

『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

柳原 恵津子（国立国語研究所）

近藤 明日子（国立国語研究所）

2022年3月31日

目次

1	はじめに	1
2	コーパスのテキスト	1
2.1	コーパステキスト(訓読文)	1
2.1.1	異訓の扱い	1
2.1.2	補読	2
2.1.3	原漢文の文意と加点の解釈が異なる場合	2
2.2	原文文字列	3
2.3	コーパステキストと原文文字列の表示法	3
2.3.1	漢文本文	3
2.3.2	ヲコト点	3
2.3.3	仮名点	4
2.3.4	濁点	4
2.3.5	補読	4
2.3.6	踊り字	5
2.3.7	句読点	5
2.3.8	余分な訓点	5
2.3.9	加点誤り	5
2.3.10	振り仮名	5
2.4	使用する文字集合	6
3	コーパステキストの語形(読み)の認定基準	6
3.1	基本的な語形認定基準	7
3.2	特別な語形認定基準	7
3.2.1	和訓と字音の読み分け	7
3.2.2	音便形	8
3.2.3	漢語サ変動詞	8
3.2.4	義注による読みの特定	8
3.2.5	敬語法	8
3.2.6	にして／にて	9
3.2.7	「唯」の読み	9
3.3	認定した語形一覧	9
3.3.1	基本語形を統一したもの	9
3.3.2	語形を複数認めるもの	11
4	形態論情報	11
4.1	形態論情報の概要	11
4.2	形態論情報の多重化	11
4.3	形態論情報の検索の注意点	12
5	コーパス検索ツール「中納言」	13
	謝辞	17
	参考文献	17
	参考 URL	18

1 はじめに

『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）は、西大寺本『金光明最勝王経』平安初期点の巻1の訓読文をコーパス化したものである。形態論情報（4節参照）の付与された本格的な訓点資料のコーパスとしてはじめて構築、公開するものである。

訓点資料とは、漢籍・仏典等の漢文で書かれた文章に、仮名やヲコト点・返り点等の符号を付して、日本語の語法にそった読み下し方を示しているもので、学問や宗教、政治の場で長く行われた漢文訓読の際の日本語の姿をとどめた資料である。そこには仮名文学作品等とは異なる語彙・語法が用いられており、当時の日本語の多様性を知る上で欠かせないものである。本コーパスの底本とした西大寺本『金光明最勝王経』平安初期点は、天平宝字6（762）年、百濟豊虫によって筆写された『金光明最勝王経』全10巻に、830年頃の白書による訓点記されたものである。質の高い訓点記が密に施されていることから平安初期を代表する訓点資料とされており、これまでに春日（1942）や総本山西大寺（編）（2013）といった研究書・影印が刊行され、日本語史上の重要な資料として扱われてきた。本格的な訓点資料のコーパスの嚆矢とするにふさわしいものであり、本コーパスの収録対象として選定した。

ただ、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ているという訓点資料の特質上、『日本語歴史コーパス』収録の他のサブコーパスとは異なる独自の処理を本文に行った点が少なくない。そのため、はじめに本解説書に目を通し、テキスト作成や形態論情報付与の際に採用した方法を把握した上で利用してほしい。

2 コーパスのテキスト

2.1 コーパステキスト（訓読文）

本コーパスでは、底本の漢文本文とそれに加点された訓点に基づいて読みを定め、漢字平仮名交じり文（訓読文の表示法の詳細については2.3節参照）で訓読文を作成し、それをコーパステキストとした。これまで広く用いられてきた春日（1942）の訓読文をそのまま用いず、独自の訓読文を作成したことになる。春日の訓読文は、漢文中に頻用されるにもかかわらずほとんど加点されない助辞類（例：「復」「若」「及」）や定型表現（例：「何以故」「善哉善哉」）の読みを特定せず、白文のままとすることがしばしばあり、形態論情報を付与するために読みを決定していく過程で、春日の訓読文とは少なからず異なりが生じるからである。さらに近年、底本の影印（総本山西大寺編2013）が刊行され、底本の加点状況の詳細が把握できる環境が整っている状況でもあるため、春日（1942）およびそれ以降の訓点語研究や日本語史研究の成果を踏まえつつ、新しい訓読文を作成することにしたものである。訓読文作成にあたっての語形（読み）の認定基準については3節を参照されたい。

2.1.1 異訓の扱い

底本では、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、複数の訓読文を作成することが可能である。そこで、訓点の記載の方法・位置によって次の(1)～(4)のように読みの優先順位を定め、優先順位がもっとも高い読みのみを採用して訓読文を作成し、コーパステキストとした。

- (1) ヲコト点およびヲコト点と組み合わせてひとつの読みとなる右傍仮名点
- (2) 仮名点 右傍（うち、ヲコト点と組み合わせて用いられていないもの。複数の読みが漢字右傍に並列して書かれている場合は、原則としてより漢字に近いもの、上下に書かれている場合は、より上に書かれたものを優先とする）
- (3) 仮名点 左傍（複数書かれている場合の優先順位は右傍に同じ）

(4) 別筆

よって、最も優先順位の高い訓以外の異訓は本コーパスでは参照できないことに注意が必要である¹。

2.1.2 補読

底本は訓点資料の中でもとりわけ加点が密に行われている資料であるが、それでも形態論情報付与のための形態素解析に適したテキスト、あるいは語形や文の区切りの理解を容易にするテキストを作成するにあたって、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。そこで適宜必要最小限の補読を行って訓読文を作成した。以下にその補読の例を示す²。

【例】

〈底本〉³ 我レ聞たまへキ。

〈コーパステキスト〉我れ聞きたまへき。(20K 西金 0830_01001, 510)⁴

〈底本〉 復梨車毗の童子有り。

〈コーパステキスト〉復た梨車毘の童子有り。(20K 西金 0830_01001, 16310)

〈原文〉 或は

〈コーパステキスト〉或いは (20K 西金 0830_01001, 30740)

〈底本〉 如キことを是

〈コーパステキスト〉是の如きことを (20K 西金 0830_01001, 420)

〈原文〉 如來は見ては彼の有情の正行を修習するを、

〈コーパステキスト〉如來は彼の有情の正行を修習するを見そなはしては、
(20K 西金 0830_01002, 96820)

