

国立国語研究所学術情報リポジトリ

古辞書の構造化記述の試み：『和名類聚抄』を例に

メタデータ	言語: Japanese 出版者: 公開日: 2021-07-16 キーワード (Ja): キーワード (En): Old Japanese Dictionary, structured description, Wamyō-Ruijushō 作成者: 藤本, 灯, 韓, 一, 高田, 智和, FUJIMOTO, Akari, HAN, Yi, TAKADA, Tomokazu メールアドレス: 所属:
URL	https://doi.org/10.15084/00003438

古辞書の構造化記述の試み

——『和名類聚抄』を例に——

藤本 灯^a 韓 一^b 高田智和^c

^a 京都府立大学 / 国立国語研究所 研究系 言語変化研究領域 [-2018.03] /

国立国語研究所 共同研究員

^b 韓国国立釜慶大学 大学院生

^c 国立国語研究所 研究系 言語変化研究領域

要旨

古代の日本の辞書には、様々な構造を持つものがあり、各辞書の構成や仕様を理解していなければ解読が困難な面があった。また注文から必要な情報を抽出するためには、限なく目視で検索する必要があった。順不同に入り組んだ注文の情報から、効率的に目的の情報に到達するためには、注文に存在する要素の属性が、それぞれ可能な限り定義づけられているべきである。本稿では、平安時代の代表的な漢和辞書である『和名類聚抄』を例として、いかにその構造を記述することが可能か、検討し、『和名類聚抄』の内容に適したタグを設計した*。

キーワード：古辞書、構造化、『和名類聚抄』

1. はじめに一古辞書利用の難しさー

古代の日本の辞書には、様々な構成、構造を持つものがある。それゆえに、各辞書の構成や仕様を理解していなければ解読が困難な面があった。加えて、注記・注文などから目的の情報を抽出するためには、限なく目視で検索することが必要であった。また、従来利用されてきた紙媒体の索引は、見出し漢字や和訓などを対象とするものが一般的であり、おのずから制約と限界があったが、この弱点は、プレーンテキストでも概ね同様であった。

例えば、図1の左図(次頁)のように、平安時代の代表的な漢和辞書である『和名類聚抄』¹(二十卷本、巻一、地部山谷類)にある「島嶼」という語の注文では、文の区切りが一見して明瞭でなく、「都皓反」「徐呂反」とあるのがそれぞれ「島」「嶼」各字に対する音注(反切)であることなども判別しづらいことが分かる。また「魑魅」(同右、同巻二、鬼神部鬼魅類)の項目では、小字による和名注が二箇所見られるが(「須太万」「古太万」)、「古太万」の和訓は見出しの「魑魅」

* 本稿は人間文化研究機構広領域連携型基幹研究「異分野融合による「総合書物学」の構築」における「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」(プロジェクトリーダー：高田智和)、JSPS 科研費(18K12407・若手研究・「色葉字類抄」)を中心とする国語辞書の語彙・系譜の研究・藤本灯)による成果の一部である。また日本語学会2017年度秋季大会で行ったブース発表(「古辞書の構造化記述の試みー和名類聚抄を例に一」)の内容に加筆修正を加えたものである。

¹ 以下、『和名類聚抄』の本文は、『日本語史研究用テキストデータ集』から引用する。

ではなく、直前にある「木魅山鬼」と対応するものであること²は、注文全体を読んでいく過程でのみ理解されるところである。このように、見出し語が熟語であっても、注は単字に対するものであったり、注文内の小字が見出し語に対応するものでなかったりすることは、決して稀な例ではない。

上記のような事柄は、日本語・日本文学研究者の大部分が必ず熟知しているという性格のものではなく、たとえ十分な知識があったとしても、本例のような、順不同に入り組んだ情報から目的の情報に到達する（あるいは目的の情報からそれを含む見出し語に到達する）には、多大な労苦を伴うことが、従来は当然であった。

