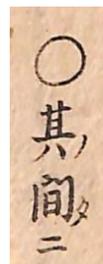


図 3

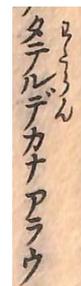


図 4

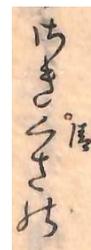


図 5

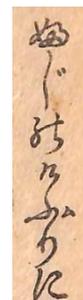


図 6

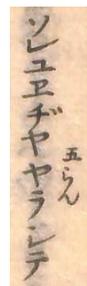


図 7

これらの傍記のうち、活用語尾、助詞の小書きは、旧国名や氏名官職の間に付される「の」を除いて本文文字列としてデータに反映されている（市村 2023）。そのほかの傍記は、「中納言」上ではその性質に関わらず振り仮名欄に表示されるため、実際の傍記の位置や内容は、「中納言」で参考リンクとして示される早稲田大学所蔵本を確認することを推奨する。

2.7 記号の処理

底本では、宣長が和歌に俚言訳を施す際、序詞などの元の句を省略する場合には訳を施さなかった句を示す印をつけられている。図1～3に示すように、上の句を省略する場合の「上」、一句目を省略する場合の「一」などが、本文中に見られる。

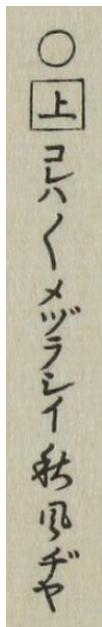


図8 「上」の例（巻四）

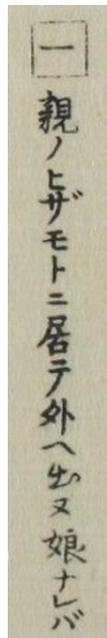


図9 「一」の例（仮名序）

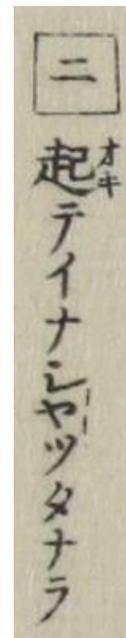


図10 「二」の例（仮名序）

上記のような箇所は、「上」や数字のみテキストに示し、品詞を「記号」とした。実際の表記は、中納言で参考リンクとして示される早稲田大学所蔵本を参照されたい。

3.3 本文情報

本コーパスのデータは全て人手による修正を加え、コアデータとして公開する。

3.4 本文情報

3.4.1 本文種別

『古今集遠鏡』コーパスでは、本文種別を次のように区別している。

- ・ 古今集 『古今和歌集』の詞書や地の文
- ・ 古今集-歌 『古今和歌集』の和歌にあたる部分
- ・ 古今集-作者 『古今和歌集』の和歌の作者
- ・ 宣長訳 ○から始まる、本居宣長による和歌の俚言訳
- ・ 千秋 横井千秋による解説
- ・ 空欄 地の文や他の資料の引用などを含む宣長の注釈

これらの本文種別は「中納言」検索結果の「本文種別」欄に表示する。

3.4.2 話者

本コーパスの「中納言」検索結果の話者名には、和歌の作者が表示される。詠み人が不明な和歌については、空欄となる。和歌作者は本文中での表記を用いて表示している。また、和歌の作者の性別も、「中納言」検索結果の性別列に表示されるが、同時代編「Ⅰ洒落本」「Ⅱ人情本」「Ⅲ近松浄瑠璃」で表示される身分・地域・年齢といった属性は記していない。

3.5 作品情報

「中納言」検索結果の各欄に表示される内容は下記の通りである。

- ・ ジャンル
「ジャンル」には「注釈書」と表示される。
- ・ 作品名
「作品名」には『古今集遠鏡』と表示される。
- ・ 成立年
『古今集遠鏡』の成立年は「1793年ごろ」とする先行研究が多く、明確ではないが、「中納言」上では「1793」とした。
- ・ 卷名等
底本の卷名である、「古今集遠鏡例言」「古今集遠鏡例言」「古今和歌集序遠鏡」「古今和歌集巻第一遠鏡」～「古今和歌集巻第二十遠鏡」がそれぞれ表示される。
- ・ 作者

『古今集遠鏡』の著者である「本居宣長」が表示され、国立国会図書館典拠データ検索・提供サービス (NDLAuthorities) へリンク付けを行った。

3.6 底本情報

3.6.1 底本

底本は全6冊で、次の通りに収録されている。

- ・ 「古今集遠鏡 一」 序文、例言、巻第一・二
- ・ 「古今集遠鏡 二」 巻第三～巻第六
- ・ 「古今集遠鏡 三」 巻第七～巻第十
- ・ 「古今集遠鏡 四」 巻第十一～巻第十三
- ・ 「古今集遠鏡 五」 巻第十四～巻第十六
- ・ 「古今集遠鏡 六」 巻第十七～巻第二十

