

文字情報と図版情報を有する

近世版本コーパスの構築とその応用

間淵 洋子 (人間文化研究機構 国立国語研究所)

国立国語研究所では、江戸時代後期の版本を対象としたコーパスを構築している。このコーパスは、表記や図版の情報を中心に、原典の版面の情報に関して詳細なアノテーションを施したもので、新奇的な試みとなるものである。本稿では、このコーパスの特長とデータの仕様について紹介すると共に、構築と研究利用の実践報告として、表記アノテーションの方法とアノテーション試行結果の提示を行う。更に、本コーパスの応用として開発を進めている、字母と Unicode とで層別した字形（文字画像）データベースと、前近代の書物に親しむ契機として提供する一般向け Web サイトの構想について述べる。

The Construction and Application of a Corpus Annotated with Information on Orthography and Inset Figures Sourced from Books Published in Edo-period

MABUCHI Yoko (National Institute for Japanese Language and Linguistics)

The National Institute for Japanese Language and Linguistics has constructed a corpus of books printed in the late Edo period. This corpus is a novel undertaking with detailed annotation on information in the original text, focusing on orthography and graphics. In this report, I introduce the features of this corpus and the specifications for its data. I also present details about orthographic annotation and annotation trial results, as a practical report on construction and research utilization. In addition, I describe the conceptualization of two applications of this corpus: a database of images of characters indexed by base form and by Unicode ID, and a website for the lay-person to familiarize themselves with pre-modern Japanese texts.

1. まえがき

現在、国立国語研究所では、『日本語歴史コーパス』(The Corpus of Historical Japanese, 以下 CHJ) に収録する 1 資料として、江戸時代後期の人情本を対象としたコーパス(「CHJ 江戸時代編 II 人情本」)を構築中である。既に公開済みの上代から近代にかけての資料と共通の形態論情報を付与し、日本語の通時研究や近世語研究に資するコーパスとして整備を進めているものであるが、この人情本のコーパスは、底本となる版本の画像収集と、それに基づく忠実な翻刻を経てコーパス本文を作成した点で、他資料と一線を画す[1]。この試みは、前近代の日本語表記の特徴である行草体や連綿、異体仮名等の表記に関する情報、文字を補完する挿絵等の図版要素に関する情報といった、原資料の紙面情報を豊富に有し、より多角的で精緻な言語研究あるいは人文情報学分野の研究に資するデータとして、本コーパスを整備することを意図して行われているものである。

そこで、本発表では、近世版本の紙面情報を精緻に写した文字・図版情報付き電子化テキスト「表記・画像情報付き版本コーパス」の作成とその応用について述べる。

2. 文字・図版情報付きコーパス構築の意義

今回電子化対象とした江戸時代後期の人情本のように、活版印刷以前の近近代資料は、筆記を版下とした木版印刷に由来する連綿文字列(複数の文字を連続して送筆したもの)、行草体、多種多様な異体仮名の使用など、表記面で近代以降の資料と大きく異なっている。これまでの表記史や書記史に関する研究においては、これらの前近代資料にみられる表記特性について、文字の連綿と語や句、表現の切れ続きとの相関[2]、異体仮名の用字傾向(仮名文字遣い)[3]といった国語学的な指摘がなされてはきたが、計量的な検討から実態を明らかにするような研究は多くなかった。

そこで、国立国語研究所では、2016年度より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」の元で、文献学と言語計量的手法により、言語単位と表記・書記単位、また書物や版面の形状との相関関係を明らかにすることを目的に、言語・表記(文字)・書物の重層構造を精緻に記述した言語コーパスのプロトタイプ構築を目指すことと

した。その中で、詳細な表記情報を付与し表記・書記分野の計量的研究に資するデータの電子化仕様を検討し、連綿文字列や変体仮名の字母情報の付与に関する試行を行ってきた[4][5]。一方で、変体仮名の電子化を目指し、国際規格 ISO/IEC 10646 への収録提案など、文字コード標準化の取り組みも行ってきたが、これについては、昨年6月出版の Unicode 10.0 に変体仮名 285 が収録されるという、大きな進展を見た。変体仮名の Unicode への収録は、前近代資料の文字情報処理において極めて有用と思われるが、実装や研究利用を通して、文字セットの妥当性や運用の問題点などについて検証が必要となるはずである。