そこで筆者らは、古辞書の利用者が、目的の情報により到達しやすい環境を整えることを目的とし、二十卷本『和名類聚抄』（元和古活字版）巻一を例として、構造化記述の方法を検討することとした。二十卷本『和名類聚抄』を対象とした理由は、画像やテキストデータが公開されていることや版本が流通したことにより、現在に至るまでよく日本人に利用され、注文の内容も充実していることによる。また『和名類聚抄』は特に、「漢文を用いた」「日本人のための」「辞書」という幾重にも特殊な要素のある文献である。筆者らの構造化記述の提案は、そのまま他の古辞書の構造化に敷衍できるものとはならないであろうが、中国の字書、古代漢文、古代日本語文、また同時代のより単純な構造を持つ日本の辞書といった多様な文献に記載された情報と適切に連携させるために、まず本書の内部を精査し、各要素を規定し、各要素の記述法の確立を目指すものである。

一方で、人文科学のテキストの符号化・交換のための規格を定める TEI コンソーシアムによる TEI P5 ガイドラインを用いた、日本古辞書の効率的な符号化モデルの設計を試みた論考として、岡田（2020）が挙げられる。岡田自身が「特定の辞書の構造に近づきすぎないことを目指して設計する」（p. 33）とするように、日本の古辞書間の構造の類似性や差異といった点についても俯瞰的に捉えようとするものであるため、本研究とは志向を異にするものと位置付けられるものの、古辞書の構造化をめぐることは、国際標準である TEI への対応も視野に入れながら幅広く検討していく必要があるであろう。上述のように、本研究は、日本の古辞書の一つである『和名類聚抄』という特殊な書物に収められた文字情報について可能な限り原本の形を再現可能な状態にして提供し、また内容についても研究者による「解釈」を含めた詳細な記述を目指すものであり、岡田のような取り組みにも大いに貢献するものと考えられる。

