

『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

2021年 3月 31日 松崎安子

1 はじめに

『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) (以下「本コーパス」と呼ぶ) は、『新編日本古典文学全集 71 松尾芭蕉集(2)』(小学館) 所収の「日記・紀行編」を底本としている。コーパス本文の作成に際してはその一部を校訂しているものの、それらの情報はXML データに記述しデータ構築を図った。本稿では底本に加えた校訂とそれらが検索アプリケーション「中納言」によって検索・表示される場合の各種情報についてその概要を述べる。

2 本文の範囲とテキスト校訂

本コーパスのテキストは『新編日本古典文学全集 71 松尾芭蕉集(2)』に「紀行・日記編」として収録されている「鹿島詣」、「野ざらし紀行」、「更科紀行」、「笈の小文」、「嵯峨日記」、「おくのほそ道」を使用している。ただし、『日本語歴史コーパス』へ収録するにあたり、以下に記すような項目について校訂を行った。

2.1 一部の片仮名表記の置き換え

底本において助詞、副詞の一部、活用語尾の一部、助動詞、サ変動詞などに散見する片仮名表記部分については、UniDic への登録語としての統一化を考慮し、平仮名表記へと置換した。置換前の文字列は、「中納言」検索画面において「原文文字列」および「原文 KWIC」に表示される。

[1] <底本> 去来^{らくししや}ガ^{いたる}落柿舎に到。 → <校訂後> 去来^{らくししや}が^{いたる}落柿舎に到。

(51-芭蕉 1691-02001, 200)

[2] <底本> おかしき事の^ミ多し。 → <校訂後> おかしき事の^み多し。

(51-芭蕉 1689-01001, 1630)

[3] <底本> 暫^{しばら}ク^{まん}学^ぐで愚^{きたら}を暁^ん事を → <校訂後> 暫^{しばら}く^{まん}学^ぐで愚^{きたら}を暁^ん事を

(51-芭蕉 1691-01001, 1630. 1700)

[4] <底本> 何がしの方に音信^ル。 → <校訂後> 何がしの方に音信^る。

(51-芭蕉 1694-01009, 150)

[5] <底本> 大^{おほ}キ^{なるくり}成栗^{こかげ}の木陰を → <校訂後> 大^{おほ}き^{なるくり}成栗^{こかげ}の木陰を

(51-芭蕉 1694-01013, 2340)

[6] <底本> 御山^{おやま}に詣^{けいはい}拜^す。 → <校訂後> 御山^{おやま}に詣^{けいはい}す。

(51-芭蕉 1694-01006, 110)

2.2 一部の漢字の置き換え

本コーパスのテキストの文字入力は Unicode に準拠している。Ver.0.4 では底本の字体のうち現行の字体に適切なものがない「暹」,「璚」は各々「𠄎碗」(51-芭蕉 1689-01001,9390),「𠄎岐」(51-芭蕉 1691-02001,22270)のように「𠄎」(U+3013)へ置き換えてコーパスの本文表記とした。

一方、底本の字体のうち『新編日本古典文学全集』の頭注などで異字体表記とのことわりが明記され、あわせて現行表記による字体が表示されている場合は、現行表記の字体を以てコーパスの本文表記とした(表1)。

表1 『日本語歴史コーパス 江戸時代編IV随筆・紀行』Ver.0.4における異体字置換一覧

	字体		作品名
	底本	校訂本文	
1	𠄎	盡	野ざらし紀行、更科紀行、 笈の小文、おくのほそ道
2	宵	宵	笈の小文、嵯峨日記
3	騷	騷	おくのほそ道
4	装	装	おくのほそ道
5	霧	霧	おくのほそ道
6	寮	寮	おくのほそ道

2.3 踊り字の校訂

仮名1字分の踊り字(ゝ, ゞ, っ, っ)は、想定される仮名に置き換えた。置換前の踊り字表記は「中納言」検索結果画面では「原文文字列」および「原文 KWIC」に表示される。

一方、2字以上の繰り返しに相当するくの字点や、漢字一字の繰り返しを表す「ゝ」は置き換えの対象としない。

ただし、例[10]のように、文節末の同語をくりかえし、かつ、文節の頭となる場合は、想定される語を仮名に開いたテキストとしている。

[7] <底本> 若き僧共, 紙・硯をかゝへて → <校訂後> 若き僧共, 紙・硯をかかへて
(51-芭蕉 1694-01043,2400)

