

# 『日本語歴史コーパス 平安時代編Ⅲ漢文系資料』（短単位データ 0.9） 解説書

山本 久（国立国語研究所）・呉 子凡（国立国語研究所）  
2026年3月31日

## 1 はじめに

『日本語歴史コーパス 平安時代編Ⅲ漢文系資料』には、『法華百座聞書抄』『高山寺本古往来』『尾張国解文』のデータを収録する。このうち高山寺本古往来と尾張国解文については、原本に付された訓点（『高山寺本古往来』院政期点、真福寺本『尾張国解文』正中二年点）をコーパスの作成に使用しており、形態論情報の付与された訓点資料のコーパスとして、『日本語歴史コーパス 平安時代編Ⅱ訓点資料』に収録された西大寺本『金光明最勝王経』平安初期点に続くものである。

『日本語歴史コーパス 平安時代編Ⅱ訓点資料』の解説書である柳原・近藤（2022）に記される通り、漢文による原文と、原文を読み下すための訓点から構成される訓点資料には、コーパス化に際して、『日本語歴史コーパス』収録の他資料とは異なる独自の処理を本文に行った点が多い。利用にあたっては、本解説書に目を通し、テキスト作成や形態論情報付与の方法を把握した上で利用することを推奨する。

## 2 資料の概要

コーパスの解説に先立ち、『日本語歴史コーパス 平安時代編Ⅲ漢文系資料』に収録した3資料の概要を説明しておく。

『法華百座聞書抄』は、天仁3（1110）年2月28日から300日間にわたって講じられた法華経・阿弥陀経・般若心経の講説（説教・説話）の聞書である。唯一の伝本である法隆寺本は、300日の講説のうち10人の説教僧による20日分の講説を収めた抄出本であり、カタカナ主体の漢字仮名交じり文で記されている。院政期の話し言葉を伝える資料として重要視されており、古文学祕籍複製会による複製本、勉強社文庫による影印が公開されているほか、小林編（1975）に翻刻が収録されている。

『高山寺本古往来』は、平安後期頃成立の往来物（書簡文例集）であり、中世から近世にかけて隆盛した往来物の中でも最古のものにあたる。全56通の書簡から成り、高山寺所蔵『表白集』の紙背に残された写本が唯一の伝本である。文体は和習を含む漢文体（和化漢文）であり、院政期末～鎌倉初期と推定される豊富な訓点（傍訓・声点・合符・返点）が付されていることから、音韻史や漢文訓読史上の重要資料として利用されてきた。また、書簡の内容から平安時代の国衙行政や農業経営の実態を知ることができ、社会的な観点からも重要な史料とされている。高山寺典籍文書総合調査団編（1972）に影印・翻刻・訓読文が収録されている。

『尾張国解文』は、永延2年（988）11月8日付で、尾張国の郡司・百姓らが国司藤原元命の非法を朝廷に訴えた31ヶ条から成る文書である。文体は随所に和習を含む和化漢文だが、全体として対句を多用した駢儷文である。平安後期の地方行政の実態を伝える史料として著名であるだけでなく、13世紀末から14世紀前半頃の古写本3点（早稲田大学本・東京大学史料編纂所本・真福寺本）が現存しており、いずれも訓点（傍訓・声点・合符・返点、真福寺本のみヲコト点）を含むことから、漢文訓読史上の重要な資料としても知られている。早稲田大学本および東京大学史料編纂所本はカラー画像がインターネット公開されており、真福寺本も愛知県史編さん委員会編（2009）にてカラー写真が公開されているほか、宇都宮（2000）で真福寺本の訓読文が公開されている。また、梅村（2020）や森（2024）に注釈・現代語訳が収録されている。本コーパスでは真福寺本を主本文とし、冒頭の欠部を早稲田大学本で補った。

### 3 コーパスのテキスト

本節と次節では、本コーパスに収録した二種類の本文「コーパステキスト」と「原文文字列」、およびそれに関する諸点について解説する。

本コーパスでは、原本を忠実に翻刻した本文を底本に定め、底本に基づいて作成した漢字ひらがな交じり文をコーパステキストとして表示した。その具体的な方法は以下に述べていくが、コーパステキストの作成方法を一言で述べると、法華百座聞書抄は小林編

(1975)の翻刻文を、高山寺本古往来は高山寺典籍文書総合調査團編(1972)、尾張国解文は宇都宮(2000)に翻刻された訓読文を底本とし、濁点の付加をはじめとする校訂を経たうえで表記を漢字ひらがな交じり文に変換したものである。

#### 3.1 コーパステキスト

以下、説明の必要な点が多い資料である高山寺本古往来を中心として解説し、必要に応じて尾張国解文にも言及する。訓点資料ではない法華百座聞抄のテキストおよび原文文字列については3.5節で解説する。また、尾張国解文のうち、依拠した原本が異なる内題・副題・第一条の前半については3.4節で解説する。

3.1節では、4節で述べる訓読方針に沿って作成した訓読文を、コーパステキストとして表示するための方法について述べる。参考のため、本解説書で用いる本文の名称と意味について以下に整理しておく。

