『日本語歴史コーパス 鎌倉時代編Ⅲ軍記』テキスト・表示項目について

2022 年 4 月 30 日　片山久留美

１．はじめに

『日本語歴史コーパス 鎌倉時代編Ⅲ軍記』（以下「本コーパス」と呼ぶ）は、『新編日本古典文学全集』41 巻所収『保元物語』および『平治物語』、45・46 巻所収『平家物語』を底本としている。電子化に際して一部テキストを校訂し、そこに様々な情報を付加することで XML データを構築した。本稿ではテキストの校訂状況や検索アプリケーション「中納言」上に表示される各種情報について概要を述べる。

２．テキストの校訂

本コーパスのテキストは、基本的に『新編日本古典文学全集』のテキストをそのまま使用しているが、コーパス化にあたり以下の各項目について校訂を行った。なお、底本に校訂を加えた箇所については「中納言」の「原文文字列」で校訂前の底本のテキストを確認することが可能である。

以下、本文書中で「底本」は『新編日本古典文学全集』のテキストを、「原本」は底本が使用した写本などの原資料を指す。

●外字の処理

底本で使用している文字は、JIS 外字であっても Unicode 内の文字であればそのままコーパス本文に使用した。ただし、以下の 1 例のみ Unicode 外の文字であるため〓で置き換えた。

例）早く烟蓬屋哀〓（阝＋帝）の究衰を忘れ、速やかに三菩提の月を翫ぶ。

（『保元』下巻『新編全集』<41>p.398）

●返読箇所・補読

底本に返り点・ルビが付されていて読みが確定できる漢文等は、訓読した形を本文とした。訓読に際し補読が必要な箇所については、最低限の文字を補った。

例）

（底本）桃李不レ言（ものいは）、春幾（はるいくばく）暮（かくれぬる）。　（『平家』巻第三『新編全集』<45>p.213）

（本コーパスのテキスト）桃李言ず。春幾か暮ぬる。

●踊り字

底本で使用されているいわゆる「くの字点」は仮名に開いた。

例）

（底本）中比は都の住ひも<u>うと〳〵しく</u>　（『平家』巻第一『新編全集』<45>p.22）

（本コーパスのテキスト）中比は都の住ひも<u>うとうとしく</u>

（底本）此事よしなし。壁に耳あり、おそろし<u>〳〵</u>　（『平家』巻第一『新編全集』<45>p.59）

（本コーパスのテキスト）此事よしなし。壁に耳あり、おそろし<u>おそろし</u>

●カタカナ表記箇所

底本には、原本で表記されていない促音・撥音・長音をカタカナ小字で補入している箇所がある。これらのカタカナ表記箇所は、すべて平仮名表記に改めた。

例）

（底本）其御子、高望の王の時、始めて平の姓を給は<u>ッ</u>て（『平家』巻第一『新編全集』<45>p.20）

（本コーパスのテキスト）其御子、高望の王の時、始めて平の姓を給は<u>っ</u>て

（底本）何事も六波羅様といひて<u>ン</u>げれば　（『平家』巻第一『新編全集』<45>p.29）

（本コーパスのテキスト）何事も六波羅様といひて<u>ん</u>げれば

３．「中納言」における表示項目と内容

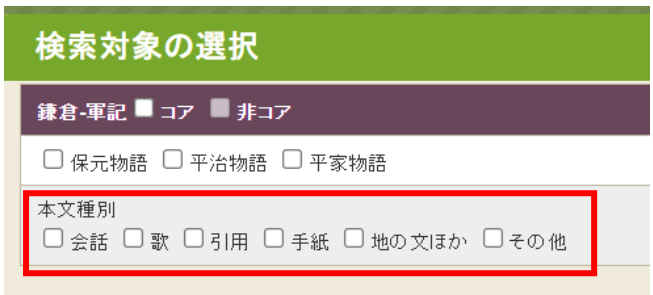
本コーパスの本文にはさまざまなタグや単語情報が付与されており、言語研究を目的とした利用に資するものとなっている。これらの情報はコーパス検索アプリケーション「中納言」上に検索結果として表示される。形態論情報の表示については他のサブコーパスと基本的に変わらないため省略し、以下では本コーパス使用に際して特に注意の必要な「本文情報」「作品情報」などの各列について表示内容の詳細を述べる。