[8] <底本> 三河の国の地つゞきにて → <校訂後> 三河の国の地つづきにて
(51-芭蕉 1691-01001,11920)

[9] <底本> 風景に魂^{ふうけい たましひ}うバゝれ → <校訂後> 風景に魂^{ふうけい たましひ}うばはれ
(51-芭蕉 1694-01013,1540)

[10] <底本> ともなふ人ふたり, 浪客の士ひとり, 〱は水雲の僧。
→ <校訂後> ともなふ人ふたり, 浪客の士ひとり, ひとり^〱は水雲の僧。

(51-芭蕉 1687-01001,950)

2.4 合略仮名の置き換え

底本における合略仮名「方」は平仮名表記「より」へと開き置き換えた(例 [11])。置換前の文字列は、「中納言」検索画面の「原文文字列」および「原文 KWIC」に表示される。

[11] 〈底本〉それ方八幡宮に詣。→ 〈校訂後〉それより八幡宮に詣。

(51-芭蕉 1694-01009,1280)

2.5 テキストへの補い

底本において助詞等が振り仮名として表記されている場合、それを元のまま保持しながら、校訂本文にも平仮名で補った(例 [12] [13])。

[12] 〈底本〉岩沼宿 → 〈校訂後〉岩沼の宿 (51-芭蕉 1694-01019,30)

[13] 〈底本〉朋友信有哉 → 〈校訂後〉朋友に信有哉 (51-芭蕉 1687-02001,1820)

2.6 字順転倒箇所の校訂

底本には、漢字文字列に対して返り点や振り仮名により字順を置き換えて読むことが指示されている箇所がある。本コーパスでは、それらの指示通りの語順となるよう字順を転倒させたり、漢字ひらがな交じりで書き下し文としたりする校訂を加えた箇所がある。校訂前の文字列は、「中納言」検索画面の「原文文字列」および「原文 KWIC」に表示される。

[14] 〈底本〉ゆきは不申先むらさきにつくばかな
→ 〈校訂後〉ゆきは申ず先むらさきにつくばかな (51-芭蕉 1687-01001,4460)

[15] 〈底本〉題落柿舎 → 〈校訂後〉落柿舎に題す (51-芭蕉 1691-02001,14030)

[16] 〈底本〉史邦・丈艸被訪。→ 〈校訂後〉史邦・丈艸訪る。(51-芭蕉 1691-02001,14920)

2.7 不読字への対応

底本の漢字文字列に付された振り仮名の通りに読むと不読となる漢字については、コーパスのテキストには表示しない。処理前の文字列は「中納言」検索結果画面中の「原文文字列」および「原文 KWIC」に表示される。

[17] 〈底本〉将軍恵美朝臣朝獺修造而 → 〈校訂後〉将軍恵美朝臣朝獺修造

(51-芭蕉 1694-01021,1060)

2.8 漢詩作品部分

本コーパスでは、『日本語歴史コーパス 江戸時代編』として収録されている他のコーパス同様、漢詩部分は字順転倒などの処理を行わず、形態論情報を付与せず、品詞を「漢文」としたうえで、文字列検索にのみヒットするようにした。

[18] ふかくがほうにたいしててうぎよをともなふ 深対岷峯伴鳥魚くわうにつきやじんのきよににたるをよるこぶ 就荒喜似野人居しとういまかくせききうのらん 枝頭今欠赤虬卵

青葉分題堪学書 (51-芭蕉 1691-02001,15030~. 15100~. 15170~.15240~)

[19] たつてあんじやうをみだしてしんきゆうをいづ 強攬怨情出深宮いちりんのしうげつやそんのかぜ 一輪秋月野村風せきねんわづかにきんめんをもとめえたり 昔年僅得求琴韻いづごごふんちくじゆのうち 何処孤墳竹樹中

(51-芭蕉 1691-02001,15390~.15460~.15530~.15600~)

[20] もんをとちてくをもとむあんむきもんをとちてくをもとむあんむき 杜門寛句陳無己きやくにたいしてふでをふるよしんせういう 杜門寛句陳無己，对客揮毫秦少游。

(51-芭蕉 1691-02001,16010~.16090~)

