

ひまわり版「西大寺本金光明最勝王經
平安初期点訓読文コーパス」
Ver. 0.1 解説書

柳原恵津子（国立国語研究所プロジェクト非常勤研究員）

2021年11月

目次

第1章 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の概要	1
第2章 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の様	2
1. ひまわり版データの使い方	2
1.1. 「ひまわり」へのインストール方法	2
1.2. 「ひまわり」を使ったコーパスの検索方法	3
2. ひまわり版データ利用上の留意点	6
2.1. 複数の訓読文がある場合の本文について	6
2.2. 検索文字列の選択	7
2.3. ブラウザ表示画面でのヲコト点 id 表示	8
2.4. ブラウザ表示画面でのルビの記載方法	8
2.5. 単字検索の推奨	9
2.6. ヲコト点のブラウザ画面での表示について	9
参考文献	10

第1章 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の概要

本パッケージは、訓読文（XML）から作成した、文字列検索システム「ひまわり」用のデータセットである。このデータを「ひまわり」にインストールすることで、わかりやすいユーザーインターフェイスによるコーパスの検索・閲覧が可能となる。検索結果画面をクリックすると、本文のブラウザ表示の画面にリンクし、前後の文脈を広範囲にわたって確認することができる。さらにブラウザ表示の画面では、異訓、発話（心話を含む）・韻文部分の確認、漢字に付された符号類、ヲコト点・仮名点についての情報を詳細に確認することも可能である。ただし仕様上、複数字単位での検索が難しい、訓読文（XML）に記されたすべての情報をブラウザ表示画面に再現しきれていない、といった留意点がある。そのあたりも含めて、以下本コーパスについて解説する。

本解説書に記された要素・属性名は、注記がないかぎり「西大寺金光明最勝王経平安初期点 巻一 訓読文（xml）」のものである。「巻一 訓読文（xml）」で使用した要素・属性については、「西大寺本金光明最勝王経平安初期点 巻一 訓読文（xml）」解説書を適宜参照されたい。

※本パッケージはクリエイティブ・コモンズ 表示 - 非営利 4.0 国際 ライセンスの下に提供されています。

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の開発に携わったスタッフは次のとおりである。

●開発担当者

柳原恵津子、近藤明日子

●開発協力者

高田智和、小木曾智信、間淵洋子、鴻野知暁、佐伯俊源、高山倫明、
田中草大、月本雅幸

また、本コーパスは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」の構築」の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」、科研費基盤研究（B）「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」（18H00674）の成果の一部として構築されたものである。これらの助成にも厚く感謝申し上げる。

第2章 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の仕様

1. ひまわり版データの使い方

1.1. 「ひまわり」へのインストール方法

データの「ひまわり」へのインストールは次の手順で行う。

- ①データ saishookyo_himawari.zip をダウンロードする。Windows 機の場合は、saishookyo_himawari.zip を右クリックし、[プロパティ] > [全般] でセキュリティのブロックが解除されていることを必ず確認する。
- ②saishookyo_himawari.zip を解凍すると「saishookyo_himawari」フォルダが現れる。その中に次のファイルがあることを確認する。
 - Corpora フォルダ…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 データを格納したフォルダ
 - config_saishookyo.xml…「ひまわり」用設定ファイル
 - .himawari_package_info…パッケージインストール設定ファイル
- ③データの対応するバージョンの「ひまわり」をインストールする。国立国語研究所 Web サイトの「トップ」—「データベース」—「ツール」ページ (<https://www.nijjal.ac.jp/database/type/tools/>) から「全文検索システム「ひまわり」」のページに移動する。説明に従い「ひまわり」のインストールを行う。
- ④「ひまわり」をインストールすると「himawari_X」(Xには「ひまわり」のバージョンに対応した数字が入る) フォルダが現れる。その中の「himawari.exe」をダブルクリックすると「ひまわり」の起動画面(図1)が開く。画面上部の「ファイル」メニュー—「インストール」(図2)を選択し、解凍した「saishookyo_h