そこで、本研究では、現在整備中の人情本のコーパスに対して、従来国語研で『日本語歴史コーパス』の資料記述用に定義してきた XML 文書型を拡張し、原典の文字・表記・書記の状態をできる限り精緻に捉えるためのアノテーションを試みた。

さらに、版面の情報をテキストの解釈や理解に役立てる「総合書物学」の観点から、表記情報に加えて、図版の情報付与を試みることにした。

3. コーパスの概要と仕様

3.1 表記・図版情報付きバージョンコーパスの概要

本研究において構築中の「表記・画像情報付きバージョンコーパス」は、CHJ に格納するための人情本のコーパスの一部に対して、CHJ では提供していない詳細な表記情報と画像情報を付与して公開する予定の、サブセットである。

表 1 に、CHJ に格納対象となる人情本の資料を編・巻数、語数と共に示す。

表 1 CHJ 江戸時代編 II 人情本の採録資料

作品名	編・巻	語数
比翼連理花酒志満台	初編(3巻)	12,801
	二三四編(9巻)	38,839
春色梅児与美	初編(3巻)	12,056
	二三四編(9巻)	35,366
おくみ惣次郎春色江戸紫	初二三編(12巻)	30,511
梅暦余興春色辰巳園	12巻	47,947
小三金五郎仮名文章娘節用	前後三編(9巻)	43,986
浮世新形態の花染	初二三編(9巻)	35,850
春色連理の梅	初二三四五編(15巻)	52,130
浦里時次郎明鳥後の正夢	初二三四五編(15巻)	67,911

本研究で表記・図版アノテーションを行うのは、このうち、網掛けにした「比翼連理花酒志

満台」初編全3巻約12,800語と、「春色梅児与美」初編全3巻約12,000語の、計約50,000語のデータセットである。

3.2 表記・図版情報付きバージョンコーパスの仕様

本コーパスは、CHJ のサブセットであるため、CHJ が有する研究用付加情報をすべて持つ。すなわち、JISX0213 に準ずる文字セットによる原文文字列情報、形態素解析用本文情報、原文校訂情報、頁・行の開始位置やルビ文字列・割書等の一部版面情報、引用・発話・文等の文書構造情報、その他書誌情報等メタ情報、そして語の切れ目、読みや品詞、活用等を含む形態論情報である。これらは、XML 形式で構造化され表 2 に示すようなタグセットで記述されている。

表 2 人情本コーパスのタグセット (抜粋)

情報種別	タグ(要素)	説明
構造	corpus	コーパス全体
	text	テキスト1冊のまとまり
	body	本文
	p	本文のひとかたまり
	s	文
	speech	会話文
	quotation	引用、手紙など
	speaker	話者
版面	pb	頁開始位置
	lb	行開始位置
	warigaki	割書き
	r	右ルビ
文字	goji	合字
	g	外字などの特殊文字
	unclear	原本の不鮮明箇所
校訂	odoriji	踊り字展開箇所
	vMark	濁点無表記箇所
	corr	本文修正箇所
形態論	SUW	短単位

本研究では、これらの情報に加えて、詳細な文字情報と図版情報を格納するため、CHJ の XML 文書型に対して、拡張した仕様を用いることとした。本研究のために拡張して追加・変更した要素は以下の通りである。

【表記情報用拡張仕様】

rb 要素：連綿文字列の切れ目を示す。

kana 要素：現行仮名と異体の仮名(変体仮名)を表す。属性として、字母の情報(jibo)と文字コードの情報(ref)を持つ。

choice 要素：テキスト中の同じ場所で、異なる符号化記述をまとめる．原文 (sic 要素) と校訂本文 (corr 要素)，本行文字列 (r 要素) とルビ文字列 (rt 要素)