- ・ 原本…古写本原本、またそれに記された本文。
- ・ 底本…書籍や論文雑誌において既に公開されている、原本を忠実に翻刻した本文。具体的には、法華百座聞書抄は小林編(1975) p3~159の本文、高山寺本古往来は高山寺典籍文書総合調査團編(1972) p123~169の本文、尾張国解文は宇都宮(2000)の本文がそれにあたる。本コーパスのコーパステキストおよび原文文字列は、これらの翻刻された本文を元としていることから、底本と呼ぶ。
- ・ コーパステキスト…形態素解析の対象として作成し、形態論情報を付与した本文。コーパス検索アプリケーション「中納言」の検索対象であり、「中納言」の検索結果に表示される本文である。以下、単に「テキスト」と呼ぶ際はこのコーパステキストを指す。
- ・ 原文文字列…「中納言」での検索時に原本の加点状況を知ることができるよう、原本を忠実に訓読し翻刻した本文。原本を忠実に再現した訓読文である点で底本と同様であり、本コーパスの原文文字列も基本的には上記の底本に基づいているが、再現の方法には一部異なる点もある(後述)。「中納言」の検索結果では「原文文字列」欄に表示されるほか、「前文脈」「キー」「後文脈」欄の下段(原文 KWIC)にも前後の文脈付きで表示される。
- ・ 訓読文…原本の漢文を日本語の語順に表記し直し、振り仮名の推読や送り仮名・付属語の補読などを行って、日本語文として理解できるようにしたもの。本解説書では翻刻・テキスト化に至る前の抽象的な本文のことを指す。高山寺本古往来と尾張国解文における「底本」「コーパステキスト」「原文文字列」は、いずれも抽象的な訓読文を翻刻・テキスト化という方で具体化させたものといえる。

##### 3.1.1 異訓・異本注記の扱い

原本では、漢文中の漢字1字(または漢字文字列)に複数の加点がなされている場合があり、底本でも複数の加点が再現されていることから、複数の訓読文を作成することが可能である。コーパステキストでは、右傍訓を優先的に採用し、右傍訓に誤りがあると思われる場合や、左傍訓に基づいた方が自然な訓読文を作成できると考えた場合にのみ左傍訓を採用するという方針を採った。採用しなかった異訓をテキストでは参照できないことに

注意されたい。

また、尾張国解文の原本には、全編にわたって、漢文中の漢字に対する異本注記（「イ〜」）がみられる。これについても、漢文本行に示された漢字を優先的に採用した。

### 3.1.2 補読

原本には全体にわたって加点が行われており、特に高山寺本古往来の加点は非常に密に行われている。しかし、形態論情報付与のための形態素解析に適し、かつ語形や文の区切りの理解が容易であるテキストを作成するにあたっては、加点にはない送り仮名・付属語・句読点・振り仮名を補う必要がある。そこで、必要最小限の補読を行った訓読文によるテキストを作成した。以下にその補読の例を示す。どのような送り仮名や付属語を補読するかという点については基本的に底本を踏襲している（3.3.4節も参照）。

#### 【例】

〈原本〉	空 <sup>シク</sup> 過 <sup>テ</sup>
〈底本〉	空 <sup>(ムナ)</sup> シク過 <sup>(ス)</sup> <u>(キ)</u> テ
〈コーパステキスト〉	空しく過 <sup>ぎ</sup> て (20L 高山 1186_00001, 760, 530)

〈原本〉	必 <sup>キ</sup> 招 <sup>マネク</sup> 嘲 <sup>チヨウロウ</sup> 弄 <sup>ヲ</sup> 歎 <sup>歎</sup>
〈底本〉	必 <sup>(カナラ)</sup> <u>(ス)</u> 嘲 <sup>チヨウロウ</sup> 弄 <sup>ヲ</sup> 招 <sup>マネク</sup> 可 <sup>(ヘ)</sup> キ歎 <sup>(カ)</sup>
〈コーパステキスト〉	必 <sup>ず</sup> 嘲 <sup>弄</sup> を招 <sup>く</sup> 可 <sup>き</sup> 歎 (20L 高山 1186_00005, 1010, 680)

〈原本〉	年と結 <sup>ケ</sup> 解 <sup>ケ</sup>
〈底本〉	年と <sup>(ネンノ)</sup> <u>(ノ)</u> 結 <sup>(ケツ)</sup> 解 <sup>ケ</sup>
〈コーパステキスト〉	年年 <sup>の</sup> 結 <sup>解</sup> (20L 高山 1186_00015, 100, 60)

### 3.1.3 仮名点

3.3.2節でも述べる通り、原本の仮名点やヲコト点は、基本的にコーパステキストに忠実に反映している。その中で、説明を要すると思われる点のみ以下に示す。

#### 3.1.3.1 音便形

高山寺本古往来の原本には、音便形に対して、多様なながらもある程度の一貫性を持った表記が行われていることが知られている（肥爪 2023 など）。コーパステキストでは、底本を踏襲し、原本の表記を可能な限り尊重する形で、次のような表示を行った。

- ・「ム」表記→「む」

#### 【例】

〈原本〉	慣 <sup>タム</sup> テ
〈底本〉	慣 <sup>(ノ)</sup> <u>ム</u> テ
〈コーパステキスト〉	慣 <sup>む</sup> で (20L 高山 1186_0015, 2940, 2090)

※「慣む」（語彙素：頼む）の活用形は「連用形-撥音便」とする。

- ・特殊符号（「ン」）表記→「ん」

#### 【例】

〈原本〉	可 <sup>シ</sup> 足 <sup>タシ</sup> ヌ
〈底本〉	足 <sup>タシ</sup> ヌ可 <sup>(ヘ)</sup> シ
〈コーパステキスト〉	足 <sup>ん</sup> ぬ可 <sup>し</sup> (20L 高山 1186_00040, 1410, 980)

※「足ん」（語彙素：足りる）の活用形は「連用形-撥音便」とする。

- ・「ツ」表記→「つ」

#### 【例】

〈原本〉	追 <sup>ヲツ</sup> テ
------	-------------------

〈底本〉 追<sup>ヲ</sup>ツテ  
〈コーパステキスト〉 追<sup>つ</sup>て (20L 高山 1186\_00009, 650, 480)

※「追ツ」(語彙素：追う)の活用形は「連用形-促音便」とする。

・零表記→(表記せず)

【例】

〈原本〉 進<sup>タテマツテ</sup>  
〈底本〉 進<sup>タテマツテ</sup>  
〈コーパステキスト〉 進<sup>て</sup> (20L 高山 1186\_00014, 1220, 920)