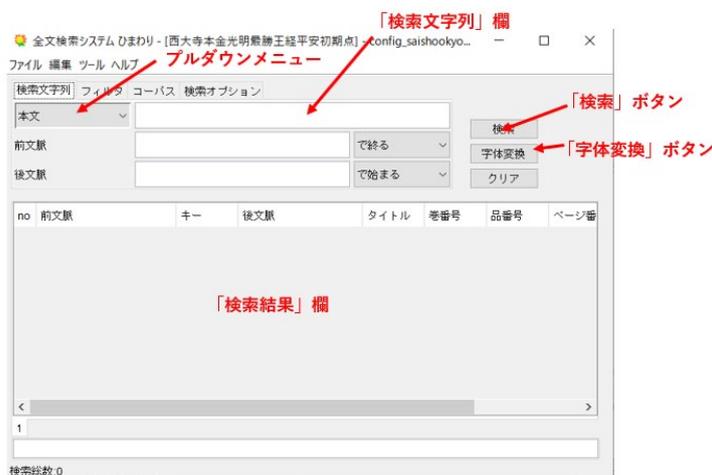


図1 「ひまわり」の起動画面

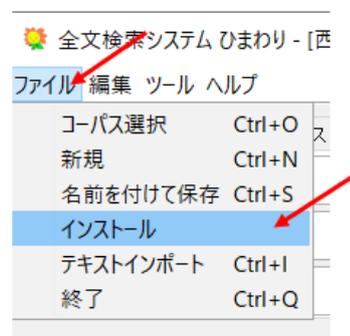


図2 「ファイル」メニュー
—「インストール」

imawari」フォルダを指定して「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 データをインストールする。

1.2. 「ひまわり」を使ったコーパスの検索方法

「ひまわり」にインストールしたコーパスデータの基本的な検索・閲覧方法を説明する。

まず、「ひまわり」の起動画面（図1）上部の「ファイル」メニュー—「新規」を選択する（図2）。設定ファイルを指定するための画面が現れるので、「config_saishookyo.xml」を選択する（本データのインストール直後や前回起動時の設定が保存されている場合は、この手順は省略できる）。次に、「プルダウンメニュー」（図1参照）で検索対象を指定する。検索対象のリストを表1としてあげる。なお、「検索文字列」欄に表示される「完全一致」「部分一致」は検索対象と検索文字列との照合方法を表す。ルビ・不読字以外の通常の本文部分を検索したい場合には「本文」を選ぶ。

表1 「ひまわり」検索対象リスト

プルダウンメニュー表示	検索対象
本文	本文テキスト部分
ルビ（完全一致）	rt要素（ひまわり版corpus.xmlでは、ruby要素
ルビ（部分一致）	rubyText属性）
不読字（完全一致）	mute要素text属性
不読字（部分一致）	

次に「検索文字列」欄に検索したい文字列を入力する。漢字文字列を検索する場合、「字体変換」ボタンをクリックすると、入力文字列に異体字がある場合は「検索文字列」欄の入力文字列が変換される（図3）。