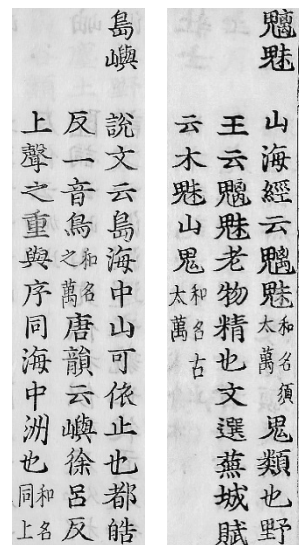


図1 「和名類聚抄 20 卷」（国立国会図書館デジタルコレクション）より
「島嶼」「魑魅」

²『箋注倭名類聚抄』（十卷本）には、「文選蕪城賦云」以下の記述がない。

的には小書きや改行，また内部的には文頭・文末などの情報も潜在している。図1「島嶼」の情報分解すれば，図4のようになる。

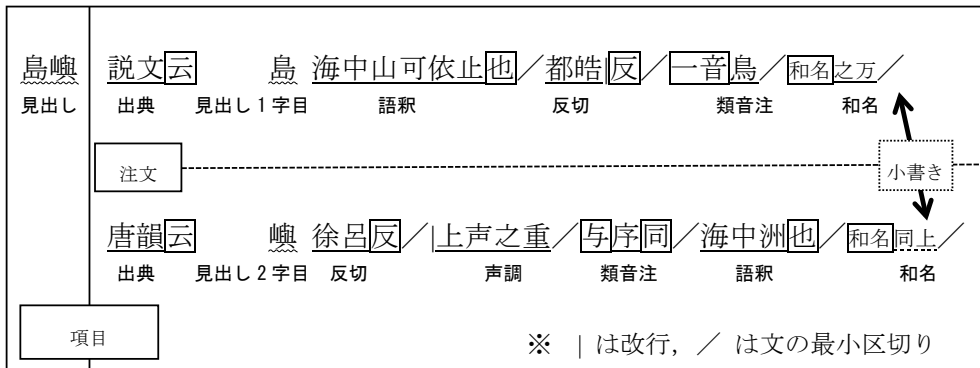


図4 「島嶼」の項目の構造

図4の注文のうち，□で囲んだ情報は，それ単独では意味を成さないが，特定の情報を示すマーカー的役割を持つものである（ただし，必須でない場合や型が一定でない場合もある）。また，図4の上半分が見出し語の一字目に対する注，下半分が二字目に対するものであると同時に，前者は『説文解字』，後者は『唐韻』を出典とするものであると記されている（和名部分を除く）。しかし『和名類聚抄』中の出典名には多く佚書が含まれており，我々現代の利用者にとって，引用部分がどこまでであるのかという線引きが決定しづらい場合も少なくない。また図4注文末の「同上」は無論これ自体が和名であるのではなく，実際には「之万（シマ）」を指すものであるが，「島嶼」の項目においては，一字目と二字目に対する和名が示されるものの，熟字への和名は存在しないということになる（ただし一般には，「島嶼」全体でも「シマ」と訓じるものと理解されるところではある）。このように，佚文中の不確定要素や，研究者に解釈の委ねられる部分が少なからずあるものの，全体としては図4で示したような明示的な情報についてマークアップすることは可能であると考えられる。次節では，仮にタグ一覧を設定し，それらを用いて『和名類聚抄』テキストにタグ付けを行う。

3. 『和名類聚抄』テキストへのタグ付与

前節に示したように，『和名類聚抄』巻一の全文につきその構造を検討し，本書に即したタグを設計した（図5）³。実際に本タグを利用してXML化した例を下に示しながら，タグの説明を行う（XML宣言はすべて省略する。また属性は今回利用したものにつき記載した）。

まず，シンプルな例として，図3にあった「日」の項目について見る。各項目を大きく<entry></entry>で括り，属性にIDを持たせる。見出し語<headword>は必須項目である。注文は，

³ タグの名称と属性の名称は仮のものである。

その注記対象が変わるごとに <note> で括り、属性に“char”（漢字文字列）と“hw”を設ける。“hw”は“char”の内容が <headword> のうち何文字目に相当するかを示したものであり、見出し語全体である場合は属性値「all」（1文字目の場合は「1」、準見出し語—しばしば見出し語下にある「〇〇附」—の場合は「sub」、見出し語全体に別字を加えた新しい熟語の場合は「add」、3文字以上による見出し語全体のうち複数文字にあたる場合は「part」、見出し語の要素に重ならない場合は「other」等）を用いている。<note> の中は原則として文 <s> の連続である。下例のように語釈 <interpretation> の文中に注記対象語が必須要素として含まれる場合は、<interpretation> 内部にその箇所を <re> でマークしておく。

項目	<entry>（属性 id）
見出し語	<headword>
準見出し語	<sub_headword>
注文	<note>（属性 char, hw）
一文	<s>
出典	<source>（属性 ref）
語釈	<interpretation>
見出し語再出	<re>（属性 hw）
字体注	<form>
反切	<fanqie>
声調	<tone>
類音訓注	<assonance>（属性 word）
その他の音注	<p>
和名	<jw>（属性 ref）
関連語	<associated_word>
注文内の小書き（割書）	<warigaki>
本行の改行	<lb/>
割書内の改行	<wlb/>

図5 古活字版和名抄用タグ一覧

○「日」の例（原文：日 造天地経云仏令宝応菩薩造日）

```
<entry id="w0101.001001.001107">
  <headword> 日 </headword>
  <note char=" 日 " hw="all">
    <s><source ref="# 造天地経 "> 造天地経 </source> 云 <interpretation>
      仏令宝応菩薩造 <re hw="all"> 日 </re></interpretation></s>
  </note>
</entry>
```

次に示す「牽牛」の例（図6）は、「出典曰 A (=) B」（A は注記対象語、B は和名や反切以外の名詞や文相当のもの）の構造を持つ例であり、この場合は B の箇所を <interpretation> で括る。注記対象語の同義語・類義語等（「関連語」）は <associated_word> で示す。割書内の情報は <warigaki><s> ～ </s></warigaki> の形で記述する。<jw> は和名、<lb/> は改行を示す。⁴

⁴ 割書内改行のマークアップについては、王・永崎（2018）に複数の提案と問題点の指摘がある。『和名類聚抄』においても、版面再現のためには必須のものと考えるが、現段階では、本行の改行と区別し、割書内改行箇所には、仮に <wlb/> でマークしておくのみとする。

