HOKKAIDO UNIVERSITY

Title	平安時代漢字字書総合データベースの構築
Author(s)	池田, 証壽
Citation	北海道大学文学研究科紀要, 142, 79(左)-90(左)
Issue Date	2014-03-25
Doc URL	http://hdl.handle.net/2115/54954
Type	bulletin (article)
File Information	05_IKEDA.pdf

Instructions for use

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

北海道大学文学研究科紀要　142　（2014 年 3 月）抜刷

平安時代漢字字書総合データベースの構築

池　田　証　壽

北大文学研究科紀要 142 (2014)

平安時代漢字字書総合データベースの構築

池 田 証 壽

1 データベース構築の目的

平安時代漢字字書総合データベースは，漢字字体史研究と漢字字書編纂史研究とに資することを目的とする。

本稿の筆者は，これまで『篆隷万象名義』と『類聚名義抄』(原本系・改編本系）の掲出字データベースを構築してきたが，新たに『新撰字鏡』の掲出字データベースを構築し，それらを関連付けた平安時代漢字字書総合データベースを構築する計画である。以下，その概要を述べる。

2 漢字字書総合データベースの構成とデータベース構築の方法

2.1 漢字字書総合データベースの構成

漢字字書総合データベースは,次に掲げる中国・日本の漢字字書データベースから構成される。

○中国

『玉篇』(梁・顧野王撰，543 年成，現存 2,087 字）(原本玉篇）

『大広益会玉篇』(宋・陳彭年等撰，1013 年成，約 22,800 字）(宋本玉篇）

『龍龕手鏡』(遼・行均撰，997 年成，約 26,000 字，高麗版）

○日本

『篆隷万象名義』(空海撰， 9 世紀初，約 16,000 字）

『新撰字鏡』（天治本，昌住撰，10 世紀初，約 21,000 字）

『類聚名義抄』（図書寮本，11 世紀初，約 3,600 項目）

『類聚名義抄』（観智院本，12 世紀後半，約 32,000 項目）

平安時代漢字字書総合データベースは，その編纂材料となった中国の漢字字書のデータベースを利用して構築する。

2.2　データベース構築の方法

漢字字書総合データベースは，第一段階として掲出字画像データベースを構築，第二段階として全文テキストデータベースを構築することで進める。

掲出字画像データベースの構築では，掲出字画像切り出しとユニコード符号化との二つの作業を行う。掲出字画像切り出しでは，掲出字の字形を一字ごとに画像情報として切り出して，所在情報に対応した画像ファイル名を付与する。ユニコード符号化では，所在情報に対応させて，ユニコードの CJK 統合漢字 20,941 字，拡張漢字 A 6,582 字，拡張漢字 B 42,111 字の範囲で掲出字のテキスト入力を行う。

2.3　掲出字画像データベース構築の進捗状況

日本の漢字字書は，掲出字画像切り出しがほぼ終了し，ユニコード符号化を進めている。ユニコード符号化は，『篆隷万象名義』，図書寮本『類聚名義抄』，『新撰字鏡』がほぼ終了している。観智院本『類聚名義抄』は，難字が多く難航している。

中国の漢字字書のうち，原本『玉篇』は東方文化学院の複製本，『大広益会玉篇』は宮内庁所蔵南宋版により掲出字画像切り出しを行い，ユニコード符号化をほぼ終了。『龍龕手鏡』は高麗版（京城帝国大学刊複製）の巻 3 と巻 4 について，掲出字画像切り出しを行い，ユニコード符号化を進めている（巻 1 と巻 2 は未着手）。

2.4　全文テキストデータベース構築の進捗状況

漢字字書の全文テキストデータベースは，掲出字テキストデータベースと

注文テキストデータベースより構成される。掲出字テキストデータベースは，掲出字画像データベースに組み込まれているので，注文テキストデータベース構築が課題となる。

現在，漢字字書のいくつかの部首を試験的に入力して注文テキストデータベース構築の方法を検討している。体例（本文の体裁）について，中国の漢字字書と日本の漢字字書とを比べると，前者は単純，後者は複雑である。