※「進」(語彙素：奉る)の活用形は「連用形-促音便」とする。

### 3.1.3.2 仮名遣いの異例

原本には、歴史的仮名遣いに相違する仮名遣いがみられる。コーパステキストでも底本を踏襲し、原本の表記を基本的にそのまま表示した。

【例】

〈原本〉 心事<sup>ヲ</sup>期<sup>コス</sup>面<sup>ヲ</sup>  
〈底本〉 心事<sup>(シムシ)</sup> ヲ面<sup>(メン)</sup> ヲ期<sup>コス</sup>  
〈コーパステキスト〉 心事<sup>わ</sup>面<sup>を</sup>期<sup>す</sup> (20L 高山 1186\_00033, 1260, 860)

### 3.1.4 原漢文の文意と加点の解釈が異なる場合

訓点資料は、原漢文を訓読文の形式となるように解説し、それに即した読みの手がかりをヲコト点や仮名点で記したものである。したがって、稀に解釈に誤りが生じ、原漢文から自然に理解できる文意とは異なる読みが加点されることがある。そのような場合には、訓点を優先して加点の通りに訓読文を作成した。コーパステキストの中に文脈的に不自然なものが見受けられた場合には、加点誤りの可能性も想定して原本を確認することが望ましい。

【例】

〈原本〉 无<sup>ク</sup>馬<sup>ノ</sup>煩<sup>ワツライ</sup>侍<sup>リ</sup>  
〈底本〉 馬<sup>(ウマ)</sup> <sup>ノ</sup>煩<sup>ワツライ</sup>无<sup>(ナ)</sup> <sup>ク</sup>侍<sup>(ハムヘ)</sup> <sup>リ</sup>  
〈コーパステキスト〉 馬<sup>の</sup>煩<sup>い</sup>无<sup>く</sup>侍<sup>り</sup> (20L 高山 1186\_00040, 580, 380)

※本例は、競馬で用いる馬を用意する必要のある差出人が、受取人に馬の貸出を依頼する文脈で用いられている。このような文脈上「馬無く、煩い侍り」という訓読文が期待されるが、「馬」字の右傍に「ノ」の訓があるため、加点者は「馬の煩い無く侍り」と訓読している。原本に加点があることを尊重し、加点に忠実な訓読文を作成した。

## 3.2 原文文字列

3.1節に示した例からも知られるように、コーパステキストとして表示される訓読文では、テキストのどの部分が原本の訓点と対応し、どの部分が訓読文作成者による補読なのかという情報を示していない。一方で、言語研究においては加点状況に関する情報を参照できるようにすることも必要である。

そこで、コーパス検索ツール「中納言」の検索結果に表示される「原文文字列」において、原本での加点状況を識別できる本文を表示することとした。これにあたって、原文文字列の表示方法には、底本などで採用されている現今の訓点資料研究の方法とは一部異なる点があることに注意されたい。「中納言」の検索結果では、原文文字列欄の他に、「原文 KWIC」において原文文字列の表示法による前後文脈付き用例が表示される。

### 3.3 コーパステキストと原文文字列の表示法

コーパステキストと原文文字列の表示法について、対比しながら以下に説明する。なお、原文 KWIC の表示法は原文文字列の表示法と同様である。

### 3.3.1 漢文本文

原本の漢文本文で使用されている漢字は、底本を踏襲し、コーパステキスト・原文文字列ともにその漢字で示した。漢字の電子化の詳細については3.5節を参照されたい。

ただし、訓読する際に不読字とするべき漢字については、コーパステキストには表示せず、原文文字列では該当する位置に表示した。この際、底本をはじめとして現今の訓点資料研究で行われている、亀甲括弧〔〕で括る処理は行っていない。

#### 【例】

〈原本〉	依 <sup>テ</sup> 志 <sup>ノ</sup> 之 <sup>切</sup> ナルニ
〈底本〉	志 <sup>(コ、ロサシ)</sup> ノ〔之〕切 <sup>(セツ)</sup> ナルニ依 <sup>(ヨ)</sup> テ
〈コーパステキスト〉	志の切なるに依て (20L 高山 1186_00022, 3000, 2160)
〈原文文字列〉	志 <sup>(コ、ロサシ)</sup> ノ之 <sup>切</sup> ナルニ依 <sup>(ヨ)</sup> テ

「不(ズ)」「可(ベシ)」のように、日本語の助詞・助動詞に相当する漢字については、高山寺本古往来のコーパステキストには漢字のまま表記し、尾張国解文のコーパステキストにはひらがなで表記している。資料間で処理が異なることに留意されたい。原文文字列では原本通り漢字で表記している。

#### 【例】

〈原本〉	不可延引 <sup>インス</sup>
〈底本〉	延 <sup>(エン)</sup> 引 <sup>インス</sup> 可 <sup>(ヘ)</sup> (カラ)不 <sup>(ズ)</sup>
〈コーパステキスト〉	延引す可から不 (20L 高山 1186_00030, 1590, 1170)
〈原文文字列〉	延 <sup>(エン)</sup> 引 <sup>(イン)</sup> ス可 <sup>(ヘ)</sup> (カラ)不 <sup>(ズ)</sup>

#### 【例】

〈原本〉	敢 <sup>テ</sup> 不言 <sup>モノイハ</sup>
〈底本〉	敢 <sup>テ</sup> 言 <sup>モノイハ</sup> 不
〈コーパステキスト〉	敢て言はず (20L 尾張 1325_00201, 3000, 2110)
〈原文文字列〉	敢 <sup>(モノイ)</sup> 言 <sup>ハ</sup> 不

原本の漢文本文に誤字があると判断され、かつ形態論情報付与に困難が生じる程度のものである場合、コーパステキストでは校訂した本文を示し、原文文字列では原本の漢字をそのまま表示した<sup>1</sup>。校訂の方法については底本に従うことを基本としている。