- 「牽牛」の例（原文：牽牛 爾雅註云牽牛一名何鼓和名比古保之
又以奴加比保之）

```
<entry id="w0101.001011.002202">
<headword> 牽牛 </headword>
<note char=" 牽牛 " hw="all">
  <s><source ref="# 爾雅註 "> 爾雅註 </source> 云 <re hw="all"> 牽牛
  </re><interpretation> 一名
  <associated_word> 何鼓 </associated_word></interpretation></s>
  <warigaki><s> 和名 <jw> 比 <wlb/> 古保之 </jw><lb/> 又 <jw> 以奴加
  <wlb/> 比保之 </jw></s>
  </warigaki>
</note>
</entry>
```

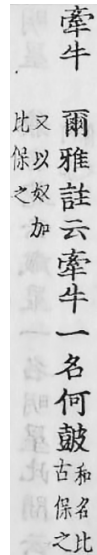


図6「牽牛」

さて、前節で取り上げた「島嶼」（図1）は、概ね上記の方法によって、次のように記述することが可能となる。反切 <fanqie> の前後では文を区切り、声調 <tone>、類音注 <assonance> を用いている。<note> の属性が「島」（hw="1", 1 字目）「嶼」（hw="2", 2 字目）のみであることから、「島嶼」全体に明示的に付された注は存在しないことが分かる。また、<jw ref="# 之万 "> 同上 </jw> とすることで、注文内に潜在している「同上（= 之万）」の情報も ID 参照により与えている。なお「同上／上同」とある箇所、本来あるべき語を注記するタグ属性“word”は、『和名類聚抄』全巻を通して、大部分が「訓」「和名」について用いられるものであるが、稀に「類音注」等を指す場合もある（巻十三「牒」の項に「音同上」）。

- 「島嶼」の例（原文：島嶼 説文云島海中山可依止也都皓反一音鳥和名之万
唐韻云嶼徐呂反上声之重与序同海中洲也和名同上）

```
<entry id="w0102.004011.007203">
<headword> 島嶼 </headword>
<note char=" 島 " hw="1">
  <s><source ref="# 説文解字 "> 説文 </source> 云 <re hw="1"> 島 </re><interpretation> 海中山可依
  止也 </interpretation></s>
  <s><fanqie> 都皓 </fanqie><lb/> 反 </s>
  <s> 一音 <assonance word=" 島 "> 鳥 </assonance></s>
  <warigaki><s> 和名 <jw id=" 之万 "> <wlb/> 之万 </jw></s></warigaki>
</note>
<note char=" 嶼 " hw="2">
  <s><source ref="# 唐韻 "> 唐韻 </source> 云 <re hw="2"> 嶼 </re></s>
```

```

<s><fanqie> 徐呂 </fanqie> 反 </s>
<s><tone><lb/> 上声之重 </tone></s>
<s> 与 <assonance word=" 嶼 " > 序 </assonance> 同 </s>
<s><interpretation> 海中洲也 </interpretation></s>
<warigaki><s> 和名 <jw ref="# 之万 " ><wlb/> 同上 </jw></s></warigaki>
</note>
</entry>

```

注文内に、見出し語そのものについての情報がない／少ないものの例としては、見出しの熟語を単字単位に分割して示すもののみならず、次例「露」(図7)のように、見出し語を含んで熟語形を生産するようなものもある。<note>の属性が「hw="add"」となっているのは、前述の通り、注記対象が見出し語を含む熟語形になっていることを示す。またこれらの(見出し語と完全一致しない)熟語についても、それを対象とする記述が含まれる<note>の内では<re>を用いて再掲を示すこととしている。

さて、「露」の項目の見出しにある「露」という語自体への注記は、実は「音路」と「和名豆由」のみであり、それ以外の部分は「白露」「寒露」「甘露」についての説明である。構造化すると次のとおり、一目瞭然である。

○「露」の例(原文: 露 三礼義宗云白露八月節寒露九月節音路

白虎通云甘露美露也降則物無不美盛矣和名豆由)

```

<entry id=" w0101.003014.006107">
<headword> 露 </headword>
<note char=" 白露 " hw="add">
<s><source> 三礼義宗 </source> 云 <re hw="add"> 白露 </re><interpretation>
  八月節 </interpretation></s>
</note>
<note char=" 寒露 " hw="add">
<s><re hw="add"> 寒露 </re><interpretation> 九 <lb/> 月節 </interpretation></s>
</note>
<note char=" 露 " hw="all">
<s> 音 <assonance word=" 露 " > 路 </assonance></s>
</note>
<note char=" 甘露 " hw="add">
<s><source ref="# 白虎通 " > 白虎通 </source> 云 <re hw="add"> 甘露 </re>
<interpretation>
<associated_word> 美露 </associated_word> 也 </interpretation></s>