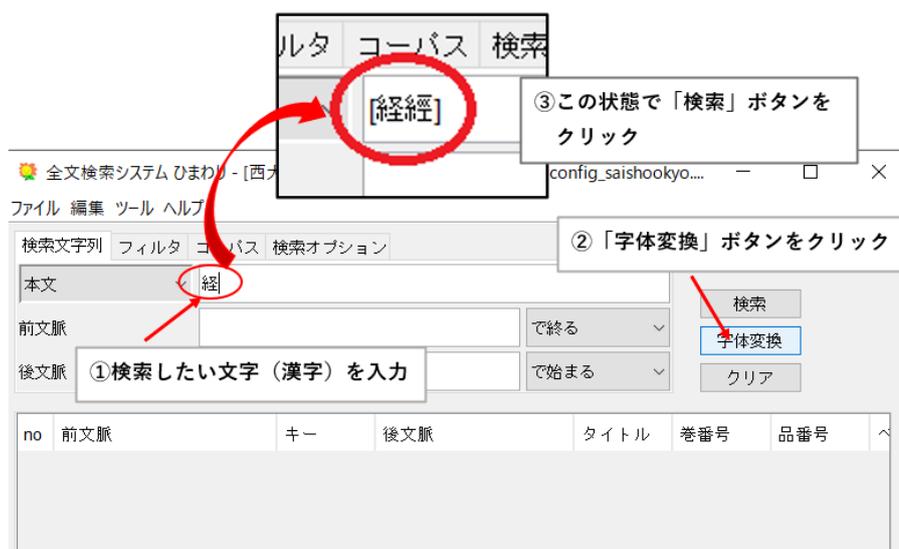


図3 「検索文字列」欄入力と「字体変換」ボタン

そのまま「検索」ボタンをクリックすると「検索結果」欄に検索結果が KWIC 形式で表示される（図 4）。

「検索結果」欄に表示される列のリストを表 2 として示す。

全文検索システム ひまわり - [西大寺本金光明最勝王経平安初期点] - config_saishookyo.xml
 ファイル 編集 ツール ヘルプ

検索文字列 フィルタ コーパス 検索オプション

本文 [経経] 検索

前文脈 [] で終る

後文脈 [] で始まる

字体変換

クリア

no	前文脈	キー	後文脈	タイトル	巻番号	品番号	ページ番号	文章種別	訳文番号
1	勝ること王のごとくいます	経	序品第一 三藏法師義浄	西大寺本...	01	01	004		
2	行と精勤とをも無量の劫を	経	てし、 諸に超せる静	西大寺本...	01	01	005		
3	妙法は、 最勝にして諸の	経	の王といます。 甚深	西大寺本...	01	01	009	verse	
4	を 著 著す應し。 此の妙	経	王の 甚深にして佛の	西大寺本...	01	01	010	verse	
5	に大衆の爲に、 是の如き	経	を宣へ説かむ。 并せ	西大寺本...	01	01	010	verse	
6	もせよ 受持せよ。 此の	経	の威力に由りて、 能	西大寺本...	01	01	010	verse	
7	のたまふ。 金光明最勝王	経	如來壽量品第二 尔	西大寺本...	01	02	011		
8	深行の諸の菩薩との 持	経	者を擁護して、 諸の	西大寺本...	01	01	011	verse	
9	沙に過ぎたること、 是の	経	を誦讀せむ者は、 當	西大寺本...	01	01	011	verse	
10	し尊重の心を以て、 是の	経	を聽聞せむ者は、 善	西大寺本...	01	01	011	verse	
11	を起さしめよ。 若し是の	経	を聽かむと欲はば、	西大寺本...	01	01	011	verse	
12	。 若し有らむとは是の	経	を聞き、 能く他の爲	西大寺本...	01	01	011	verse	
13	むることを爲らむ。 是の	経	を供養せむ者は、 前	西大寺本...	01	01	011	verse	
14	してせむ。 我れ當に是の	経	の 甚深にして佛の行	西大寺本...	01	01	011	verse	
15	讀たるい、 方にして是の	経	、 懺悔の法をば開く	西大寺本...	01	01	011	verse	
16	さ不。 如來の所説の甚深の	経	典をも、 亦た受持し讀	西大寺本...	01	02	015	speech	
17	生さしめ、 佛世尊の所説の	経	教の於に、 速く當に受	西大寺本...	01	02	015	speech	
18	りといふ想を生じ、 所有る	経	典をも悉く皆受持して	西大寺本...	01	02	017	speech	
19	、 釋迦牟尼如來、 今金光明	経	の甚深の法要を演説し	西大寺本...	01	02	017	speech	
20	常に鷲山に在りて、 此の	経	寶を宣説す、 衆生を	西大寺本...	01	02	018	verse	
21	哉釋迦牟尼如來、 今金光明	経	の甚深の法要を演説し	西大寺本...	01	02	018	speech	
22	を至して是の金光明最勝王	経	を聽きたてまつる應し	西大寺本...	01	02	019	speech	
23	ること能は不所なり。 此の	経	は能く無量無邊の福德	西大寺本...	01	02	019	speech	
24	聽きたてまつる應し。 諸の	経	の中に最も殊勝にいま	西大寺本...	01	02	019	speech	
25	無き者とならば、 云何にぞ	経	の中に、 涅槃すること	西大寺本...	01	02	022	speech	