【図版情報用拡張仕様】

figure 要素：図版とそれに付随する要素を示す．

graphic 要素：図版の URL を示す．

head 要素：図版のタイトル相当の文書要素を示す．

caption 要素：図版に付随する文書要素 (タイトル以外) を示す．

figDesc 要素：subject 要素 (対象)，sex 要素 (性別)，name 要素 (名前)，hair 要素 (髪型)，clothes (着物)，additional 要素 (付随して描かれているもの．内部を item 要素で記述) により，図の内容を記述する．

```
<figure>
<graphic url="https://goo.gl/TRuuSC"/>
<head>
<s><lb/>梅の阿由が義妹竹長吉</s>
</head>
<caption><s><lb/>そも / \ 和朝に寄浄留理の元祖は
京都<lb/> 四条の河原にて (中略) たへて<lb/>なしと
ぞ</s></caption>
<figDesc>
<subject>Human</subject><sex>F</sex>
<name>お長</name><hair>若衆鬘</hair>
<clothes>袴</clothes>
<additional><item>見台</item><item>湯呑</item>
</additional> </figDesc> </figure>
```

図 3 図版アノテーションの例
 (『春色梅児誉美』巻1口絵3丁表)

図 1 に、収録対象の原本画像を、図 2, 3 に表記と図版のアノテーション例を示す．

これらの新たな要素を検討するにあたっては、構造化テキスト記述の国際標準を目指す TEI P5[6]に準ずるものとなるよう設計を行ったが、前近代の日本語テキストに特有の連綿文字列や変体仮名についての仕様については、研究目的に応じて独自の仕様とした．

4. アノテーションの実践

4.1 表記情報アノテーション

本研究では、原典の版面から得られる文字表記に関する情報として、まずはこれまでも検討を続けてきた(1)文字列の連綿の情報、(2)変体仮名の情報の2点について、アノテーションを行った．

連綿の情報については、近世文学を専攻とし、近世版本を読み慣れている大学院生アノテーター2名によって、連綿の切れ目を判定する作業を行ってもらった．切れ目の判定に関しては、できるだけ文脈等に依存せず、送筆の切れ続きに留意して行うよう指示し、2名の判断が重なった部分を「切れ目 (rb)」として認定した．また、一方のみが切れ目と判断した箇所については、原則として「切れ目なし」としたが、見落としなどの可能性も考えられるため、作業管理者 (著者) が確認の上認定することとした．

アノテーター2名による判定の突き合わせを行ったのは、版本における連綿の切れ目の判定が必ずしも容易でなく揺れの生じる可能性を含むためである．

例えば、図 4 における「貧」「し」の間には、空白があるように見え、これを連綿の切れ目と判定するアノテーターもいる．しかし、版本においては、印刷の状態による空白は、必ずしも送筆に



図 1 原本画像の例 (左：『比翼連理花廻志満台』初編上5丁表，右：『春色梅児誉美』巻1口絵3丁表)

```
<s>「どう<rb/>して<pb n="17"originalN="5オ"/><lb/>
/ \ <rb/>。</s>
<s>おいらん<rb/>で<rb/>も<rb/>か< kana jibo="奈"
ref="U+1B082">な</kana>やア<rb/>< kana jibo="志"
ref="U+1B048">し</kana>ねへ<rb/>の<rb/>ヨ<rb/>。
</s>
<s>そ< kana jibo="連" ref="U+1B100">れ</kana>
< kana jibo="可" ref="U+1B01A">か</kana>ら<rb/>ま
< kana jibo="多" ref="U+1B060">た</kana><rb/>
<choice><r>泊</r><rt>と< kana jibo="満"
ref="U+1B0C6">ま</kana><rb/></rt></choice>り
< kana jibo="越" ref="U+1B11A">を</kana><rb/>
<lb/>< kana jibo="飛"ref="U+1A0EF">ひ</kana>い
< kana jibo="天"ref="U+1B073">て</kana><rb/>
<choice><r>四</r><rb/><rt>< kana jibo="志"
ref="U+1B048">し</kana>ひや<rb/></rt></choice>
< kana jibo="可" ref="U+1B01A">か</kana><rb/>
<choice><r>五</r><rb/><rt>ごひやく
<rb/></rt></choice>も<rb/>とる
```