〈原文〉 以て大智慧を能ク爲照明したまふこと

〈コーパステキスト〉大智慧を以て、能く照明したまふこと、(20K 西金 0830_01002, 39990: 40100)

2.1.3 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読し、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

1 本コーパスに先んじて国立国語研究所ウェブページ「日本語史研究用テキストデータ集」において公開した、同じ底本によるコーパスデータとして、『ひまわり版「西大寺本金光明最勝王経平安初期点訓読文コーパス」Ver.0.1』、「巻一 訓読文 (xml)」、「巻一 訓読文 (txt)」がある

(https://www2.ninjal.ac.jp/textdb_dataset/sksh/)。これらのデータには形態論情報は付与していない一方で、別訓・別筆、仮名点の記載位置、ヲコト点の形式など加点に関する詳細な情報を付与している。異訓を参照したい場合はこれらのデータを利用されたい。

2 振り仮名の表示方法は特殊なため、ここには例示しない。詳細は2.3.10節を参照のこと。

3 底本の用例は、漢文の漢字を出現順に示し、各漢字に加点された訓点をその漢字の直後に小書きで示す。訓点はヲコト点を平仮名、仮名点を片仮名、句点相当の符号を「。」、読点相当の符号を「、」で示す。

4 コーパステキストの用例では、末尾の () 内にサンプル ID (表 2「サンプル ID」項参照)・開始位置 (表 2「開始位置」項参照) を記し、下線部の文字の出現位置を示す。

キストの中に語法として不自然なものが見受けられた場合には、加号誤りの可能性も想定して原漢文を確認することが望ましい。

【例】

〈底本〉

図1 参照

〈コーパステキスト〉 如来の壽命長遠なりと説きたまふを聞きて

(20K 西金 0830_01002, 59250)

※「如来の壽命長遠なり」は、直前の文脈に記された憍陳如（婆羅門）の発話。婆羅門に「たまふ」の待遇は不自然である。原漢文の語順からは、憍陳如が「如来の寿命は長遠である」と語ったと解釈できるが、「説」に「たまふ」のヲコト点があるため（漢字左下「一」）、加号者は、その場にいる「四はしらの如来」、または釈迦自身の発話と判断していたことが推測される。底本に加号があることを尊重し、敬語形のまま訓読文を作成した。



図1

2.2 原文文字列

2.1 節に述べたように、コーパステキストとした訓読文では、テキストのどの部分が底本の訓点と対応し、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号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」(5 節参照)の検索結果に表示される「原文文字列」において、底本での加号状況を識別できるテキストを表示することとした。原文文字列には、従来の訓点資料研究で作成されてきた訓読文で広く用いられてきた表示法(2.3 節参照)を使用し、これまでの研究との連続性を保てるようにした。また、「中納言」では原文文字列の他に、「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。

2.3 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。

2.3.1 漢文本文

底本の漢文本文で使用されている漢字は、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2.4 節を参照のこと。

ただし、訓読する際に不読字とするべき漢字は、コーパステキストには表示しなかった。原文文字列では、不読字を該当する位置に表示し、全角角括弧 [] で括って示した。

【例】

〈底本〉

於其の四面に各有り上妙の師子の之座有り。

〈コーパステキスト〉 其の四面に各上妙の師子の座有り。

(20K 西金 0830_01002, 4820: 4940)

〈原文文字列〉

[於] 其の四面に各上妙の師子の [之] 座有り。

2.3.2 ヲコト点

ヲコト点は、コーパステキスト・原文文字列ともに平仮名で示した。

【例】

〈底本〉

満月の

〈コーパステキスト〉 満月の (20K 西金 0830_01002, 39860)

〈原文文字列〉

満月の

2.3.3 仮名点

仮名点は、コーパステキストでは平仮名で、原文文字列では片仮名で示した。

【例】

〈底本〉 如^ク
〈コーパステキスト〉 如_ク (20K 西金 0830_01002, 39870)
〈原文文字列〉 如^ク

底本の仮名点には、上代特殊仮名遣いのうち、「コ」のみ甲類（「古」）と乙類（「己」）の区別が残存するが、コーパステキストではすべて「こ」で示した。原文文字列では、甲類の「コ」は「コ」、乙類の「コ」は全角山括弧で括り「〈コ〉」で示した。

【例】

〈底本〉⁵ 彼^己に
〈コーパステキスト〉 彼_己に (20K 西金 0830_01002, 89670)
〈原文文字列〉 彼^コに

また、底本の巻2以降で、仮名点にア行（「衣」）とヤ行（「江」）の区別が見られる。コーパステキストでは区別せずすべて「え」で示すこととした。原文文字列では、ア行の「エ」は「エ」、ヤ行の「エ」は全角山括弧で括り「〈エ〉」と示すこととした。ただし、底本巻1にはヤ行の「エ」の例はなく、本コーパス（短単位データ0.4）において原文文字列に「〈エ〉」は存在しない。

さらに、底本では、仮名点に交じって「事（コト）」「如（ゴトシ）」「人（ヒト）」「物（モノ）」「命（ノタマフ）」「申（マラス）」などの頻用語を和訓とする漢字が、仮名に準じて用いられる。このような仮名点の一部として記された漢字は、コーパステキストでは「こと」「ごとし」「ひと」「もの」「のたまふ」「まをす」のように平仮名に開いて示し、原文文字列では「コト」「ゴトシ」「ヒト」「モノ」「ノタマフ」「マラス」のように片仮名に開いて示した。

2.3.4 濁点

底本ではすべて清音の訓点を使用されるが、コーパステキストでは濁音が期待される部分は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に倣い清音の仮名で表記した。