#### 【例】

〈原本〉	雖 <sup>シ</sup> 補 <sup>ツクノイ</sup>
〈底本〉	補 <sup>ツクノイ</sup> 雖 <sup>〔難〕</sup> (カタ)シ
〈コーパステキスト〉	補い難し (20L 高山 1186_00017, 2770, 2000)
〈原文文字列〉	補 <sup>(ツクノ)</sup> イ雖 <sup>(カタ)</sup> シ

また、尾張国解文の原本には、本文の右傍などに脱字を挿入した箇所が全編にわたってみられる。これに対しては、挿入箇所が助辞1字のみである場合など、脱字による形態論情報の付与や本文の文脈理解にそれほど困難が生じないと判断した場合を除き、脱字を補った本文をコーパステキストとして示すことを基本とした。テキストに脱字を補った場合、補わなかった場合とも、原文文字列では底本に倣って挿入箇所を《補》～として示している<sup>2</sup>。

<sup>1</sup> 原文文字列では校訂を行わない。ただし、高山寺本古往来の原文には不読字に対する誤字とみられる箇所があり、これに限っては、期待される漢字を表示しなければコーパステキストと原文文字列の整合的な理解に困難が生じる恐れがあるため、原文の漢字の直後に期待される漢字を〔〕に括って表示した。

<sup>2</sup> 仕様上、補った箇所を「中納言」で検索した際の「原文文字列」欄は、最初の形態素は空欄で表示さ

### 【例】

〈原本〉	政 <sup>コト</sup> 濁 <sup>リ</sup> 涙 <sup>タ</sup> 澄 <sup>テ</sup>
〈底本〉	政 <sup>(マツリ)</sup> コト濁 <sup>リ</sup> 涙 <sup>(ナミ)</sup> タ <u>[《補》澄(ミ)て]</u>
〈コーパステキスト〉	政濁 <sup>リ</sup> 涙澄 <sup>みて</sup>
〈原文文字列〉	政 <sup>(マツリ)</sup> コト濁 <sup>リ</sup> 涙 <sup>(ナミ)</sup> タ <u>《補》澄(ミ)て</u>
〈原本〉	彼憚 <sup>テ</sup> 權 <sup>-</sup> 公 <sup>カ</sup> 之 <sup>ニ</sup> 威 <sup>ニ</sup>
〈底本〉	彼(ノ)權 <sup>-</sup> 公 <sup>カ</sup> <u>[《補》之]</u> 威 <sup>ニ</sup> 憚 <sup>(リ)</sup> テ
〈コーパステキスト〉	彼の權公が威に憚りて
〈原文文字列〉	彼(ノ)權公カ <u>《補》之</u> 威 <sup>ニ</sup> 憚 <sup>(リ)</sup> テ

### 3.3.2 仮名点・ヲコト点

高山寺本古往來の原本にはヲコト点は使用されておらず、助詞「テ」「ニ」なども含めてカタカナによる仮名点が施されている。仮名点は、コーパステキストでは平仮名で、原文文字列では原本通り片仮名で示した。

### 【例】

〈原本〉	故 <sup>コト</sup> サラニ不 <sup>ス</sup> 參 <sup>候</sup> セ
〈底本〉	故 <sup>コト</sup> サラ <u>ニ</u> 參 <sup>(サムコウ)</sup> <u>セ</u> 不 <sup>ス</sup>
〈コーパステキスト〉	故 <sup>ニ</sup> 參 <sup>候</sup> <u>せ</u> 不 <sup>ス</sup> (20L 高山 1186_00034, 140, 110)
〈原文文字列〉	故 <sup>(コト)</sup> サラ <u>ニ</u> 參 <sup>(サムコウ)</sup> <u>セ</u> 不 <sup>(ス)</sup>

一方、尾張国解文にはヲコト点を使用されており、底本はヲコト点をひらがなで表示している。これに対し、本コーパスではヲコト点についても、仮名点と同様に、コーパステキストではひらがなで、原文文字列ではカタカナで示した<sup>3</sup>。

### 【例】

〈原本〉	請 <sup>被</sup> と <sup>任</sup> て <sup>-</sup> 官 <sup>-</sup> 府 <sup>の</sup> 旨 <sup>に</sup> <u>裁</u> <u>断</u> <sup>下</sup> <u>イ</u>
〈底本〉	官 <sup>-</sup> 府 <sup>の</sup> 旨 <sup>-</sup> <u>に</u> <u>任</u> <u>ニ</u> (セ) <u>て</u> <u>裁</u> <u>断</u> <sup>下</sup> <u>イ</u> (セ) 被 <sup>(レム)</sup> と請 <sup>(フ)</sup>
〈コーパステキスト〉	官符 <sup>の</sup> 旨 <sup>に</sup> <u>任</u> <u>せて</u> <u>裁</u> <u>断</u> <u>せ</u> <u>られ</u> <u>む</u> と請 <sup>ふ</sup> (20L 尾張 1325_00202, 50, 40)
〈原文文字列〉	官府 <sup>ノ</sup> 旨 <sup>ニ</sup> <u>任</u> (セ) <u>テ</u> <u>裁</u> <u>断</u> (セ) 被 <sup>(レ)</sup> (ム) <u>ト</u> 請 <sup>(フ)</sup>

また、尾張国解文の原文には、全編にわたって非常に多くの左右両訓がみられる。

3.1.1 節で述べた通り、コーパステキストの作成にあたっては右傍訓を中心として単一の訓による訓読文を作成したが、採用しなかった方の傍訓も原文文字列上では《左》～などとして示したことがある。