図 2 表記アノテーションの例
 (『比翼連理花廻志満台』初編上5丁表)

おける切れ目を意味しない。「貧」と「し」の間には明らかな連続性が見られるため、このような例は連続している例として<rb/>タグ挿入位置とは認めなかった。これらの判断の積み重ねによって、判別のルールを精緻化する必要がある。

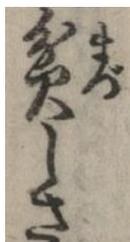


図 4 連綿の判定が揺れる例

変体仮名の情報については、同様に近世版本を読み慣れている大学院生アノテーター1名が、(1) 現行平仮名との異同、(2) 現行平仮名と異なる字形の場合はその字母（仮名の元となった漢字）、(3) 当該字形が Unicode の字体に割り当てられる場合はその字体、の3点について判定し、作業管理者（著者）がすべてを点検の後、字体を認定することとした。

大学院生アノテーターの判定と作業管理者の間で、判定に差の見られる例があった。字体の認定の揺れについての対応例を以下に示す。

- 現行仮名「し」における「し」「し」（U+1B045）
- 現行仮名「な」における「な」「な」（U+1B082）
- 現行仮名「や」における「や」「や」（U+1B0DD）
- 現行仮名「ゆ」における「ゆ」「ゆ」（U+1B0E5）
- 現行仮名「る」における「る」「る」（U+1B0FB）
- 現行仮名「を」における「を」「を」（U+1B11C）

表 3 字体の判定

音価	現行仮名	変体仮名	弁別の観点
し			右方向への運び
な			3画目の点
や			「ゆ」に近い形
ゆ			中央のとめ
る			上の横画
を			2画目の角度

その他、「て」と「乏」（U+1B073）、「も」と「も」（U+1B0DA）などにおいても、現行仮名と変体仮名との間で判定の揺れが多く見られた。

なお、字体の認定においては、豊島（1999）が指摘するように「暗黙の文字包摂を排除出来ない」が[7]、その字体包摂をできる限り詳細に記録しておくことが肝要である。本アノテーションにおいては、現行の Unicode に字体を集約することを試みたため、いくつかの字形について未だ疑念のある包摂を行っている。ここにそのいくつかを示し、今後の議論、検討の材料としたい。

【「毛」を字母とする変体仮名「毛」（U+1B0DE）】

字形としては、① ② のように、1画で書かれており、また送筆の方向も異なる。これらは、「学術情報交換用変体仮名」[8]における、字体「350020040」「350020050」にそれぞれ対応するものであるが、現行仮名との差異と草体化の進度から U+1B0DE に包摂した。なお、両字形は、当該資料において下方への連綿があるものは①、ないものは②と相補分布している。

【「本」を字母とする変体仮名「本」（U+1B0C0）】

当該資料においてはより草体化が進んだ「不」に近い字形 であり、「学術情報交換用変体仮名」における「300050020」に対応するものであるが、字母が同一であることから U+1B0C0 に包摂した。

このような、字体同定における検討や包摂の基準の整理も、変体仮名の標準規格化における評価と運用において重要な課題である。

4.3 表記アノテーションの結果

現在、連綿のアノテーションは全て終了し、変体仮名のアノテーションについては、アノテーターの作業が一通り終了し、作業管理者の点検と基準の作成、整理、集計を行っている。

本稿では、アノテーション結果の利用に関して、変体仮名アノテーション結果について一部を報告する。集計の済んだ『比翼連理花廻志満台』初編巻一のアノテーション結果を稿末の付録に示す。これを用いて、以下に本アノテーションの特長を活かした簡単な調査結果の提示を行う。