図 4 検索結果画面

表 2 「ひまわり」検索結果列リスト

列名	備考
前文脈	
キー	
後文脈	
タイトル	text要素title属性に対応
巻番号	text要素volume属性に対応
品番号	article要素nb属性に対応
ページ番号	pb要素n属性に対応
文章種別	quotation要素type属性に対応
訳文番号	yakubun要素nb属性に対応

「検索結果」欄の任意のセルをダブルクリックすると、Web ブラウザが起動し、当該用例前後の訓読文が閲覧できる（図 5）。ブラウザでは、「検索結果」欄でクリックした行の当該例が黄色マーカー付き文字で表示され、西大寺本金光明最勝王経平安初期点巻 1 の訓読文全体が閲覧できる。ブラウザ画面に表示された訳文の凡例を表 3 にあげる。

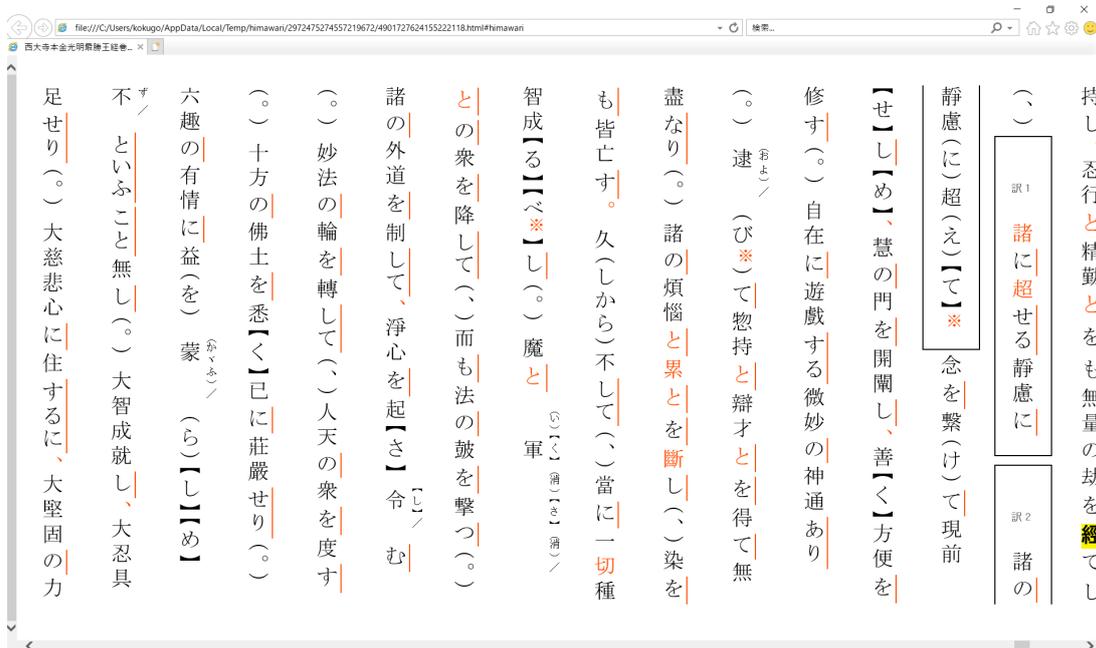


図5 ブラウザでの本文閲覧

表3 「ひまわり」検索結果ブラウザ画面訳文凡例

表記	説明	ポップアップ表示
漢字（黒、ポップアップ表示のある場合は赤橙）	本文漢字	kanji要素の各属性の表示
漢字（ [] ）	不読字	
平仮名（黒、ポップアップ表示のある場合は赤橙） + 赤橙傍線	okototen要素の内容（平仮名）とヲコト点の区切り（赤橙傍線）	id（当該ヲコト点が表示範囲がルビと本文に跨る場合）・position属性（「と」のみ）の表示
平仮名（ [] ）	kanaten要素の内容（平仮名）と仮名点1字の区切り（ [] ）	
平仮名（ () ）	補読された本文	
背景青	quotation要素（引用）の内容	
枠線	yakubun要素（異訓並記）の内容	
ルビ ※ルビ内の括弧などについては、第2章2の2.4（p.8）参照	rt要素（ひまわり版corpus.xmlでは、ruby要素rubyText属性）に対応	
句読点	period要素に対応	句読点位置の表示
※	ポップアップあり	v Mark要素・g要素、kanaten要素position属性の表示
*	ポップアップあり	betsuhitsu要素の表示