【例】

〈底本〉 有ル^ハシ
〈コーパステキスト〉 有る_ハし (20K 西金 0830_01002, 11960)
〈原文文字列〉 有ル^ハシ

〈底本〉 其の名をは
〈コーパステキスト〉 其の名を_ば (20K 西金 0830_01001, 8930)
〈原文文字列〉 其の名を^ハ

2.3.5 補読

2.1.2 節に述べたように、コーパステキストでは底本の加点とは対応しない送り仮名・付属語・句読点・振

5 底本の用例では、乙類「コ」の仮名点を「己」で示す。

り仮名を適宜補い、平仮名で示した。原文文字列ではこれを全角丸括弧（ ）で括り示した。

【例】

〈底本〉 我レ聞たまへキ。
〈コーパステキスト〉 我れ聞きたまへき。(20K 西金 0830_01001, 510)
〈原文文字列〉 我レ聞(き)たまへキ。

2.3.6 踊り字

底本の仮名点で使用される踊り字は、コーパステキストではその踊り字が示す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】

〈底本〉 佛に護ラルゝ童子
〈コーパステキスト〉 佛に護らるる童子 (20K 西金 0830_01001, 17150)
〈原文文字列〉 佛に護ラルゝ童子

2.3.7 句読点

底本の漢文本文の漢字左下のやや漢字から離れたところに記された句切点（星点）を句点または読点とし、訓読文作成者の判断で句点・読点の区別を行った。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。

2.3.8 余分な訓点

底本にどうしても訓読文に反映できない余分な訓点がある場合、コーパステキストではその訓点はテキストとして示さなかった。また、原文文字列でも、該当訓点を示すことで読みにくさを招く恐れがあるため、その訓点はテキストとして示さなかった。

【例】

〈底本〉 煩に惱を
〈コーパステキスト〉 煩惱を (20K 西金 0830_01001, 31280)
〈原文文字列〉 煩惱を

2.3.9 加点誤り

底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テキストを示した。原文文字列では、底本の訓点のままのテキストを示した。

【例】

〈底本〉 知り數へルこと
〈コーパステキスト〉 知り數ふること (20K 西金 0830_01002, 16890)
〈原文文字列〉 知り數へルこと

2.3.10 振り仮名

底本の訓点のうち、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」（5 節参照）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。その表示法は、2.3.2 節～2.3.9 節で示した原文文字列の表示法と同様とした。コーパステキストの表示法とは異なる点に注意されたい。

【例】

〈底本〉 旦いふ
〈コーパステキスト〉 旦ふ (20K 西金 0830_01002, 420)
〈振り仮名〉 い
〈原文 KWIC のキー〉 旦いふ

〈底本〉 策ハケみ
〈コーパステキスト〉 策み (20K 西金 0830_01002, 26200)
〈振り仮名〉 ハケ
〈原文 KWIC のキー〉 ハケ 策み

※「策み」の語形は「ハゲミ」であるが、振り仮名の表示法は「ハケ」と清音になる。

〈底本〉 惡クみ
〈コーパステキスト〉 惡み (20K 西金 0830_01002, 8620)
〈振り仮名〉 (に)ク
〈原文 KWIC のキー〉 (に)ク 惡み

2.4 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 0213:2004 に準拠した⁶。

底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。ただし、底本の漢文本文は天平宝字 6(762)年に書写されたものであり、当時広く行われた初唐の標準的な字体が用いられている。また、一回的に用いられた誤字と判断できる字体なども使用され、JIS の包摂規準を適用できない漢字字体が多く使用される。それらの字は、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（例：底本「明」→コーパステキスト「明」）。

以上の処理を行ってもなお電子化できない漢字は、「■」（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。ただし、本コーパス（短単位データ 0.4）では■入力の例はなかった。

以上のように、底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、字体の確認が必要な場合は影印（総本山西大寺編 2013）等を参照されたい。

3 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ためには、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。底本にはヲコト点・仮名点等による詳細な加点があるが、それらは語尾部分や付属語、一部の自立語に偏っているため、これら以外の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。

6 ただし、①JIS X 0213 附属書 7 2.1b) に掲載される、戸籍法施行規則付則別表「人名用漢字許容字体表」（昭和 56 年法務省令 51）の漢字、及び常用漢字表（昭和 56 年内閣告示第 1 号）のかっこ書き内の漢字（「いわゆる康熙字典体」）のうち、JIS X 0208 で包摂していた漢字、②JIS X 0213:2004 において UCS との互換のために追加された 10 字、についてはこれを用いない。

3.1 基本的な語形認定基準

(1) 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 西大寺本『金光明最勝王経』平安初期点（全 10 巻）に和訓による加点例があればそれによる。また、字音の語形を示す仮名点・漢字による音注や合符が付された例があれば、字音語とする。
- ② 築島裕（2007-2009）『訓点語彙集成』に収録された他の訓点資料での加点例のうち、時代の近い（平安初・中期）訓点資料での和訓の加点例を参照する。
- 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などにより、和訓または字音読み of 語の確例を参照する。

(2) (1) で読みが確定しなかった場合は、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』記載の呉音で読む。

(3) サンスクリット語の固有名詞や植物名も音訳に用いられている漢字の呉音読みを原則とする。個別に平安初期の発音や標準的な表記を判断するのが困難であり、呉音読みがサンスクリット語の原音に近い事例がしばしば見られるためである。

3.2 特別な語形認定基準

3.2.1 和訓と字音の読み分け

3.1 節にあげた (1) の手順で和訓が定まる場合でも、平安初期仏典で字音で読む方が妥当である場合は字音で読む。字音語で読むことが妥当と判断した語については、文脈によって字音で読んだり和訓で読んだりといった読み分けは、原則しない。

【例】力 語形：リキ （×チカラ）

【例】身 語形：シン （×ミ）

ただし、底本 10 巻分の加点例などを鑑み、文脈や当該例での加点の状況から読み分けを認める場合には、以下のA)～C)の要領で字音で読む場合と和訓で読む場合とを個別に判断する。

A) 文中での役割が異なる場合

同じ漢字（あるいは漢字文字列）で記されているものの、品詞や文中での機能が異なる場合は、読み分けを行う。

【例】衆

人を指す場合には「シュ」を語形とする。

〈コーパステキスト〉大苾芻の衆九万八千人（20K 西金 0830_01001, 1110）

連体修飾的な場合には「オオク」を語形とする。

〈コーパステキスト〉帝の青瑠璃と種種の衆の寶とをもちて、（20K 西金 0830_01002, 4220）

【例】垢

原則として平安初期仏典での読みとして妥当な「ク」を語形とする。

〈コーパステキスト〉心を淨くして垢を無から令め、（20K 西金 0830_01001, 37020）

ただし、複合語の一部と解せる場合は「アカ」を語形とする。

〈コーパステキスト〉垢つき穢れたること有る者は（20K 西金 0830_01002, 8800）

B) 原本に複数の読みの加点例がある場合

【例】生す

「ショウス（サ変）」「オコス（四段）」どちらの活用語尾の加点例も見られるが、加点からはどちら

の活用型の例か判断できない例も多い。個々の例の読みを特定するにあたっては、加点があれば加点に沿い、区別のつかない連用形・終止形については、自動詞的な用法の場合は「ショウス」、他動詞的な場合は「オコス」と読むことにする。