### 3.3.3 濁点

原本では清音の訓点を使用され、底本でも訓点を清音として示しているが、コーパステ

---

れ、その次の形態素に「《補》(最初の形態素)(その次の形態素)」として表示されることに注意されたい。

<sup>3</sup> ただし、形態素解析に直接関わらないヲコト点はひらがなで示したことがある。具体的には、尾張国解文の原本には、助詞を仮名点とヲコト点の両方で表示した箇所が一定数みられる。このような場合、形態素解析の対象となるテキストでは仮名点のみを採用して訓読を行った一方、原文文字列では仮名点とヲコト点の両方を表示し、その際のヲコト点をひらがなで示している。いずれにせよ原文文字列上でのヲコト点の表示方法は徹底しているとはいいがたいので、研究の目的によって仮名点とヲコト点を区別することが必要な場合には、底本の本文、さらには上述の影印・画像を参照してほしい。

キストでは濁音が期待される部分を濁点付きの仮名で示した。読みの候補として清濁両形ある場合は、より古形とされる形や連濁を起こす前の形を選んだ。原文文字列では底本と同様に清音の仮名で表記した。

【例】

〈原本〉	為 <sup>ニ</sup> 充 <sup>アテムカ</sup>	
〈底本〉	充 <sup>アテムカ</sup> 為 <sup>(タメ)</sup> ニ	
〈コーパステキスト〉	充てむが <sup>ニ</sup> 為 <sup>ニ</sup>	(20L 高山 1186_00003, 90, 60)
〈原文文字列〉	充 <sup>(ア)</sup> テムカ <sup>ニ</sup> 為 <sup>ニ</sup>	

唯一の例外として、高山寺本古往来の次の箇所には、漢字の振り仮名(3.3.8節)として用いられた仮名に濁声点が付されている。この例に対しては、原文文字列の振り仮名を濁音「ガ」とした。

【例】

〈原本〉	藤葛 <sup>カ(平濁)ツラ</sup>	
〈原文文字列〉	藤 <sup>(フチ)</sup> 葛 <sup>(ガツラ)</sup>	(20L 高山 1186_00029, 11640, 8580)

### 3.3.4 補読

3.1.2節に述べたように、コーパステキストでは、底本を踏襲する形で原本の加点とは対応しない送り仮名・付属語・句読点・振り仮名を適宜補い、平仮名で示した。原文文字列では、やはり底本と同様、これを全角丸括弧 ( ) に括って示した。

〈原本〉	雖 <sup>トモ</sup> 賜 <sup>タマフ</sup> 拜 <sup>ハイ</sup> 謁 <sup>エツヲ</sup>	
〈底本〉	拜謁 <sup>ハイエツヲ</sup> 賜 <sup>タマフ</sup> (ト) 雖 <sup>(イヘ)</sup> トモ	
〈コーパステキスト〉	拜謁を賜ふと雖も	(20L 高山 1186_00001, 120, 80)
〈原文文字列〉	拜 <sup>(ハイ)</sup> 謁 <sup>(エツ)</sup> ヲ賜 <sup>(タマ)</sup> フ (ト) 雖 <sup>(イヘ)</sup> トモ	

この例からもわかるように、コーパステキストには補読の有無に関する情報が付与されていない。よって、特に付属語や送り仮名の認定にあたっては、原本の訓点をもとに認定したのか、訓読文作成者の判断で認定したかのをテキストのみから区別することはできない。本コーパスにおける付属語や送り仮名の表記について検討する際には、「原文文字列」の表示を参照し、原本における訓点の有無を確認してほしい。

なお、テキストの作成に際して、どのような付属語や送り仮名を補読するかという点については、底本を踏襲することを原則としている。ただし、底本を踏襲しても形態素解析や語形・文の区切りの理解になお困難が生じる場合などに、底本では補読されていない付属語や送り仮名等をテキストでは補読した場合がある。この場合、もともと底本で補読されていた文字を ( ) で括って原文文字列上に表示する一方、テキストでのみ補読した文字はそもそも原文文字列上に表示しないことによって、可能な限り原文文字列と底本とが近い本文になるようにしている(振り仮名については3.3.9節も参照)。

【例】

〈原本〉	不 <sup>ス</sup> 遂 <sup>トケ</sup> 其 <sup>ノ</sup> 思 <sup>ヲ</sup>	
〈底本〉	其 <sup>(ソ)</sup> ノ思 <sup>(オモヒ)</sup> ヲ遂 <sup>トケ</sup> 不 <sup>ス</sup>	
〈コーパステキスト〉	其の思 <sup>ひ</sup> を遂げ不	(20L 高山 1186_00001, 610, 440)
〈原文文字列〉	其 <sup>(ソ)</sup> ノ思 <sup>(オモヒ)</sup> ヲ遂 <sup>(ト)</sup> ケ不 <sup>(ス)</sup>	

※コーパステキストでは「思ひ」の送り仮名「ひ」を補読したが、原文文字列では底本と同様に補読を行わず、「思」字の推読の振り仮名として表示した。

以上に示してきた例からも知られるように、原文文字列の本文は、基本的に底本の本文とほぼ同一のものである。ただし、前述の通り不読字(3.3.1節)とヲコト点(3.3.2節)の表示方法が異なるほか、関連する先行研究を勘案して例外的に底本の訓読方法とは異なる訓読方法でテキストを作成した場合など、ごくまれに底本と原文文字列の訓読文が

異なることもある(3.4節も参照)。原本の表記を正確に知りたい場合には、底本の本文、さらには上述の影印・画像を参照してほしい。

### 3.3.5 踊り字・くの字点・同の字点

原本の漢文本文ならびに訓点で使用される踊り字・くの字点・同の字点は、底本では原本通りに翻刻されているが、コーパステキストではその踊り字が示す漢字または仮名に置き換えて示した。原文文字列では踊り字のまま表示した。

#### 【例】

〈原本〉	被 <sup>ラ</sup> ル <sup>ハ</sup> 差 <sup>シ</sup> 定 <sup>メ</sup>
〈底本〉	差 <sup>(サ)</sup> シ定 <sup>(サタ)</sup> (メ) 被 <sup>ラ</sup> ル <sup>ハ</sup>
〈コーパステキスト〉	差し定め被 <sup>ル</sup> る (20L 高山 1186_00005, 170, 110)
〈原文文字列〉	差シ定 (メ) 被 <sup>(ラ)</sup> <u>ルハ</u>