プルダウンメニュー表示で「ルビ（完全一致／部分一致）」、「不読字（完全一致／部分一致）」を選択すると、ひまわり版 XML でタグ内の属性値として記されているため「本文」検索ではヒットしないルビや不読字部分の用例を検索することが出来る。

ただし、ルビの場合はキー欄に当該ルビの付された漢字が記され（図 6 上）、ブラウザ画面では当該漢字が黄色マーカー付きで表示される（図 7 右）。不読字の場合、キー欄は空欄となり（図 6 下）、ブラウザ画面では当該不読字の位置は色付きでは示されず、角括弧内に表示される（図 7 左）。

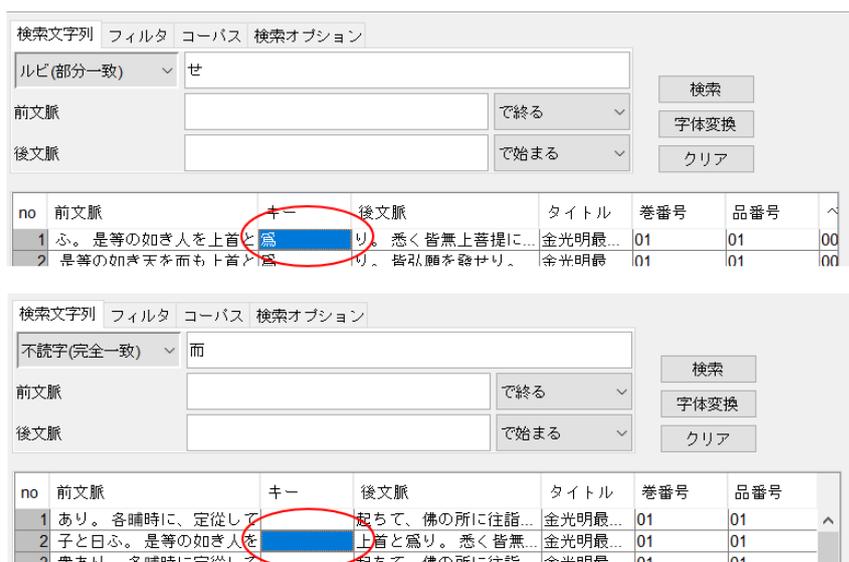
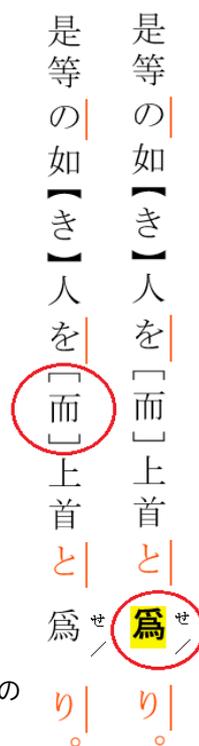


図 6 「ルビ (「せ」)」（上）、「不読字 (而)」(下) 検索結果画面

図 7 「ルビ (「せ」)」（右）、「不読字 (而)」(左) のブラウザ表示画面



2. ひまわり版データ利用上の留意点

本データには、単語・形態素ごとの品詞や語形、活用などについての形態論情報が付与されていない。それに加え、訓点類・訓読文を表記するための括弧類やルビが多い、複数の読みが振られている部分がある、といった訓点資料特有の性質を抱えてもいる。そのため、通常の資料と異なる処理を行った点がある。以下にあげる本コーパス利用上の留意点を確認した上で、ひまわり版で検索する際には、検索文字列を 1 字、あるいは極力短い単位とすることを推奨する。

2.1. 複数の訓読文がある場合の本文について

本資料にはヲコト点と仮名点、右傍仮名点と左傍仮名点、通常の筆と大字や擦り消しなどの書き分けによって、複数の読みが加点されている漢字文字列が多くある。春日(1942)は仮名点の筆致を7種ほどあるとした上で、「これら異筆の仮名は少なくとも本点の時代と同一のものとして取り扱つても、さして無理は生じない¹⁾」と述べている。本コーパスでは煩雑な読み分けを避け、「大字」「小字(通常の子)」の2種に区別してタグ付けしているが、検索をする際には、いずれも同時代、等価値の用例として検索対象にする必要がある。そのため、複数の種類の読みが記された箇所については、すべての読みを連続して検索用本文(検索に使用され、検索結果欄の前文脈・キー・後文脈に表示される本文)に並記することにした。