3 「中納言」における表示項目と内容

本コーパスの本文に付したさまざまなタグや単語情報はコーパス検索アプリケーション「中納言」上に検索結果として表示される。表示については他のサブコーパスと統一を図っているため概略・省略しながら、各項目について本コーパスの表示内容を加え記す。

3.1 コーパス情報

3.1.1 時代名，サブコーパス名

時代名は「5 江戸」，サブコーパス名は「江戸随筆・紀行」と表示する。

3.1.2 サンプルID

『日本語歴史コーパス』に統一的なサンプルID付与のルールに従いながら，全作品に付与した。本コーパスでは表2に示す4, 5桁目を「芭蕉」とすることで，芭蕉の紀行文であることを示している。

表2 サンプルID付与のルール

1-2桁目	3桁目	4-5桁目	6-9桁目	10桁目	11-12桁目	11-15桁目
時代	区切り記号	作品ID	成立時期	区切り記号	同成立年作品の区別	同作品内の出現順通し番号

3.1.3 コア

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。

3.1.4 主本文

本コーパスでは、検索結果の用例が掛詞のものであれば「0」を、主たる本文のものであれば「1」を表示する。

3.1.5 多重化種別

本コーパスでは掛詞の技法が用いられている箇所にその情報を持たせており、それは「副本文」として扱われる。検索箇所に掛詞の情報が付与されている場合、「検索動作」において「副本文を検索対象に含まない」状態（図 1）で検索を行ったとしても、「列の表示」において「多重化種別」にチェックを入れ、列の表示を増やす（図 2）と検索結果画面において多重化種別を見ることができる。

多重化種別にチェックを入れると、図 3 のようにデフォルトの列表示から検索表示列が増える。ここで多重化種別の列を見ると「掛詞」という多重情報の種類が表示されていることから、主本文であり表層形である「松」には掛詞の技法が用いられていることが判る。

図 1 「検索動作」において「副本文を検索対象に含まない」を選択した状態

検索動作 設定を隠す

文脈中の区切り記号 | 文脈中の文区切り記号 # 前後文脈の語数 20

副本文 副本文を検索対象に含まない 共起条件の範囲 文境界をまたがない

図 2 「列の表記」で「多重化種別」にチェックを入れた状態

列の表示 設定を隠す

コーパス情報

時代名 サブコーパス名 サンプル ID 開始位置 連番 コア 層 層内連番 主本文 多重化種別

図 3 「掛詞」の情報が付与されている主本文・表層形「松」の検索結果

サンプル ID	開始位置	コア	主本文	多重化種別	前文脈	キ	後文脈	語彙素読み	語彙素	語形	品詞
51-芭蕉 1694- 01019	2010	1	1	掛詞	た(くまの)松(みせ)申せ(運)松(と)挙白(と)云もの(の)鏡(別)したり(け)ね(ば、# 松)より	松	(は)二(木)を(三)月(越)し	マツ	松	マツ	名詞-普通名詞-一般
					た(くまの)松(みせ)申せ(運)松(と)挙白(と)云もの(の)鏡(別)したり(け)ね(ば、# 松)より	まつ松	(は)二(木)を(三)月(越)し				

3.2 本文情報

3.2.1 原文文字列

踊り字、カタカナ表記、不読字等を校訂する前の本文は「原文文字列」および「原文 KWIC」に表示される。ただし、これらはいくまで補助的な表示であるため、底本の詳しい状況を確認

認ずる場合は、ページ数を元に底本を直接ご参照いただきたい。

3.2.2 本文種別

本コーパスでは、本文種別を表3のように区別し、「中納言」検索結果の「本文種別」欄に表示する。なお、地の文は「本文種別」を空欄とした。

《本文種別の階層表記》

- 「会話」は登場人物の発話箇所である。
- 「前書」は紀行において詠まれた俳句の題や、詞書のような部分である。
- 「韻文」は紀行中に読まれた俳句、和歌に付与した。特に連句として表示されている場合は、何句目かが分かるように表示した。
- 「引用」は作者・芭蕉が他から引用した部分である。下位の階層としての「-典拠」は、文献等からの引用を表す。引用の内容が判明する箇所には「・」で区切り表示した。
- 「漢詩」は本コーパスでは形態論情報付与の対象ではないが、文字検索にヒットするため念のため本文種別を付与している。
- 「漢詩題」は「漢詩」に付された題名部分である。