本コーパスは、ルビと本文行とを区別した上で、字体判別の集計を行うことができるため、本文行とルビとで使用に有意差のある字母や字体の存在に関しても、計量的に示すことができる。

例えば、現行仮名「か」においては、本文行で現行仮名「か」(56例)と「可」に由来する変体仮名「う」(212例)「ゝ」(3例)が用いられるものの、ルビではほぼ「う」の使用に限られる(105例中104例)といった事象が、計量的に明らかになる。その他、「こ」における「み」、
「し」における「し」、
「す」における「は」なども同様に本文行にのみ現れる字体が多く、これらは概ね、ルビにおいてより草体化の高い字体が選好されていることを示している。

ルビにおいて特定の草体化の高い字体が用いられ、本文行においてバリエーションの豊富な字体が用いられていることは、これまで国語学的に扱われてきた仮名字体の使い分けにおける「仮名文字遣い」のような、言語形式や内容に基づく字体の使い分けとは異なる、物理的な観点による分析も必要となることの証左となる事象として捉えることができる。

その他、連綿の内部または頭・末尾における使用、行における位置、序文・跋文と本文など、書物としての物理的位置における仮名字体の使い分けは、今後分析を進める予定である。

5. コーパスの応用

5.1 字形データベースの構築

本研究では、文字・図版情報付きコーパスの応用として、2つのコンテンツ開発を検討している。

その一つは、字母と Unicode で層別した字形(文字画像)データベースの構築である。現在人文学オープンデータ協同利用センターが公開する日本古典籍字形データセット[9]は、翻字現行文字ごとにくずし字の字形データを収録しているが、すべての変体仮名に Unicode 情報を付与することで、字体や字母の別により字体を参照することができ、くずし字の機械認識においても、有用なデータとなりうる。



図4 字形データベースのコンテンツイメージ

5.2 版本に親しむ Web サイトの構築

本コーパスの応用として検討しているコンテンツの今一つは、一般向けに前近代の書物に親しむ契機として提供する Web サイトの構築である。具体的には、人情本の資料性(会話中心の現代で言うライトノベル的性質)を活かした娯楽的・教育的サイト(仮称『江戸のライトノベル 人情本: 春色梅児誉美 Light』)を準備中である。このコンテンツの特長は主に以下4点である。

- (1) コーパスに付与する図版の情報などを活かし、発話と人物図・場面図を対応付ける
- (2) コーパスが有する会話アノテーション(会話範囲及び発話者情報)を活かし、SNSアプリケーション LINE のトーク画面風な会話表示を行う
- (3) 内容に即し、より親しみを持てるような、ライトノベル風文体による口語訳を提供する
- (4) 文章のブロック(会話やト書きなど)を単位として、既に国語研で公開している IIF 対応原典画像へのリンクを実装する

これらの特長により、親しみやすく魅力的なコンテンツとすることで、前近代資料への興味の足掛かりを提供すること、また、原文と口語訳、原文と原典画像との対照を可能とすることで、古典やくずし字の学習ツールを提供することを目指す。