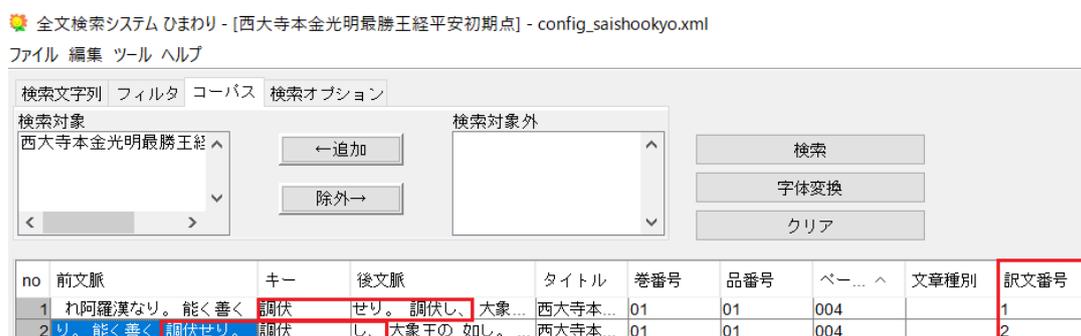


図8 複数の読みがある箇所の検索結果画面表示

図8の「前文脈」「キー」「後文脈」赤枠部分で示した部分が、2つの読み(調伏せり、調伏し)が加点された本文部分である。異訓の存する場所は「訳文番号」欄(図8右端赤枠部分)に、「1」「2」「3」のように並記した訳文のうちのいくつ目の訳文に該当するかが数字で記されている。ここに記載がある場合は、キーを含む前後の文脈に異訓があり、すべての訳文が並記されている。ブラウザ画面上では、複数の読みのある箇所は訳文ごとに枠線で囲って示す(図9)。

訓読文を作成する際、ひとつのヲコト点や仮名点を複数の訳文の中で読むことは避け、2つ目以降の訳文では補読として読む。したがって、あるヲコト点・仮名点で加点された語が、ヲコト点・仮名点で記された例として複数回ヒットすることはない(2つ目以降は補読の例となる)。

2.2. 検索文字列の選択

「而」「於」のように本文として読まれる場合と不読字とされる場合がある漢字は、読まれている例は検索文字列で「本文」を選択した際、不読字の例は「不読字(完全一致/部分一致)」を選択した



図9 複数の読みのブラウザ画面表示

¹⁾ 本文篇 p.20(著作集版(春日(1985))). 引用に際し、旧字体の漢字を新字体に改めた。

際にしか検索出来ない。よって、「而」「於」のような漢字の全用例を見るには、「本文」「不読字（完全一致／部分一致）」双方で検索しなければならない。

2.3. ブラウザ表示画面でのヲコト点 id 表示

「といふ」「います」などルビと本行にまたがる範囲の読みがひとつのヲコト点で加点されている際には、okototen タグにひとつのヲコト点で記された内容であることを示す id 属性を付与している。ひまわり版ブラウザ表示では、本行にはこの id 属性をポップアップ表示内に示しているが、ルビ部分には示していない(たとえば、図10「います」「といふ」の赤橙部分(「す」「と」「ふ」)のみポップアップ表示あり)。本コーパス内では、「といふ」「います」「せり」「せる」「する」「なり」のヲコト点による加点部分に、このような id 付与を行った例がある。



図10
ブラウザ画面での
okototen 要素 id
属性表示

2.4. ブラウザ表示画面でのルビの記載方法

仮名点・ヲコト点・符号類の別筆や擦り消し、仮名点の記載位置に関する情報を、ブラウザ画面の本行部分ではポップアップ表示内に記している。ルビ部分ではそのような表示が仕様上できないため、ルビそのものに符号類を記して、別筆・擦り消し関係、仮名点の位置についての情報を示す。ルビテキスト関係→仮名点記載位置関係→別筆関係、の順で、表4 凡例の要領で示す。また、実際のルビ表示の例を図11に示す。