表3 『日本語歴史コーパス 江戸時代編Ⅳ随筆・紀行』 Ver0.4 (芭蕉の紀行文) の本文種別

上位階層	以下の階層	上位階層	以下の階層
会話	ナシ -引用-典拠・和歌-韻文 -韻文-俳句	引用	-典拠・俳句-韻文 -典拠・和歌-韻文 -典拠・連句-韻文
前書	ナシ		-典拠・連句・発句-韻文
韻文	-俳句 -和歌 -連句・発句 -連句・脇句 -連句・第三句 -連句・第五句 -連句・第四句		-典拠・連句・脇句-韻文 -手紙 -典拠・漢文 -典拠・漢詩題 -典拠・その他
		漢詩	ナシ
		漢詩題	ナシ

3.2.3 話者

本コーパスにおける会話文の話者名については、可能な限り本文中での呼称を用い表示している。なお、同時代編「Ⅰ洒落本」、「Ⅱ人情本」、「Ⅲ近松浄瑠璃」とは異なり、話者の性別・身分・地域・年齢といった属性は記していない。

俳句や和歌の読み手については俳号や雅号を以て表示している。その一部は国立国会図書館典拠データ検索・提供サービス (NDLAuthorities) にリンクしている。

3.3 作品情報

●ジャンル

「ジャンル」には一律「紀行」と表示される。

●作品名

「作品名」には個別の作品名が表示される。

●成立年

収録作品のうちには成立年に諸説ある作品も含まれるが、『日本古典文学大辞典』（岩波書店）を参考とし成立年を定めた。

●巻名等

「おくのほそ道」については、校注者が設けた亀甲括弧付きの節番号を使用した。

以上を含め、作品名、サンプルID、巻名等の対応を右表4として示しておく。

●部

本コーパスでは空欄表示となる。

●作者情報

作者欄には作品の作者名を表示し、基本的な情報を得られるよう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。

3.4 底本情報

●底本

「底本」には「新編全集<71>」のように対応する底本を表示する。

●ページ番号

当該例が出現する『新編日本古典文学全集 71 松尾芭蕉集（2）』のページ番号を表示する。

表4 作品名とサンプルIDおよび巻名等の対応

作品名	サンプルID	巻名等
鹿島詣	51-芭蕉1687-01001	
野ざらし紀行	51-芭蕉1687-02001	
更科紀行	51-芭蕉1689-01001	
笈の小文	51-芭蕉1691-01001	
嵯峨日記	51-芭蕉1691-02001	
おくのほそ道	51-芭蕉1694-01001	〔一〕
おくのほそ道	51-芭蕉1694-01002	〔二〕
おくのほそ道	51-芭蕉1694-01003	〔三〕
おくのほそ道	51-芭蕉1694-01004	〔四〕
おくのほそ道	51-芭蕉1694-01005	〔五〕
おくのほそ道	51-芭蕉1694-01006	〔六〕
おくのほそ道	51-芭蕉1694-01007	〔七〕
おくのほそ道	51-芭蕉1694-01008	〔八〕
おくのほそ道	51-芭蕉1694-01009	〔九〕
おくのほそ道	51-芭蕉1694-01010	〔一〇〕
おくのほそ道	51-芭蕉1694-01011	〔一一〕
おくのほそ道	51-芭蕉1694-01012	〔一二〕
おくのほそ道	51-芭蕉1694-01013	〔一三〕
おくのほそ道	51-芭蕉1694-01014	〔一四〕
おくのほそ道	51-芭蕉1694-01015	〔一五〕
おくのほそ道	51-芭蕉1694-01016	〔一六〕
おくのほそ道	51-芭蕉1694-01017	〔一七〕
おくのほそ道	51-芭蕉1694-01018	〔一八〕
おくのほそ道	51-芭蕉1694-01019	〔一九〕
おくのほそ道	51-芭蕉1694-01020	〔二〇〕
おくのほそ道	51-芭蕉1694-01021	〔二一〕
おくのほそ道	51-芭蕉1694-01022	〔二二〕
おくのほそ道	51-芭蕉1694-01023	〔二三〕
おくのほそ道	51-芭蕉1694-01024	〔二四〕
おくのほそ道	51-芭蕉1694-01025	〔二五〕
おくのほそ道	51-芭蕉1694-01026	〔二六〕
おくのほそ道	51-芭蕉1694-01027	〔二七〕
おくのほそ道	51-芭蕉1694-01028	〔二八〕
おくのほそ道	51-芭蕉1694-01029	〔二九〕
おくのほそ道	51-芭蕉1694-01030	〔三〇〕
おくのほそ道	51-芭蕉1694-01031	〔三一〕
おくのほそ道	51-芭蕉1694-01032	〔三二〕
おくのほそ道	51-芭蕉1694-01033	〔三三〕
おくのほそ道	51-芭蕉1694-01034	〔三四〕
おくのほそ道	51-芭蕉1694-01035	〔三五〕
おくのほそ道	51-芭蕉1694-01036	〔三六〕
おくのほそ道	51-芭蕉1694-01037	〔三七〕
おくのほそ道	51-芭蕉1694-01038	〔三八〕
おくのほそ道	51-芭蕉1694-01039	〔三九〕
おくのほそ道	51-芭蕉1694-01040	〔四〇〕
おくのほそ道	51-芭蕉1694-01041	〔四一〕
おくのほそ道	51-芭蕉1694-01042	〔四二〕
おくのほそ道	51-芭蕉1694-01043	〔四三〕
おくのほそ道	51-芭蕉1694-01044	〔四四〕
おくのほそ道	51-芭蕉1694-01045	〔四五〕
おくのほそ道	51-芭蕉1694-01046	〔四六〕
おくのほそ道	51-芭蕉1694-01047	〔四七〕
おくのほそ道	51-芭蕉1694-01048	〔四八〕
おくのほそ道	51-芭蕉1694-01049	〔四九〕