〈原本〉	恐 <sup>ト</sup> 謹言
〈底本〉	恐 <sup>ト</sup> 謹言 (クキヨウ / \ キンケン)
〈コーパステキスト〉	恐恐謹言 (20L 高山 1186_00002, 3260, 2360)
〈原文文字列〉	恐 <sup>ト</sup> <u>ト</u> (クキヨウ / \) 謹言 (キンケン)

### 3.3.6 句読点

原本には句切点等は使用されていない。基本的には底本を踏襲する形で句点・読点を補い、コーパステキスト・原文文字列ともに句点を「。」、読点を「、」で示した。

### 3.3.7 声点・合符・返点

高山寺本古往来と尾張国解文の原本には全編にわたって声点・合符・返点が付されており、底本にも反映されている。これらの符号は、漢語語形の認定や和訓読み／字音読みの認定、訓読文の作成のために最大限利用したが、コーパステキストおよび原文文字列には、1例の例外(3.3.3節)を除いて反映しなかった。

### 3.3.8 加点誤り

原本の加点に誤りがあると判断され、かつ形態論情報付与に困難が生じる程度のものである場合、コーパステキストでは校訂した本文を示した。原文文字列では、原本の訓点のままの本文を示した。校訂の方法については底本に従うことを基本としている。

#### 【例】

〈原本〉	寒 <sup>ハ</sup> 来 <sup>リ</sup> 暑 <sup>ソ</sup> 往 <sup>ユ</sup> イテ
〈底本〉	寒 <sup>(カン)</sup> ハ来 <sup>(キタ)</sup> リ暑 <sup>(シヨ)</sup> <u>ソ</u> (ハ) 往 <sup>(ユ)</sup> イテ
〈コーパステキスト〉	寒は来り暑は往いて (20L 高山 1186_00049, 520, 360)
〈原文文字列〉	寒 <sup>(カン)</sup> ハ来 <sup>(キタ)</sup> リ暑 <sup>(シヨ)</sup> <u>ソ</u> 往 <sup>(ユ)</sup> イテ

### 3.3.9 振り仮名

原本の訓点のうち、訓読文で漢字の振り仮名に相当すると考えられるものは、コーパス検索ツール「中納言」の検索結果において「振り仮名」項や「原文 KWIC」の「キー」の振り仮名として示すこととした。また、当該語形を形態論情報の付与に際しての語形認定(4.1節参照)のために最大限利用している。

#### 【例】

〈原本〉	如 <sup>キノ</sup> 此之輩 <sup>トモカラ</sup>
〈底本〉	此 <sup>(カク)</sup> (ノ) 如 <sup>(コト)</sup> キノ [之] 輩 <sup>トモカラ</sup>
〈コーパステキスト〉	此の如きの輩 (20L 高山 1186_00014, 1750, 1270)
〈原文文字列〉	此 <sup>(カク)</sup> (ノ) 如 <sup>(コト)</sup> キノ之輩 <sup>(トモカラ)</sup>

※「輩」の語形は「トモガラ」とする。

補読(3.3.4節)と同様に、テキストには振り仮名の情報が付与されていないため、ある漢字の読みの認定にあたって、振り仮名をもとに認定したのか、訓読文作成者の判断で認定したかのをテキストのみから区別することはできない。本資料における漢字の読みを検討する際には、「原文文字列」の表示を参照し、原本における振り仮名の有無を確認してほしい。

なお、高山寺本古往来の底本である高山寺典籍文書綜合調査團編(1972)では、原本に振り仮名の存しない漢字のほぼ全てに対し、原本に存する振り仮名と区別した形で推読による振り仮名を表示している。本コーパスでもこれに倣い、高山寺本古往来の原文文字列においては、推読による振り仮名を( )に括った形で積極的に表示している。

本コーパスでテキストに対して付与した語形(4.1節を参照)と底本の推読にはまれに異なる箇所があるが、その場合でも、参考に供するために原文文字列上の振り仮名には底本の推読を示した場合がある。本コーパスの利用者は、原本における振り仮名の有無を確認したうえで、原本に振り仮名が存しない場合には、テキストに付与された語形と原文文字列に表示された推読による振り仮名とを比較・吟味し、適宜取捨選択を加えてほしい。

### 3.4 尾張国解文の内題・副題・第一条の前半について

尾張国解文について、本コーパスで主底本とした真福寺本は、内題・副題・第一条の前半部分を欠いている。底本の宇都宮(2000)は、この欠部に相当する早稲田大学本の本文を、訓点を省略した白文で示しているが、本コーパスでは早稲田大学本の訓点に基づいて独自の訓読文を作成し、これをコーパステキストとした。

補填箇所のコーパステキストと原文文字列の作成方法について、上記の説明と異なる点はない。ただし、補填箇所には原文文字列の元となる底本の訓読文が存在しないという関係上、この箇所の原文文字列では( )に括る形での補読表示を行わず、原本の漢文と仮名点のみを日本語の語順通りに表示するという例外的な形式をとっている。

#### 【例】

〈原本〉	領 <sup>テ</sup> 能田 <sup>ヲ</sup> 以不請正税 <sup>ヲ</sup>
〈底本〉	領能田以不請正税
〈コーパステキスト〉	能田を領りて以て正税を請けず (20L尾張 1325_00101, 1600, 1170)
〈原文文字列〉	能田ヲ領テ以正税ヲ請不

### 3.5 法華百座聞書抄のコーパステキストおよび原文文字列

法華百座聞書抄のコーパステキストおよび原文文字列については、底本に基づいて作成した漢字ひらがな交じり文をテキストとして用い、原文文字列には原本の表記に従ったカナタナ交じり文を表示した。