サンプルID	開始位置	連番	コア	前文脈	キー	後文脈	語彙素読み	語彙素	語形	品詞	活用型	活用形	原文文字列	振り仮名	本文種別	話者	ジャンル	作品名	成立年	巻名等	部	作者	生年	底本	ページ番号	底本リンク	参考リンク
30-保元 1223_01011	15060	10260	1	なる|のみ|にあら|ず|、|馬|の|上|、|歩立|、|惣|て|天|を|翔る|翅|、|地|を|	走る	|獣|も|、|目|を|惣ナ|と|惣ナ|つる|もの|を|射|留めず|といふ|事|なし|。#将門	ハシル	走る	ハシル	動詞-一般	文語四段-ラ行	連体形-一般	走る				軍記物語	保元物語	1223	新院御所各門々固めの事仰すたり軍評定の事	上			新編全集 <41>	252	JK	
				めしげなるのみにあらず、馬の上、歩立、惣て天を翔る翅、地を	走る	獣も、目を惣ナと惣ナつるものを射留めずといふ事なし。#将門-純友にも貞任-宗任にも勝れたり。																					
30-保元 1223_02001	37850	25580	1	なり|。#あな|、|いたづら事|や|」#とて|、|長刀|を|取り直し|て|、|先|に|立ち|て|そ|	走り	|ける|。#惟行|、|鹿|廿|八|、|身|の|盛り|と|見え|たり|。#大|の|男|の|したたか者	ハシル	走る	ハシル	動詞-一般	文語四段-ラ行	連用形-一般	走り				軍記物語	保元物語	1223	白河殿へ義朝夜討に寄せらるる事	中			新編全集 <41>	281	JK	
				勲功の事は、命の生きての上の事なり。#あな、いたづら事や」#とて、長刀を取り直して、先に立ちてそ	走り	ける。#惟行、故廿八、身の盛りと見えたり。#大の男のしたたか者、弓は三人張り、矢束は十三束、下針をも射ん																					

3.1. 本文情報

●本文種別

本コーパスでは「中納言」の「検索対象の選択」画面において、「本文種別」を「会話」「歌」「引用」「手紙」「地の文ほか」「その他」の中から選択して検索することができる。



「検索対象の選択」で選べるのは大まかな本文種別の分類であるが、検索結果の「本文種別」の列には下記のような詳細な本文種別情報が表示される。

大分類	中分類	小分類	
会話			登場人物の発話箇所
	引用		発話の中で典拠のある文言を引用している箇所
		韻文	上記のうち典拠が和歌・漢詩などの韻文のもの
	歌		発話の中で自作の歌を詠んでいる部分
	発話引用		発話の中で他者の発言を引用した部分
引用			典拠のある文言を引用している箇所
	韻文		上記のうち典拠が和歌・漢詩などの韻文のもの
手紙			書状、書き付けなど
歌			作中人物の詠んだ和歌
その他	奥書		作品末尾の奥書部分

●話者

本コーパスにおいては、現時点では会話部分に対する話者情報の付与は行っていない。

3.2　作品情報

●ジャンル

一律「軍記物語」と表示される。

●作品名

『保元物語』『平治物語』『平家物語』のいずれかが表示される。

●成立年

各作品の成立年が表示される。なお成立年の決定にあたっては、築島裕ほか（2011）『古語大鑑』収録の「用例出典一覧」を参考にした。

●部
『保元物語』『平治物語』においては上・中・下巻の別を、『平家物語』においては巻一～巻十二および「灌頂巻」といった巻号を表示する。

●巻名等
「部」の下位分類として、各巻の章段名を表示する。章段名は『新編日本古典文学全集』で各章段に付されているものを使用した。

3.3　底本情報
●底本・ページ番号
「底本」には当該例の出現する『新編日本古典文学全集』の巻数、「ページ番号」には『新編日本古典文学全集』でのページ番号が表示される。

3.4　外部リンク
●底本リンク
ジャパンナレッジの『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。なお、利用にはジャパンナレッジの利用登録が必要となっている。

【参考文献・参考 URL】
国立国語研究所（2022）『日本語歴史コーパス　鎌倉時代編Ⅲ軍記』
https://clrd.ninjal.ac.jp/chj/kamakura.html#gunki（2022 年 3 月 31 日確認）
築島裕・編集委員会代表（2011）『古語大鑑　第 1 巻』東京大学出版会