3.5 外部リンク

●底本画像リンク

「中納言」検索画面の底本リンク列において「JK」のロゴをクリックすると Japan Knowledge『新編日本古典文学全集』の該当ページへのリンクが表示される。リンク先では本文，現代語訳，註釈が参照できる。ただし，利用には Japan Knowledge への利用登録が必要である。

●参考画像リンク

収録作品のうち「笈の小文」，「更科紀行」，「おくのほそ道」については，国文学研究資料館（NIJL）鶴飼文庫の影印画像とリンクしている。「中納言」検索画面の参考リンク列において「Nijl」のロゴをクリックすると，NIJL と人文学オープンデータ共同利用センター（Center for Open Data in the Humanities / CODH）とが共同開発しウェブ上で公開している IIF Curation Viewer を通して影印画像が無償で参照できる。

【参考規程・概要書】

池上 尚（2016）『『日本語歴史コーパス 鎌倉時代編 I 説話・随筆』形態論情報の概要』

<https://ccd.ninjal.ac.jp/chj/doc/morph-kamakura-2016.pdf>

国立国語研究所コーパス開発センター（鴻野知暁）編（2017）『『日本語歴史コーパス 鎌倉時代編』短単位規程集 Ver.1.0』

https://ccd.ninjal.ac.jp/chj/doc/morph_kamakura_v1_0.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://ccd.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://ccd.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://ccd.ninjal.ac.jp/chj/doc/morph-at-chikamatsu-2020.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://ccd.ninjal.ac.jp/chj/doc/abstract-share-2019.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』『同 江戸時代編 II 人情本』形態論情報の概要』（2019年3月29日更新）

<https://ccd.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://ccd.ninjal.ac.jp/chj/doc/abstract-ninjo-2019.pdf>

【参考文献】

『日本古典文学大辞典』岩波書店

『新編日本古典文学全集 70 松尾芭蕉集 (1)』小学館

『新編日本古典文学全集 71 松尾芭蕉集 (2)』小学館

【関連する URL】

国立国語研究所 (2021) 『日本語歴史コーパス』バージョン 2021.3

<https://ccd.ninjal.ac.jpchj/>

日本古典籍データセット <http://codh.rois.ac.jp/pmjt/>

IIIF Curation Viewer (人文学オープンデータ共同利用センター)

<http://codh.rois.ac.jp/software/iiif-curation-viewer/>

JapanKnowledge <https://japanknowledge.com/>