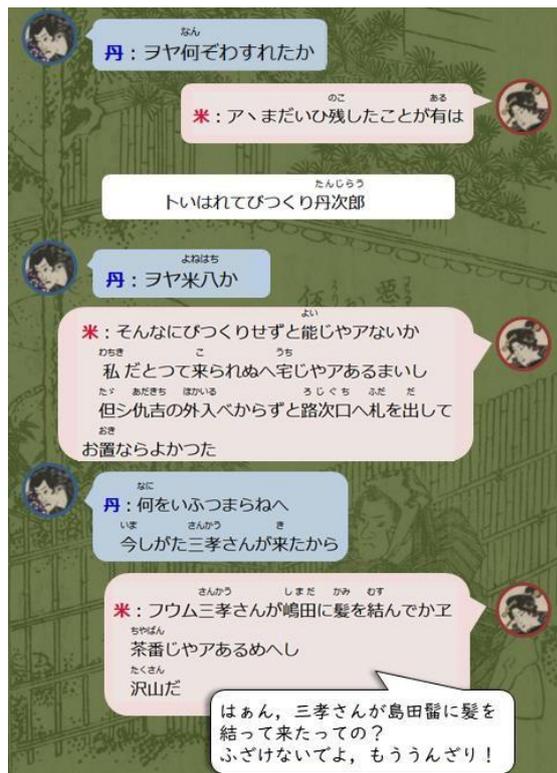


図 5 『江戸のライトノベル 人情本：春色梅児誉美 Light』のコンテンツイメージ

このコンテンツ作成に際しては、本コーパスのアノテーションに加えて、口語訳の作成が別途必要となる。現在は、コーパスの s 要素（文）を単位として、コーパスデータとは別途口語訳を作成しているが、今後は、s 要素の仕様を拡張し id 属性を持たせることで、原文と口語訳文との対応付けを行った上で、ロールオーバー時に口語訳文が表示される等の表示を実装する。

6. おわりに

本稿では、現在国語研において構築中の、詳細な表記・図版情報を有する近世版本について、その概要を述べると共に、構築におけるアノテーション実践の方法と結果を報告した。また、今後、本コーパスを言語研究や表記研究に利用するだけでなく、他分野や専門家以外の社会一般向けに応用する方策として開発している、Unicode のコードポイントによって層別した字形データベースの構築と、前近代書物に親しむための Web サイトの構築について、その概要を述べた。

『CHJ 江戸時代編 II 人情本』の公開は今年度末（2019年3月）を予定している。本コーパスは、表記・図版アノテーションを今年度中に終了し、公開される形態論情報と合わせて再構築した上で、応用コンテンツと共に、2019年中に公開することを検討している。今後もその仕様について

検討を続け、より良いコンテンツとして提供することを目指す。

付記

本研究は、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」国語研ユニット「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」による成果の一部である。

参考文献

- [1] 藤本灯, 北崎勇帆, 市村太郎, 岡部嘉幸, 高田智和. 「人情本コーパス」の設計と構築. 国立国語研究所論集, 2017, Vol. 12, p.1-12.
- [2] 小松英雄. 日本書紀史原論 補訂版, 笠間書院, 2000, 402p.
- [3] 今野真二. 仮名表記論攷, 清文堂出版, 2001, 670p.
- [4] 藤本灯, 高田智和. 「人情本コーパス」の表記情報アノテーション, 漢字字体史研究 2: 字体と漢字情報, 勉誠出版, 2016, p.222-243.
- [5] 高田智和. 国立歴史民俗博物館蔵『源氏物語』「鈴虫」の仮名字体記述: ISO/IEC 10646 提案文字による翻字シミュレーション. 古写本『源氏物語』触読研究ジャーナル, 2017, Vol. 2, p.91-102.
- [6] “TEI: P5 Guidelines”. <http://www.tei-c.org/guidelines/p5/>, (参照 2018-08-24).
- [7] 豊島正之, 横山詔一・笹原宏之・野崎浩成・エリック=ロング編著『新聞電子メディアの漢字-朝日新聞 CD-ROM による漢字頻度表』(書評), 日本語科学, 1999, No.6, p.91-102
- [8] “学術情報交換用変体仮名 公開サイト”. <https://kana.ninja.lac.jp/home/>, (参照 2018-08-24).
- [9] 北本朝展, 山本和明. 人文学データのオープン化を開拓する超学際的データプラットフォームの構築. 人文科学とコンピュータシンポジウム じんもんこん 2016 発表論文集, 2016, p.117-124.