「中納言」検索画面における底本の箇所には、該当する冊名が表示される。

3.6.2 ページ番号

ページ番号には、『本居宣長全集』（参考リンク）の該当箇所のページ番号が表示される。

3.6.3 参考リンク

本コーパスの底本である文化十三年の刊記を持つ市村太郎架蔵本は、市村太郎氏の HP で公開されているが、「中納言」との連携は難しいため、参考リンクとして同版とみられる早稲田大学古典籍総合データベースの該当箇所が閲覧可能である。略称ロゴ (Waseda) をクリックすると、該当箇所の画像が表示される。

【参考規程・概要書】

片山 久留美 (2019) 「『日本語歴史コーパス 室町時代編』形態論情報規程集 Ver.1.0」

https://clrd.ninjal.ac.jp/chj/morph_muromachi_v1_0.pdf.pdf

村山 実和子 (2018) 「『日本語歴史コーパス 江戸時代編 I 洒落本』テキストの凡例と『中納言』表示項目について」 (2019年3月29日更新)

<https://clrd.ninjal.ac.jp/chj/doc/abstract-share-2019.pdf>

村山 実和子 (2019) 「『日本語歴史コーパス江戸時代編 I 洒落本』『同 江戸時代編 II 人情本』形態論情報概要」 <https://clrd.ninjal.ac.jp/chj/doc/morph-edo-2019.pdf>

村山 実和子 (2019) 「『日本語歴史コーパス 江戸時代編 II 人情本』テキストの凡例と『中納言』表示項目について」 <https://clrd.ninjal.ac.jp/chj/doc/abstract-ninjo-2019.pdf>

片山 久留美 (2020) 「『日本語歴史コーパス 江戸時代編 III 近松浄瑠璃』テキスト・表示項目について」 <https://clrd.ninjal.ac.jp/chj/doc/abstract-chikamatsu-2020.pdf>

片山 久留美 (2020) 「『日本語歴史コーパス江戸時代編 III 近松浄瑠璃』形態論情報の概要」

<https://clrd.ninjal.ac.jp/chj/doc/morph-chikamatsu-2020.pdf>

片山 久留美 (2020) 「『日本語歴史コーパス 江戸時代編Ⅲ近松浄瑠璃』副本文の形態論情報の概要」

<https://clrd.ninjal.ac.jp/chj/doc/morph-at-chikamatsu-2020.pdf>

松崎 安子 (2021) 「『日本語歴史コーパス 江戸時代編Ⅳ随筆・紀行』Ver.0.4 (芭蕉の紀行文) テキストの凡例と『中納言』表示項目について」

<https://clrd.ninjal.ac.jp/chj/doc/abstract-zuihitsu-202103.pdf>

松崎 安子 (2023) 『日本語歴史コーパス 江戸時代編Ⅳ随筆・紀行』Ver.0.8 テキストの凡例と「中納言」表示項目について」

<https://clrd.ninjal.ac.jp/chj/doc/morph-zuihitsu-202303.pdf>

【参考文献】

今西 祐一郎校注 (2008) 『古今集遠鏡 1』東洋文庫 770、平凡社

今西 祐一郎校注 (2008) 『古今集遠鏡 2』東洋文庫 722、平凡社

大久保 正編 (1969) 『本居宣長全集第三卷』筑摩書房

久保 柁子、市村 太郎、小木曾 智信 (2023) 「『古今集遠鏡』コーパスの設計と構築」『人文科学とコンピュータシンポジウム 2023 論文集』 pp.279-pp.284

市村 太郎 (2023) 「『古今集遠鏡』を対象とするコーパス構築の試み」岡部嘉幸・橋本行洋・小木曾智信編『コーパスによる日本語史研究 近世編』 pp.259-pp.278、ひつじ書房

【関連する URL】

市村 太郎、久保 柁子、岩崎 凜太郎編 (2024) 『Himawari 版古今集遠鏡コーパス Ver.1.0』

<https://sites.google.com/view/ichimurat/%E5%B8%82%E6%9D%91%E7%A0%94%E7%A9%B6%E5%AE%A4/%E7%A0%94%E7%A9%B6%E8%B3%87%E6%96%99?authuser=0>

国立国語研究所 (2024) 『日本語歴史コーパス』

<https://clrd.ninjal.ac.jp/chj/>

早稲田大学図書館 <https://www.waseda.jp/library/>