原本の漢字はテキスト・原文文字列ともにその漢字で表示し、原本のカタカナはテキストではひらがなで、原文文字列ではカタカナで表示した。また、原本の仮名遣いに対する処理、誤字に対する校訂、濁点の処理、踊り字の処理、句読点の処理については、上記と同様の方針で行っている(3.1.3.2節、3.1.4節、3.3.3節、3.3.5節、3.3.6節を参照)。

法華百座聞書抄に特有の問題として、原本に相当数の欠損箇所があるという問題がある。小林編(1975)の翻刻文は、残画を根拠とした推読字には「△」、文脈に基づく推読字(および脱字箇所に対する推読字)には「×」印を付している。これに対して、本コーパスでは、テキストでは「△」「×」ともに推読字をそのまま表示し、原文文字列では「×」字のみを( )で括る形で表示した。

#### 【例】

〈底本〉	コ△ノ×ヒト文字カキヲヘテ
〈コーパステキスト〉	このひと文字かきをへて (20L 法華 1110_00005, 2430, 1560)
〈原文文字列〉	コ (ノ) ヒト文字カキヲヘテ

### 3.6 使用する文字集合

テキストの電子化で使用する文字集合は、「JIS X 0213」に準拠した。原本の仮名点で  
使用される変体仮名は、現行の仮名字体で電子化した。

漢字については、JIS の包摂規準を適用し文字集合内の文字で電子化した。JIS の包摂  
規準を適用できない漢字字体については、学術的な知見・慣例に基づいて現行の字体（新  
旧の別のあるものは旧字体）に置き換えて入力した。以上の処理を行ってもなお電子化で  
できない漢字は、ユニコード番号がある場合にはユニコード番号で表示し、ユニコード番号  
がない場合には「=」（げた記号、JIS 面区点面区点 1-02-14、U+3013）で入力すること  
とした。

以上のように、原本の漢字字体の詳細な情報はコーパス上では示されないため、字体  
の確認の必要な場合は影印や画像等を参照されたい。

## 4 コーパステキストの語形（読み）の認定基準

訓点資料の訓読文に形態論情報を付与するには、漢文本文部分を含めたすべての語の語  
形（読み）を特定しなければならない（漢字かな交じり文における漢字表記の部分につい  
ても同様である）。高山寺本古往来と尾張国解文の原本には全体にわたって仮名点による  
加点があるが、加点のない部分については、以下に述べる基準に沿って語形を認定した。

### 4.1 基本的な語形認定基準

- (1) 法華百座聞書抄と高山寺本古往来については、底本または底本に対応した語彙索引  
において語形の認定が積極的になされているので、これを利用する。ただし、国立  
国語研究所（池上尚）編（2016）に語形の認定に関する規定がある箇所は、これを  
優先する。
- (2) 法華百座聞書抄と高山寺本古往来において語形認定に検討の余地が残るとされる  
箇所、および尾張国解文において加点のない箇所については、以下の手順で読みが  
定まれば、これに従って和訓または字音で読む。
  - ① 当該資料内に、和訓による加点例または字音による加点例、そのほか当該語の語形  
が確定できる例があればそれによる。また、声点が付された例があれば字音語とし、  
声点によって清濁や呉音／漢音を判断する。合符があれば訓合符／音合符の別を和訓  
読み／字音読みの判断に参照する。ただし、同じ形態素に属する複数の語形につい  
て、資料内に加点例のある語形よりも明らかに一般的と思われる語形が存在する場  
合などに、②③の手続きに基づいて、加点のある部分以外には一般的な語形の方を採用  
したことがある。
  - ② 『色葉字類抄』『類聚名義抄』など時代の近い（院政～鎌倉時代）古辞書の記載、  
『和泉往来』『庭訓往来』など性質の近い（和化漢文）訓点資料での加点例、『平安  
遺文』『鎌倉遺文』所収の仮名文書の実例を参照する。
  - ③ 『日本国語大辞典 第二版』での記載、『日本語歴史コーパス』収録の他のサブコー  
パス、「資料横断的な漢字音・漢語音データベース」などのデータベースにより、和  
訓または字音読みの語の確例を参照する。また、底本に示された推読を参照する。
- (3) (1) (2) で読みが確定しなかった箇所は、字音で読む。その際には、沖森卓  
也・三省堂編修所編（2014）『五十音引き漢和辞典』記載の漢音で読む。

### 4.2 特別な語形認定基準

3.1.3.1 節で高山寺本古往来の例を示した通り、音便形を表記したと認められる箇所については、その表記に即した音便形で読む。音便形・非音便形の両方の読みの可能性がある箇所については、法華百座聞書抄・高山寺本古往来は底本の音便形／非音便形の判断に従い、尾張国解文は「以（もつ）て」「於（おい）て」を音便形で、それ以外を非音便形で読む。

## 5 形態論情報

本コーパスでは、すべてのコーパステキストに対して、国立国語研究所（池上尚）編（2016）に拠って、短単位の形態論情報（語に関する情報）を付与している。長単位の形態論情報は付与していない点に留意されたい。形態論情報は、MeCab 用形態素解析用辞書 UniDic（小木曾・小町・松本 2013）による形態素解析結果を、人手で修正して付与した。

## 6 コーパス検索ツール「中納言」における表示項目と内容

コーパステキストに付した形態論情報等は、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・省略しながら本コーパスでの仕様を示す。