C) 数詞

後ろに続く助数詞が字音語の時には字音語として、それ以外は和語として読む。

3.2.2 音便形

音便形・非音便形の両形の読みの可能性のある箇所については、春日(1942、研究篇、p.50)も例外的に音便形を示す加点例が別筆にあると指摘する「おいて」「まうづ」以外、非音便形で読む。ただし「おいて」は助詞的に用いる連語の場合のみで、本動詞「置く」の例は非音便形で加点された例が同じく別筆にあるため、これに従って非音便形とする。

3.2.3 漢語サ変動詞

字音語を文語サ行変格活用の動詞として読む場合、読み添えるサ変動詞をすべて「～す」と読み、「～ず」は採用しない。

3.2.4 義注による読みの特定

底本には、ヲコト点・仮名点ではないものの、義注の形式で本文の漢字の意味を記している箇所がある。第一義的にはあくまでも意味を記すための形式ではあるが、春日(1942、研究篇、p.28)も述べるように、この義注が読み(和訓)も示しているととらえるのが自然な例が散見される。よって、義注がある場合には、義注に示された漢字の定訓で読む。

【例】

〈底本〉 図2 参照

〈コーパステキスト〉 廣く舎宅に造らむときに (20K 西金 0830_01002, 90510)

※「造」(実際は加点位置のずれにより「於」)に「至也」の義注がある。この義注は、『大広益会玉篇』(古代字書輯刊、2004、中華書局)に「造、徂皓切爲也又七到切至也」とあること、元暁(新羅)『金光明經疏』に「造者趣也」、恵沼(唐)『最勝王經疏』に「造諸(詣)」、憬興(新羅)『最勝王經略贊』に「造者進也」とあったと願暁(生年不詳～874)『金光明最勝王經玄樞』から知られることなどを踏まえると、「造」が「至」の意であるとの注記だと判断できる。そこで、この義注を読みにも採用し、この「造」(およびヲコト点、仮名点)を「いたらむときに」と読む。

廣 ク
造 ラ
於 至也
舎
宅
図2

3.2.5 敬語法

春日(1942、研究篇、p.262)は、敬語が釈迦、国王、王族、また仏法を記した経典そのものに用いられること、これらの人物などに対しては謙讓表現も用いられ、釈迦は自敬表現も用いることを指摘する。実際に底本で確認すると、春日の指摘は概ね適切で、上記のほか釈迦以外の如来にも敬語が用いられることがわかる。

また、これもすでに春日(1942、研究篇、p.253)が指摘するように、敬語の読み添えのあり方は用言と体言で異なる。そこで、用言体言それぞれの加点状況を踏まえ、以下のA)B)のように語形を認定する。

A) 用言

動作主が釈迦・四はしらの如来・経典の場合、用言については敬語形を語形として採用する。語彙的な敬語の場合は適切な敬語動詞を語形とし、それ以外の場合には補助動詞「たまふ」などを補読する。

【例】

〈底文〉 如來は見ては彼の有情の正行を修習するを、
〈コーパステキスト〉 如來は彼の有情の正行を修習するを見そなはしては、(20K 西金 0830_01002, 96810)

B) 体言

同時代の訓点資料での加点の状況を踏まえると、敬語相当の訓で読むべき名詞がある(例:「足」が釈迦の足の場合は「ミアシ」)、「容」が釈迦の顔の場合は「ミカオ」)。しかし、底本全巻にわたり、体言を敬語相当の語形で読ませるような加点の例は見られない。自明であるから加点されないとの判断も十分できるものの、動詞には加点例が見られることを踏まえ、体言については一括して、敬語相当の語を採用しないことにする。よって、そのための補読も行わない。

ただし、助数詞「柱(ハシラ)」のみ加点例が見られるため、これを認め、必要な補読を行う。

【例】

〈底文〉 四の如來
〈コーパステキスト〉 四はしらの如來(20K 西金 0830_01002, 250)

3.2.6 にして/にて

ヲコト点「に」とヲコト点「て」が一つの漢字に加点されている場合、「にて」と読むか「に(し)て」と読むかという問題が発生する。実際の加点を見ると、類似した文脈で「にして」(ヲコト点「に」・「して」)・「(に)して」(補読「に」+ヲコト点「して」)のように加点した例が散見されるため、本コーパスでは「にて」を認めることはせず、「にて」と解せる例には「し」を補読し「にして」とする⁷。

【例】

〈底本〉 一にて心を瞻仰たてまつる諸佛の殊勝の之相を。
〈コーパステキスト〉 心を一にして諸佛の殊勝の相を瞻仰したてまつる。(20K 西金 0830_01002, 32880)

3.2.7 「唯」の読み

底本において「唯」はとりわけ加点されることが稀な漢字で、巻1内の10例のうち1例に仮名点「タ」の加点が見られるのみである。ただ、副助詞「し」が加点された例が3例見られ、このような例については「ただし」と読むのが妥当である。「ただに」「ただの」と読む例があれば、それらの品詞は従来の規定に沿って形状詞と判断することになるが、「に」「の」の加点例は見られない。このような実態を踏まえ、「唯」をすべて副詞「ただ」と判断し、副助詞「し」の加点の有無によって「ただ」または「ただし」と読む。

3.3 認定した語形一覧

以下に、3.1節、3.2節に示した語形認定基準により認定した語形の一覧を挙げる。

3.3.1 基本語形を統一したもの

(1) 漢字1字名詞(語形が揺れる例のみ挙げる)

i. 字音語とした漢字1字名詞

有(ウ)、喜(キ)、言(ゴン)、食(ジキ)、生(ショウ)、性(ショウ)、侵(シン)、身(シン)、説(セツ)、仙(セン)、染(ゼン)、相(ソウ)、想(ソウ)、尊(ソン)、忍(ニン)、滅(メツ)、利(リ)、力(リキ)