```

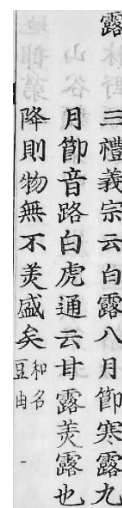


図7「露」


```

<s><interpretation><lb/> 降則物無不美盛矣 </interpretation></s>
</note>
<note char="露" hw="all">
  <warigaki><s> 和名 <jw><wlb/> 豆由 </jw></s></warigaki>
</note>
</entry>

```

以上のことから、本稿冒頭に掲げた「魍魅」（図1，巻二所収）の語についても、以下のように記述できよう。<note>の属性として「hw="other"」とすることにより、見出し語とは別の形を持った語（ここでは「木魅山鬼」）を主題とした注文を抽出することができるようになる。

○「魍魅」の例（原文：魍魅 山海経云魍魅和名須太万鬼類也野王云魍魅老物精也
文選蕪城賦云木魅山鬼和名古太万）

```

<entry id="w0205.017006.005101">
<headword> 魍魅 </headword>
<note char="魍魅" hw="all">
  <s><source ref="# 山海経"> 山海経 </source> 云 <re hw="all"> 魍魅 </re></s>
  <warigaki><s> 和名 <jw> 須 <wlb/> 太万 </jw></s></warigaki>
  <s><interpretation> 鬼類也 </interpretation></s>
  <s><source ref="# 野王"> 野 <lb/> 王 </source> 云 <re hw="all"> 魍魅 </re><interpretation>
    老物精也 </interpretation></s>
</note>
<note char="木魅山鬼" hw="other">
  <s><source ref="# 文選蕪城賦"> 文選蕪城賦 </source><lb/> 云 <re hw="other"> 木魅山鬼
    </re></s>
  <warigaki><s> 和名 <jw> 古 <wlb/> 太万 </jw></s></warigaki>
</note>
</entry>

```

4. 構造化によって可能となること

このようにして『和名類聚抄』巻一の注文全体に亘り、各構成要素を定義付けることによって、従来、紙媒体の索引でも検索可能であった出典名や見出し語、和訓のみならず、注文中に潜在する見出し語相当の語（「関連語」）やそれに対する注、見出しや本文に出現する単字に付された音注なども、格段に抽出しやすくなった（下例）。また、「○○という注記や内容を後続文に持つ見出し語／出典名」という逆引き索引を作成することや、頁移りを設定することで画像とのリンクも可能となろう。

例1 卷一のうち類音注を持つ延べ字数

→ <assonance> を抽出 → 延べ 77 字 59 項目 (巻一)

「禹」(「雨」の項), 「源」(「原」), 「宅」(「沢」), 「悪」(「壘」) など

例2 巻一のうち見出し語を含み新しく熟語化した延べ語数

→ <note char="○○" hw="add"> を抽出 ※○○が複数項目の場合もあり

→ 延べ 10 語 6 項目 (巻一)

「甲賀杣」「田上杣」(「杣」の項), 「磐石」(「磐」), 「石鍾乳」(「鍾乳」) など

一方で、各出典による引用部分がそれぞれの範囲であるかという点については、現状のタグ設計では特定不能である。これは前述のように、出典にあたる書物が現存しないために確認できないというような場合もしばしばあるためであり、完全に解決することは困難である。