### 6.1 時代名・サブコーパス名

時代名は「2 平安」、サブコーパス名は「平安-漢文系資料」と表示する。

### 6.2 サンプル ID

既存の『日本語歴史コーパス』で用いられている統一的なルールに従ってサンプル ID を付与した。

#### ・サンプル ID 付与のルール

1-2 桁目	時代
3 桁目	ジャンル
4-5 桁目	作品 ID
6-9 桁目	成立時期
10 桁目	区切り番号
11-15 桁目	作品内での出現通し番号

1-2 桁目の時代は、すでに公開されている他の「平安時代編」収録コーパスに従い「20」とした。

3 桁目のジャンルは、「平安時代編Ⅱ 訓点資料」の「K」と区別するために「L」とした。

4-5 桁目の作品 ID は、資料名の頭 2 文字を使用し、法華百座聞書抄は「法華」、高山寺本古往来は「高山」、尾張国解文は「尾張」とした。

6-9 桁目の成立時期は、次のように設定した。

- ・法華百座聞書抄…講説の開催年次（天仁 3=1110 年）から、1110。
- ・高山寺本古往来…加号年代の推定（院政期～鎌倉初期頃）から、1186。
- ・尾張国解文…真福寺本の加号年次（正中 2=1325 年）から、1325。

11-15 桁目の出現通し番号は、次のように設定した。

- ・法華百座聞書抄…講説の開催日および各日の構成（説教／説話）に基づき、00001 から 00084 までの通し番号を付与。
- ・高山寺本古往来…書簡 56 通に対して 00001 から 00056 までの通し番号を付与。
- ・尾張国解文…11-13 桁目は、本文と原本の種類により、「内題」「副題」（いずれも

早稲田大学本)は0、本文のうち早稲田大学本補填箇所は1、本文(「後文」含む)は2、「奥書」「巻首見返後筆」は3とする。14-15桁目は、「内題」「副題」はそれぞれ1と2、本文のうち「早稲田大学本補填箇所」が1、本文が1~32、「奥書」「巻首見返後筆」がそれぞれ1と2とする。結果として、00001、00002、00101、00201~00232、00301、00302の計37種類を付与。

### 6.3 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。また、本コーパスでは本文種別および話者情報は設定しなかった。

### 6.4 作品情報

「中納言」による検索結果の各欄に表示される内容は以下の通りである。

- ・ジャンル  
法華百座聞書抄は「説教」、高山寺本古往来は「往来物」、尾張国解文は「文書」と表示する。
- ・作品名  
各資料名を表示する。
- ・成立年  
各資料の成立・加添年次(6.2節)を表示する。
- ・巻名等  
各資料の構成要素を以下の通り表示する。
  - ・法華百座聞書抄…日付・講師名・説教／説話(説話通し番号)
  - ・高山寺本古往来…書簡通し番号
  - ・尾張国解文…本文の場合は条数通し番号、それ以外は本文の種類(内題／副題／奥書／巻首見返後筆)
- ・作者  
本コーパスでは設定していない。

### 6.5 底本情報

#### 6.5.1 底本

「中納言」検索画面における底本の箇所には、コーパステキストおよび原文文字列を作成する際に参照した底本を収める書籍名・論文名を以下の通り表示した。

- ・法華百座聞書抄…「法華百座聞書抄総索引」
- ・高山寺本古往来…「高山寺本古往来表白集」
- ・尾張国解文…「真福寺宝生院蔵尾張国解文正中二年点訓読文」

#### 6.5.2 ページ番号

本コーパスに収録した3資料の原本はいずれも卷子本であるため、ページ番号には、各資料の行番号を表示した。法華百座聞書抄に限っては卷子本の両面に本文が記されているため、表面の7行目であれば「o007」、裏面の3行目であれば「u003」のように、表裏の別を表す「o/u」の後に行番号を表示した。

#### 6.5.3 参考リンク

本コーパスに収録した3資料に、原本・底本の画像・影印が公開されているものはない。ただし、3.4節で述べた通り、尾張国解文の底本とした宇都宮(2000)の本文は真福寺本の訓読文だが、真福寺本の冒頭の欠部に相当する早稲田大学本の本文を白文で示して

おり、この早稲田大学本の画像は閲覧が可能である。よって、尾張国解文のうち、早稲田本の補填箇所に限って画像リンクを表示した。

## 謝辞

『日本語歴史コーパス 平安時代編Ⅲ漢文系資料』は、国立国語研究所共同研究プロジェクト「開かれた共同構築環境による通時コーパスの拡張」、JSPS 科研費 25H01243「歴史知識辞書（ナレッジベース）の構築」（2025-2030）、日本学術振興会科学研究費助成事業 基盤研究（B）「和漢の両系統を統合する平安・鎌倉時代語コーパス構築のための語彙論的研究」（2012-2015 研究課題番号：24320086 研究代表者：田中牧郎）の成果の一部である。

## 参考文献

- 愛知県史編さん委員会編（2009）『愛知県史 資料編 7(古代 2)』愛知県
- 宇都宮睦男（2000）「真福寺宝生院蔵尾張国解文正中二年点訓読文」『日本文化論叢』8
- 梅村喬（2020）『尾張国郡司百姓等解文の時代』塙書房
- 小木曾智信・小町守・松本裕治（2013）「歴史的日本語資料を対象とした形態素解析」『自然言語処理』20(5)
- 高山寺典籍文書総合調査團編（1972）『高山寺本古往來表白集』東京大学出版会
- 国立国語研究所コーパス開発センター（池上尚）編（2016）『『日本語歴史コーパス 平安時代編』形態論情報規程集』
- 小林芳規編（1975）『法華百座聞書抄総索引』武蔵野書院
- 古文學祕籍複製會編（1934）『法華修法百座』古文學祕籍複製會
- 沼本克明（1972）「高山寺本古往來の音韻」高山寺典籍文書総合調査團編『高山寺本古往來表白集』東京大学出版会
- 肥爪周二（2023）「平安時代の仮名表記—書き分けない音韻を中心に—」『言語研究』164
- 森公章（2024）『平安時代の国衙機構と地方政治』吉川弘文館
- 柳原恵津子・近藤明日子（2022）『『日本語歴史コーパス 平安時代編Ⅱ訓点資料』（短単位データ 0.4）解説書』
- 山岸徳平（開題）（1976）『法華修法一百座聞書抄』勉誠社