⁷ ただ、ヲコト点「に」とともにヲコト点「て」・「して」どちらを加点するかの異なりが「にて」「にして」の異なりである可能性もある。当該部分を用例として用いる場合は、この点留意が必要である。

ii. 和語とした漢字1字名詞

男 (オトコ)、女 (オミナ)、容 (カオ)、雑 (クサグサ)、罪 (ツミ)、称 (ナ)、報 (ムクイ)

(2) 語形が音変化で揺れる語、語形の候補が複数ある語

跌 (アナウラ)、敬 (イヤマウ)、稟 (ウケル)、大 (オオキナリ)、逮 (オヨブ)、
被・蒙 (カガフル)、行 (ギョウス)、現 (ゲンス)、退 (シゾク)、掌 (タナウラ)、促 (ツヅマル)、
同 (ドウナリ)、汝等 (ナムダチ)、樂 (ネガウ)、將 (ヒキイル)、名称 (ホマレ)、申 (モウス)、
最 (モトモ)

(3) 漢文訓読特有の語法

※左列に底本例、右列に語形をあげる。(×～～)は採用しなかった語形の候補を示す。

有所ル／所有ル	…アラルル (×アラユル)
或ときにはA、或B	…アルトキニハA、アルトキニハB
或は	…アルイハ (×アルハ)
云何をか二と爲すとならば	…イカナルヲカニトスルトナラバ
雖	…仮定条件：トイウトモ 確定条件：トイエドモ
乃	…イマシ (×スナワチ)
如是等	…コレラノゴトキ (×カクノゴトキラノ)
悉	…コトゴトク
便／即／則	…スナワチ
唯／唯し	…タダ／タダシ (×タダニ)
俱／共／俱共／與	…トモニ
A及B	…AトBト
Aすること勿	…Aスルコトナ ※「ナ」は副詞と判断する。
何以故	…ナニヲモチテノユエニ
A乃至B	…Bの末尾に「ニ」、「至」に「マテニ」などの加点がある場合 ：A、Bニイタルマデニヲモ 加点がない場合、あるいはA・B末尾に「ト」の加点がある場合 ：AトBトヲ
Aにて	…Aニシテ (×Aニテ)
の於に	…ノウエニ (×ノタメニ)
一者	…数詞が一～九の場合：ヒトツハ・フタツハ… (×ヒトツニハ) 数詞が十の場合：トオハ ※「者」にヲコト点「は」を加点する例が見られるため、「者」を「は」と読む。
方	…マサニ
亦タ／復タ／亦復タ	…マタ
若自	…ミズカラ (×オノズカラ)
若／若は	…モシ／モシハ (×モシクハ)
Aヤ不ヤ	…Aヤイナヤ
所以者何	…ユエンハイカニ

容し

…ヨシ (副詞)

3.3.2 語形を複数認めるもの

垢 通常：ク

「垢ツク」の加点がある場合：アカ (ツク)

生 (動詞用法で加点がない、または「ショウス」「オコス」の区別がつかない連用形・終止形の場合)

自動詞的：ショウス

他動詞的：オコス

※加点から区別がつく場合は、加点に従って読む。

何 副詞：イカニ

代名詞：ナニ

空 仏教の教義としての用法：クウ

実在の天体を指す用法：ソラ

前 時間を表す用法：サキ

空間を表す用法：マエ

4 形態論情報

4.1 形態論情報の概要

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) 『『日本語歴史コーパス 平安時代編』形態論情報規程集』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。

4.2 形態論情報の多重化

本コーパスではコーパステキストとなる訓読文を作成するにあたり、形態論情報の付与を見込んで、短単位の境界や並び順と文字列の境界や並び順が対応関係を持つようにした。ただし、底本の漢字本文とそれに加点された訓点との関係から、短単位と文字列との間に対応関係のとれないコーパステキストをやむを得ず作成した箇所がある。その箇所では、コーパステキストに対して直接形態論情報が付与できないため、形態論情報の多重化の技術 (小木曾 2017) を用い、コーパステキストの読みを平仮名表記したテキストを主本文とし、それに対して形態論情報を付与した。その場合、本来のコーパステキストの文字列は副本文とし、品詞「対象語無し」として未知語扱いとした (表 1)。

表 1 形態論情報の多重化の例

コーパス テキスト	テキスト の主副	主な形態論情報					
		書字形 出現形	語彙素	語彙素 読み	品詞	活用型	活用形
① 所有る	主	あら	有る	アル	動詞-非自立可能	文語ラ行変格	未然形一般
		るる	れる	レル	助動詞	文語下二段-ラ行	連体形一般
	副	所有る			対象語無し		
② 鋒	主	ほこ	矛	ホコ	名詞-普通名詞一般		
		の	の	ノ	助詞-格助詞		
		さき	先	サキ	名詞-普通名詞-副詞可能		
	副	鋒			対象語無し		

表 1 の①「所有る」(20K 西金 0830_01001, 25050) の例は、底本の漢字本文「所有」の「有」字にヲコト点「る」が加添されており、その状態を表すためコーパステキストは「所有る」としたものである。このコーパステキストと短単位(動詞「有る」+助動詞「れる」)の境界・順序が一致しないため、平仮名表記のテキスト「あらるる」を主本文とし、それに対して形態論情報を付与した。表 2 の②「鋒」(20K 西金 0830_01002, 88990) の例は、コーパステキストの漢字「鋒」1 字に対して複数の短単位の形態論情報(名詞「矛」+助詞「の」+名詞「先」)を付与する必要があるものである。コーパステキストと短単位の境界が一致しないため、平仮名表記のテキスト「ほこのさき」を主本文とし、それに対して形態論情報を付与した。

4.3 形態論情報の検索の注意点

このように、本コーパスではコーパステキストとは異なるテキストに対して形態論情報の付与を行っている箇所がある。そのため、コーパス検索ツール「中納言」(5 節参照)で検索を行う場合、以下の点で注意が必要である。