表4 ブラウザ表示画面のルビ表示凡例

	ルビ内符号・括弧	説明
ルビテキスト関係	無標 ※平仮名	ヲコト点による読み
	【 】 ※括弧内平仮名	仮名点による読み
	() ※括弧内平仮名	作業による補読の読み
	<え> <こ>	ヤ行「え」(巻1該当例なし)、上代特殊仮名遣い乙類「こ」
	/	当該漢字列へのルビ末尾
仮名点記載位置関係	(左)	左傍仮名点
	(右2)	右傍二つ目の仮名点
別筆関係	(消)	擦り消し(通常の筆の仮名点・ヲコト点)
	(大消)	擦り消し(大字)

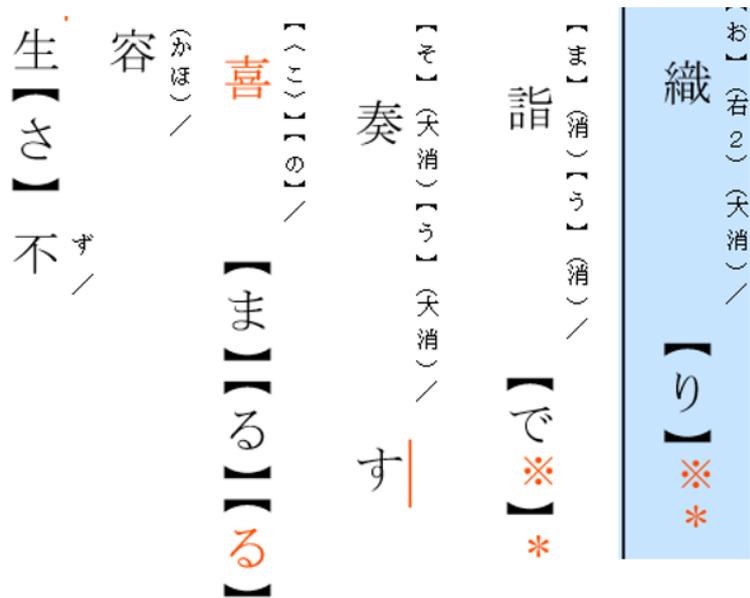


図 11 さまざまなルビ表示例

2.5. 単字検索の推奨

① 仮名文字列がルビと本文にまたがる場合

活用語の語幹と語尾など、ルビ部分と本行にまたがる範囲を指定して文字列検索することはできない。

【し】む (図 5 11 行目参照・「令」へのルビ)

→ 「しむ」での検索不可 (ルビ検索「し」または本文検索「令」などで検索可能)

② ルビ内の文字列で、仮名点・補読の括弧がはさまる場合

ルビは括弧などを含めたテキストの形で ruby 要素の属性値としているため、括弧などが間にはさまる例は文字列検索 (括弧などなしで検索した場合) では検索できない。

(い) 【く】(消) 【き】(消) (図 5 9 行目参照)

→ 「いくき」での検索不可 (本文検索「軍」、または1字ずつでルビ検索、などで検索可能)

2.6. ヲコト点のブラウザ画面での表示について

ブラウザ表示画面では、原本でヲコト点によって表示されている部分は、赤橙色の右傍線でそのヲコト点で示された内容と範囲を示す。よって、たとえば図 5 の 2 行目「にして」は「に」「して」の 2 つのヲコト点で、同行から 3 行目にかけての「とありますが」は「と」「います」「が(か)」の 3 つのヲコト点で加点されていることがわかる。

また、上記「いますが」の「が」下の「※」にカーソルを当てると、ポップアップ表示で「濁点なし」の表示がでる。これにより、当該例の「が」は「か」のヲコト点、つまり「何をか」「無かりけり」などの「か」と同じ点で記されていることが解る。

参考文献

春日政治[編]（1942）『西大寺本金光明経古点の国語学的研究』斯道文庫（ただし、作業の際には『春日政治著作集』別巻（1985年、勉誠社）を用いた）。

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