5. 古辞書の構造化についての展望

本稿では、『和名類聚抄』巻一について、構造化の検討対象としてきたが、筆者らは現在、20 巻本『和名類聚抄』の全文テキストの作成を行っている。これを機械可読のタグ付き全文テキストとすることが次なる目標となる。全体を見渡したところ、概ね、巻一の構造は全体に敷衍できる内容であるものの、職官部・国郡部などは、よりシンプルでありながら、他の部とは異なる構造を持つものである。『和名類聚抄』全体に対する構造化を行うためには更に詳しい規定が必要となるであろう。またタグ情報を利用して抽出した注記(和名等)は、『和名類聚抄』諸本をはじめとして、新旧の他文献(辞書等)と連携する際にも有効となることが期待される。そのためには、万葉仮名に、ひらがなやカタカナによる音情報を付与しておくことなども肝要なプロセスとなろう。また一般の利用に供するためには、注文の書き下しの作成・提供なども求められるであろう。今後、異なる構造を持つ漢和辞書・国語辞書や古活字版『和名類聚抄』にはない要素(声点、合点や写本上の朱墨など)についての記述方法も、検討していく予定である。更には前述のように、TEI への対応についても引き続き考えることとしたい。

参考文献

- 岩淵匡 (1971) 「倭名類聚抄における和訓の方法について—漢字一字表記語の場合—」『早稲田大学教育学部 学術研究 国語・国文学編』20: 53-66.
- 江口泰生 (1985) 「『和名類聚抄』の「俗」音注」『国語学』141: 1-17.
- 王一凡・永崎研宣 (2018) 「東アジア文献への TEI の適用をめぐる」『研究報告人文科学とコンピュータ(CH)』118(4): 1-4.
- 岡田一祐 (2020) 「日本平安期古辞書の符号化モデル: TEI をもとにした符号化」『デジタル・ヒューマニティーズ』2: 26-54. https://doi.org/10.24576/jadh.2.0_26
- 近藤尚子 (1984) 「『倭名類聚抄』の註文の考察」杉本つとむ(編)『和名抄の新研究』9-33. 東京: 桜楓社.
- 築島裕 (1963) 「和名類聚抄の和訓について」『訓点語と訓点資料』25: 28-60.
- 新野直哉 (1986) 「『和名類聚抄』の「俗云」の性格—「A 俗云 B」の場合—」『文芸研究』112: 53-68.

- 新野直哉 (1989) 「『和名類聚抄』の「一云」―「語形＋一云＋語形」の場合について」『宮崎大学教育学部紀要 人文科学』 66: 43–55.
- 宮澤俊雅 (1986) 「『和名類聚抄』の和訓について 続紹」築島裕博士還暦記念会 (編) 『築島裕博士還暦記念国語学論集』 272–286. 東京: 明治書院.
- 宮澤俊雅 (1994) 「『倭名類聚抄』の「此間」について」『国語国文研究』 95: 1–16.
- 望月郁子 (1989) 「『和名類聚抄』における冠《和名》の役割」『静大国文』 34: 48–61.
- 李安九 (2010) 「『倭名類聚抄』の本文に関する一考察―複数の出典が挙げられる場合―」『松山大学 言語文化研究』 29(2): 45–66. <http://id.nii.ac.jp/1249/00002204/>

関連 Web サイト

- 国立国語研究所『日本語史研究用テキストデータ集』 <https://textdb01.ninjal.ac.jp/dataset/kwrs/> (2020 年 12 月 10 日確認)
- 国立国会図書館デジタルコレクション『倭名類聚抄 20 巻』 <https://dl.ndl.go.jp/info:ndljp/pid/2606770> (2020 年 12 月 10 日確認)

Scheme for a Structured Description of an Old Japanese Dictionary: The Case of *Wamyō-Ruijushō*

FUJIMOTO Akari^aHAN Yi^bTAKADA Tomokazu^c

^aKyoto Prefectural University / Language Change Division, Research Department, NINJAL [–2018.03] / Project Collaborator, NINJAL

^bGraduate Student, Pukyong National University

^cLanguage Change Division, Research Department, NINJAL

Abstract

There were many kinds of dictionaries in ancient Japan. Therefore, without highly specialized knowledge, it is quite difficult for users to interpret their contents. Furthermore, users have been compelled to seek the necessary information based on visual observation of the dictionaries. To effectively find the required information within a complex description, a strict definition of the contents is required.

This paper thus proposes a model for such a structured description and discloses the tagged data by using *Wamyō-Ruijushō*, a representative dictionary of the Heian Era in Japan.

Keywords: Old Japanese Dictionary, structured description, *Wamyō-Ruijushō*