(1) 「中納言」の短単位検索における注意点

「中納言」の検索方法の一つである短単位検索において、書字形出現形を検索する場合、形態論情報の多重化を行っている箇所では主本文である平仮名表記したテキストが対象となる。副本文となるコーパステキストを検索対象としたい場合は、「検索動作」欄の「副本文」ボックスで「副本文を検索対象に含む」を選択する必要がある。ただし、「副本文を検索対象に含む」を選択した場合は、前方・後方共起の条件指定をした検索はできないことに注意が必要である。

(2) 「中納言」の文字列検索における注意点

「中納言」の検索方法の一つである文字列検索において、検索対象となるのはコーパステキストの文字列である。よって、形態論情報の多重化を行っている箇所では副本文が検索対象となり、主本文である平仮名表記したテキストは対象とならない。例えば、「ほこ」を文字列検索しても表 2 の②の例はマッチしない。「ほこ」はコーパステキストではないためである。

また、文字列検索の検索結果では、検索文字列を含む本行の短単位の範囲に対応する主本文の最後の 1 短単位の形態論情報のみが表示される。例えば、表 2 の②の例の場合、「鋒」字で文字列検索を行うと、返ってくる検索結果で表示されるのは主本文の語彙素「先」の形態論情報のみであり、主本文に紐付けられた他の語彙素「矛」「の」の形態論情報や、副本文の品詞「対象語無し」の形態論情報は表示されない(図 3)。

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
20K西金0830_01002	50740	35280		1、 口 中 白 齒 生 し て、 長 大 き し て 利 か ら む こ と	鋒	の 大 く あ ら む と き に、 方 に 佛 の 舍 利 を は 求 む べ し 。# 假使 ひ 冤 の 角 を	サキ	先	サキ	名詞-普通名詞-副詞可能			鋒	(ほ)(こ)(の)(さ)(ぎ)
				口の中に白き歯生(し)て、長	(ほ)(こ)(の)(さ)(ぎ)鋒	(の)如(あ)らむ(とき)に(、)方(に)佛(の)舍利(を)は(求)む(べ)し(。) # 假使(ひ)冤(の)角(を)持(ち)て(、)用(あ)て[於]梯(と)成(して)								

図3 「中納言」文字列検索で「鋒」を検索した検索結果表示

ただし、検索結果の「サンプル ID」列をクリックすると別ウィンドウで表示される「詳細な文脈情報」で、主本文に紐付けられた他の形態論情報を確認することができる(図4)。

詳細な文脈情報													
平安-訓点資料	20K西金0830_01002	35250	1	利から	トイ	疾い		形容詞-一般	文語形容詞-ク	未然形-補助トカラ	和	利(か)ら	
平安-訓点資料	20K西金0830_01002	35260	1	む	ム	む		助動詞	文語助動詞-ム	連体形-一般	和	む	
平安-訓点資料	20K西金0830_01002	35270	1	こと	コト	事		名詞-普通名詞-一般		コト	和	こと	
平安-訓点資料	20K西金0830_01002	35280	1	鋒	ホコ	牙		名詞-普通名詞-一般		ホコ	和	鋒	(ほ)(こ)(の)(さ)(ぎ)
平安-訓点資料	20K西金0830_01002	35280	1	鋒	ノ	の		助詞-格助詞		ノ	和	鋒	(ほ)(こ)(の)(さ)(ぎ)
平安-訓点資料	20K西金0830_01002	35290	1	鋒	サキ	先		名詞-普通名詞-副詞可能		サキ	和	鋒	(ほ)(こ)(の)(さ)(ぎ)
平安-訓点資料	20K西金0830_01002	35290	1	の	ノ	の		助詞-格助詞		ノ	和	(の)	
平安-訓点資料	20K西金0830_01002	35300	1	如く	ゴトシ	ごとし		助動詞	文語助動詞-ゴトシ	連用形-一般	和	如く	
平安-訓点資料	20K西金0830_01002	35310	1	あら	アル	有る		動詞-非自立可能	文語ラ行変格	未然形-一般	和	(あ)ら	
平安-訓点資料	20K西金0830_01002	35320	1	む	ム	む		助動詞	文語助動詞-ム	連体形-一般	和	む	

図4 「中納言」の「詳細な文脈情報」ウィンドウ

5 コーパス検索ツール「中納言」

本コーパスは、コーパステキストおよびアナノテーションのデータを、コーパス検索アプリケーション「中納言」(https://chunagon.ninjal.ac.jp/)での検索結果の形で利用者に提供する(図5)。

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	著者	ジャンル	作品名	成立年	巻名等	作者	生年	蔵本	ページ番号
20K西金0830_01001	3880	2640		1 等 の 如 時 諸 の 大 聲 聞 あり 。# 各 時 時 に、 定 從 し て 起 ち て、 は に 往 詣 し、 問 を も ち て、 右 に 坐 す 。	佛	の 所 に 往 詣 し、 問 を も ち て、 佛 の 足 を 礼 し た て ま つ り て、 右 に 坐 す 。	ホトケ	仏	ホトケ	名詞-普通名詞-一般		佛	訓点資料	西大寺本金 光明殿勝王 経平安初期 点				830	巻第一 序品第一	* (加 点)既浄 (作)	* /635	西大寺本金 光明殿勝王 経 巻第一	1	
20K西金0830_01001	4010	2750		1 には、 定 從 し て 起 ち て、 佛 の 所 に 往 詣 し、 問 を も ち て、 是 等 の 如 き 諸 の 大 聲 聞 あり 。# 各 時 時 に、 定 從 し て 起 ち て、 佛 の 所 に 往 詣 し、 問 を も ち て、 右 に 坐 す 。	佛	の 足 を 礼 し た て ま つ り て、 右 に 三 市 締 り て、 從 き て 一 面 に 坐 す 。	ホトケ	仏	ホトケ	名詞-普通名詞-一般		佛	訓点資料	西大寺本金 光明殿勝王 経平安初期 点				830	巻第一 序品第一	* (加 点)既浄 (作)	* /635	西大寺本金 光明殿勝王 経 巻第一	1	
20K西金0830_01001	7880	5500		1 た て ま つ り て、 脫 履 せ ず、 私 誓 の 心 を 發 して、 未 來 際 を 盡 す 。# 廣 く 諸 佛 に 應 事 へ た て て、 (ま)つ り て、 脫 履 せ ず、 私 誓 の 心 を 發 して、 未 來 際 を 盡 す 。# 廣 く	佛	の 所 に、 深 く 淨 因 を 種 々 た り 。# 三 世 の 法 に 於 て、 兼 生 を 悟 ル 忍 あり 。	ホトケ	仏	ホトケ	名詞-普通名詞-一般		[於]佛	佛	訓点資料	西大寺本金 光明殿勝王 経平安初期 点				830	巻第一 序品第一	* (加 点)既浄 (作)	* /635	西大寺本金 光明殿勝王 経 巻第一	1