資料（変体仮名字体判別結果一覧）

現行	字母	Unicode	UCS	本文	ルビ
あ	阿	㇀	1B004	1	
	安	あ	3042	59	57
い	以	い	3044	134	104
		𐄛	1B006	5	
う	宇	う	3046	81	125
え	衣	え	3048	3	5
		𐄜	1B011		1
お	於	お	304A	108	43
		𐄞	1B014	1	
		𐄟	1B015	1	
か	加	か	304B	56	1
	可	𐄡	1B019	3	
		𐄢	1B01A	212	104
き	幾	き	304D	31	54
		𐄤	1B025	1	
	起	𐄥	1B02A	25	27
	支	𐄦	1B026	1	
く	久	く	304F	82	85
け	介	𐄨	1B033	1	2
	計	け	3051	51	29
こ	己	こ	3053	54	121
	古	𐄬	1B038	28	
さ	左	さ	3055	91	47
		𐄰	1B03F	1	
し	志	𐄲	1B048	34	24
	之	し	3057	21	
		𐄴	1B045	152	81
す	春	𐄶	1B04F	55	24
	須	𐄸	1B050	1	
		𐄹	1B051	15	1
	寸	す	3059	2	22
せ	世	せ	305B	44	27
		𐄺	1B052	1	
そ	曾	そ	305D	78	22
	楚	𐄼	1B05B	1	
た	多	𐄾	1B05F	3	

	多	𐄾	1B060	129	100
	太	た	305F	16	
	堂	𐄿	1B05E	2	
ち	知	ち	3061	34	73
つ	川	つ	3064	38	56
		𐅀	1B069	14	
		𐅁	1B06A	35	31
	津	𐅂	1B06B	1	
	徒	𐅃	1B06D	10	
て	天	て	3066	139	46
𐅅		1B073	69		
と	止	と	3068	233	117
な	奈	な	306A	10	
		𐅇	1B080	1	
		𐅈	1B081	25	
		𐅉	1B082	135	79
な	那	𐅊	1B084	1	
に	尔	𐅋	1B08B	60	21
		𐅌	1B08C	132	8
	耳	𐅍	1B08D	1	
	仁	に	306B	5	
ぬ	奴	ぬ	306C	44	5
ね	祢	ね	306D	5	8
	祢	𐅏	1B097	14	
	年	𐅐	1B094	7	10
の	乃	の	306E	274	18
	乃	𐅒	1B099	2	
は	者	𐅓	1B0A6	66	62
	波	は	306F	6	
	八	𐅕	1B09E	168	25
	盤	𐅖	1B0A3	2	
ひ	比	ひ	3072	51	118
	飛	𐅘	1B0AF	3	
ふ	不	ふ	3075	66	44
へ	部	へ	3078	95	17
	遍	𐅚	1B0B6	1	
ほ	本	𐅛	1B0C0	24	23
ま	末	ま	307E	88	49

	満	海	1B0C5	4	
	満	毎	1B0C6	10	9
み	三	こ	1B0C9	26	70
	美	み	307F	6	7
む	武	む	3080	15	21
め	女	め	3081	41	33
も	裳	茗	1B0DC	1	
	母	母	1B0D7	1	
	毛	も	3082	2	
		毛	毛	1B0D9	3
		も	1B0DA	167	26
や	也	や	3084	49	47
		や	1B0DD	43	12
		や	1B0DE		32
ゆ	由	ゆ	3086	3	1
		ゆ	1B0E4		1
		ゆ	1B0E5	16	18
よ	与	よ	3088	49	58
ら	良	ら	3089	119	30
		ら	1B0EE	1	
り	利	り	308A	103	42
	里	里	1B0F6	13	2
る	流	流	1B0F8	3	
	留	る	308B	54	1
		る	1B0FB	49	24
	累	累	1B0FC	2	
類	類	1B0FD	1		
れ	礼	れ	308C	17	14
	連	連	1B100	89	4
ろ	呂	ろ	308D	24	31
	路	路	1B106	1	
わ	王	わ	1B10C	27	30
ゐ	為	ゐ	3090	2	12
ゑ	恵	ゑ	3091	1	17
を	越	越	1B11A	38	
	遠	を	3092	28	
		を	1B11C	31	25
ん	无	ん	3093	70	99