図5 「中納言」の検索結果の表示例

「中納言」では「検索対象」欄で検索するコーパスを選択することができる。『日本語歴史コーパス』のうち、本コーパスのみ検索対象としたい場合は、「検索対象」欄の「検索対象を選択」ボタン(図6)をクリック

クすると表示される「検索対象の選択」ウィンドウで「平安-訓点資料」の「コア⁸」にチェックを入れる（図7）。

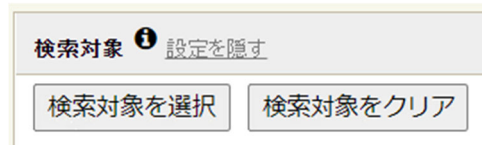


図6 「中納言」の「検索対象」欄



図7 「中納言」の「検索対象の選択」ウィンドウ

次の表2では、「中納言」の検索結果で表示されるテキスト・アノテーションのうち、初期設定で表示される項目と、初期設定では表示されないが本コーパスで特に注意が必要な項目である「主本文」「多重化種別」「性別」（表中*を付す）について説明する。

表2 「中納言」検索結果の主な表示項目

情報種別	項目名	内容
コーパス情報	サンプルID	<p>検索対象の含まれるサンプルのID。 1～2桁目 …すべて「20」。時代名「平安時代」を表す。 3桁目 …すべて「K」。作品ジャンル「訓点資料」を表す。 4～5桁目 …底本の略称を表す。 6～9桁目 …底本の成立年（下記「成立年」項参照）を表す。 10桁目 …すべて「」。区切り符号。 11～12桁目 …巻の通し番号を表す。 13～15桁目 …品の通し番号を表す。 サンプルIDをクリックすると「詳細な文脈情報」ウィンドウが開き、検索対象の前後文脈の形態論情報を見ることができる（図4参照）。</p>

⁸「コア」については表2「コア」項を参照。本コーパスはすべて「コア」であり、「コア」を選択すれば本コーパス全体を選択したことになる（「非コア」は選択できない状態になっている）。

情報種別	項目名	内容
	開始位置	「キー」の本行テキストの先頭の文字の、サンプル内における位置を表すID。10 きざみの連番。
	連番	「キー」に対応する主本文の短単位 ⁹ の、サンプル内における位置を表すID。10 きざみの連番。
	コア	検索対象の含まれるサンプルがコア ¹⁰ データであることを表す。「1」がコア、「0」が非コアを表す。本コーパスではすべて「1」が表示される。
	主本文*	検索結果として表示されている形態論情報の付与されている本文の主副(4.2 節参照)の別。「1」が主本文、「0」が副本文を表す。
	多重化種別*	形態論情報の多重化(4.2 節参照)を行っている箇所で、その多重化の種別を表す。本コーパスではすべて「振り仮名」である。
形態論情報 ¹¹	前文脈	「キー」の前方文脈。
	キー	検索対象の含まれるコーパステキスト ¹² 。
	後文脈	「キー」の後方文脈。
	原文 KWIC	上記項目「前文脈」「キー」「後文脈」に対する、従来の訓点資料の訳文(訓読文)で用いられてきた表示方法によるテキスト(2.2 節参照)。「キー」に振り仮名のある場合はそのテキストが本行テキストの上側に表示される。
	語彙素読み	「キー」に含まれる短単位 ¹³ の語彙素(下記項目「語彙素」参照)の読み。片仮名表記である。
	語彙素	「キー」に含まれる短単位の語彙素の表記。語彙素は、単語の様々なバリエーション(語形、活用形、表記形など)を統合した辞書の見出しに相当するもので、原則、一般の和語・漢語は漢字平仮名表記、外来語・人名・地名は片仮名表記である。

9 「キー」に対応する主本文に複数の短単位がある場合は、最後の短単位。

10 コーパス全体に高精度な付加情報を付与することが困難な場合、一部分のテキストに対して高精度な付加情報を付与し、コーパスの核となるデータとすることがある。『日本語歴史コーパス』ではその高精度なデータを「コア」と呼び、それ以外のデータを「非コア」と呼ぶ。

11 形態論情報の個々の項目の内容は、「前文脈」「キー」「後文脈」「原文 KWIC」「原文文字列」「振り仮名」を除き、UniDicの項目に対応している。各項目の詳細については、国立国語研究所コーパス開発センター(池上尚)(編)(2016)を参照のこと。

12 「キー」の表示範囲は設定により変更可能であるが、ここでは初期設定で表示される「キー」の範囲を指して言う(表中の他の項目も同様)。初期設定では、短単位検索の場合は検索対象とした短単位と対応する本行テキスト1短単位分、文字列検索の場合は検索対象文字列に対応する本行テキスト1短単位分(検索対象文字列が複数の短単位と対応する場合は最後の1短単位分)が表示範囲となる。

13 「キー」に対応する本文に主本文・副本文の2種がある場合は、「主本文」項で示された主副の別に対応する短単位を指す。また、「キー」に対応する本文に短単位が複数ある場合は、最後の短単位を指す。以下の「語彙素」「語形」「品詞」「活用型」「活用形」「原文文字列」も同様。

情報種別	項目名	内容
	語形	「キー」に含まれる短単位の語形。語形は、語彙素では統合される語形の別（例：語彙素「甚深」に対する「ジンシン」「ジンジン」など）や活用型の別（例：語彙素「有る」に対する「アル（五段-ラ行）」「アリ（文語ラ行変格）」など）等を区別した語の個々の形に相当するもの。片仮名表記である。
	品詞	「キー」に含まれる短単位の品詞で、UniDic の体系に基づく。学校文法における「形容動詞」は、語幹が「形状詞」、活用語尾が「助動詞」に分割される点に注意が必要である。 UniDic の体系に基づかない特殊な品詞には以下の種類がある。 対象語無し …形態論情報の多重化（4.2 節参照）を行っている箇所の副本文の文字列。 なお、この特殊な品詞の付与された短単位は、形態論情報に関する項目のうち、「前文脈」「キー」「後文脈」「原文 KWIC」「品詞」「原文文字列」「振り仮名」以外は空欄となっている。
	活用型	「キー」に含まれる短単位の活用の型。活用語の場合のみ表示される。
	活用形	「キー」に含まれる短単位の活用形。活用語の場合のみ表示される。
	原文文字列	「キー」の本行テキストの、従来の訓点資料の訳文（訓読文）で用いられてきた表示方法によるテキスト（2.2 節参照）。
	振り仮名	「キー」の本行テキストに対する振り仮名のテキスト。
本文情報	本文種別	検索対象の含まれる文が「地の文」以外の場合の、その種別。以下の種類がある。 会話 …会話・独話・心内発話等の引用 会話-韻文 …「会話」の内、形式が韻文（偈頌など）の部分
	話者	上記項目「本文種別」が「会話」「会話-韻文」の場合の話者名。不明の場合は「*」で示す。
作品情報	ジャンル	検索対象の含まれるサンプルの、文章内容に基づく分類。本コーパスはすべて「訓点資料」である。
	作品名	検索対象の含まれるサンプルが収録された訓点資料名。
	成立年	検索対象の含まれるサンプルが収録された訓点資料の加成年。
作者情報	作者	検索対象の含まれるサンプルが収録された訓点資料の原典著者・加点者の名。原典著者名は後ろに「(作)」を付け、加点者名は後ろに「(加点)」を付けて示す。名が不明の場合は「*」で示す。複数人の場合は、「/」で区切る。 原典著者・加点者が「国立国会図書館典拠データ検索・提供サービス (Web NDL Authorities)」に収録されている場合は、そのウェブページへのリンクを付与する。
	生年	検索対象の含まれるサンプルが収録された訓点資料の原典著者・加点者の生年。西暦 4 桁で示す。不明の場合は「*」で示す。複数人の場合は、「/」で区切る。 生年の認定は、各種人名事典・Web NDL Authorities 等に基づく。

情報種別	項目名	内容
	性別*	検索対象の含まれるサンプルが収録された訓点資料の原典著者・加点者著者の性別。不明の場合は「*」で示す。複数人の場合は、「/」で区切る。
底本情報	底本	検索対象の含まれるサンプルが収録された訓点資料名。資料名・巻名を示す。
	ページ番号	検索対象の底本における出現箇所。西大寺本『金光明最勝王経』では、卷子本である底本の本紙の通し番号（紙数）を示す ¹⁴ 。 なお、コーパステキストは漢文原文の訓読文であるため、底本での漢字出現順序と訓読文での漢字出現順序が異なる場合がある。西大寺本『金光明最勝王経』では、卷子本の1枚の本紙の冒頭の漢字から、次の1枚の本紙の冒頭の漢字の直前までの範囲のコーパステキストの文字列に同一ページ番号を付与する。よって、返読により訓読文を作成している箇所では、「ページ番号」と実際の紙番号とがずれている場合もあることに注意されたい。
その他	底本リンク	検索対象の底本画像へのリンク。本コーパスでは該当画像がないため空欄。
	参照リンク	検索対象の底本以外の参照本画像へのリンク。本コーパスでは該当画像がないため空欄。

謝辞

『日本語歴史コーパス 平安時代編Ⅱ訓点資料』は、国立国語研究所共同研究プロジェクト「通時コーパスの構築と日本語史研究の新展開」、JSPS 科研費 18H00674 「訓点資料訓読文コーパスの構築と古代日本語史研究の革新」、人間文化研究機構広領域連携型基幹研究プロジェクト「異分野融合による総合書物学」国立国語研究所ユニット「表記情報と書誌形態情報を加えた日本語歴史コーパスの精緻化」の成果の一部である。

参考文献

- 小木曾智信（2017）「多重の読みを持つテキストのコーパス化」『言語資源活用ワークショップ発表論文集』（1）、pp.159-162
- 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)、pp.727-748
- 沖森卓也・三省堂編修所（編）（2014）『三省堂 五十音引き漢和辞典 第二版』三省堂
- 春日政治（1942）『西大寺本金光明最勝王経古点の国語学的研究』岩波書店（春日和男（編）（1985）『春日政治著作集 別巻』（勉誠社）に拠る）
- 国立国語研究所（2022）『日本語歴史コーパス』バージョン 2022.3、<https://clrd.ninjal.ac.jp/chj/>
- 国立国語研究所（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）
<https://clrd.ninjal.ac.jp/chj/heian.html#kunte>
- 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』<https://clrd.ninjal.ac.jp/chj/doc/morph-heian-2016.pdf>

¹⁴紙の通し番号（紙数）は、春日（1942）やその再録版の本文編に収録された写真、総本山西大寺（編）（2013）でも確認できる。

総本山西大寺（編）（2013）『国宝 西大寺本 金光明最勝王経 天平宝字六年百濟豊虫願経』勉誠出版
日本国語大辞典第二版編集委員会・小学館国語辞典編集部（編）（2000-2002）『日本国語大辞典 第二版』小学館（ネットアドバンス社提供サービス「ジャパンナレッジ」コンテンツ、
<https://japanknowledge.com/contents/nikkoku/>に拠る）
築島裕（編）（2007-2009）『訓点語彙集成』汲古書院

参考 URL

Mecab <https://taku910.github.io/mecab/>
UniDic <https://clrd.ninjal.ac.jp/unidic/>
国立国会図書館典拠データ検索・提供サービス (Web NDL Authorities) <https://id.ndl.go.jp/auth/ndla>
日本語史研究用テキストデータ集 https://www2.ninjal.ac.jp/textdb_dataset/