中国の漢字字書の掲出字は単字が原則であり，注文は，音注，義注，字体注の三要素よりなる。音注は，ほとんどが反切であり，まま類音注を交える。義注（意義注）は，経典本文やその訓注である。字体注は「正」「俗」「亦」等を注記する。義注と字体注とをあわせて釈義と呼ぶ。また，掲出字と注文とをあわせて項目と呼ぶ。

日本の平安時代漢字字書は，単字を大字で掲出するのが原則だが，２字以上の熟語掲出も散見し，稀に脱字，注文への繰込みがある。注文は，音注，義注，字体注に加えて，和訓が記載される。掲出字を重複して掲出すること，注文に「在某部」のように注記することも多い。『類聚名義抄』では，声点，片仮名傍訓，ヲコト点が施され，より複雑な様相を呈する。

3　平安時代漢字字書の本文研究と出典研究

3.1 『篆隷万象名義』の本文研究と出典研究

『篆隷万象名義』は平安時代漢字字書に入れているが，実態は原本『玉篇』の一異本である。原本『玉篇』と宋本『玉篇』は，胡吉宣（1989）が本文研究・出典研究を行って有益である。『篆隷万象名義』は，反切に関して上田正（1986）が精緻であったが，近時，呂浩（2007）が出て，本文研究を大きく前進させた。

3.2 『新撰字鏡』と『類聚名義抄』の出典研究

『新撰字鏡』は，『玉篇』，『切韻』，玄応『一切経音義』の三出典を主要な材料とする。これら三大出典の受容に関わる問題は，貞苅伊徳（1959，1960，

1961）が基本的な部分を解明している。その後，『玉篇』引用部分は井野口孝（1978）と上田正（1986），『切韻』引用部分は上田正（1981，1984），『一切経音義』引用部分は池田証寿（1982，1984）がある。和訓は貞苅伊徳（1983a）が「臨時雑要字」と漢語抄類との関連を解明したが，大槻信（2002）は部首分類部分の和訓と漢語抄類との関連を追及する。

図書寮本『類聚名義抄』は出典を明記する字書であり，吉田金彦（1954a，1954b，1955，2013）をはじめとして，多数の研究が出て，進展が著しい。

観智院本『類聚名義抄』は，原撰本にあった出典名（書名・人名）の表示を削除しているので，出典箇所の特定には，引用原典との照合による他ない。和訓は，築島裕（1950）をはじめ多数の研究があり，字音は，呉音について法華経と大般若経の読誦音との関連，漢音について『切韻』系韻書との関連が注意される。掲出字の採録・増補は，貞苅伊徳（1983b）により，『玉篇』字順群の存在が指摘されている。また，掲出字に難字・僻字が多く，それらは『龍龕手鏡』に類似する点が吉田金彦（1958）により指摘されていたが，田村夏紀（2000）はこれを精査している。

3.3 『新撰字鏡』の本文研究

『新撰字鏡』は，和訓を含む項目について，つとに山田孝雄（1916）があり，京都大学文学部国語学国文学研究室（1958）はそれを更に深化させている。掲出字は，長嶋豊太郎（1958-1959）が貴重であり，反切は上田正（1984，1986）の考証が参考になる。和訓と訓詁は内田賢徳（2005）の成果があった。

3.4 『類聚名義抄』の本文研究

『類聚名義抄』は，前述のように，字音と和訓を中心に研究が進展しているが，本文の翻字とその注釈としては，蓮成院本に関する田島毓堂（1990）の試みがある。池田（2005）は図書寮本水部全文の解読を試みたものである。

4　多漢字文献のデータベース化と公開

4.1　利便性・有用性の高いオンラインデータベース

近年は，ユニコードの大規模漢字集合フォントによる多漢字文献テキスト化の推進・公開が著しい。漢字字書も例外ではない。

『説文解字』，『広韻』，『康熙字典』は複数の公開がある。次が参考になる。

中國哲學書電子化計劃（Chinese Text Project）http://ctext.org

漢字データベースプロジェクト http://kanji-database.sourceforge.net

玄応『一切経音義』は，徐時儀校注本が利用できる。

http://tls.uni-hd.de/xuanying.html

ユニコードで使用可能な漢字を検索するには CHISE IDS 漢字検索が便利である。

http://www.chise.org/ids-find

4.2　『玉篇』系漢字字書データベース化の現状

『玉篇』と『篆隷万象名義』は，華東師範大学中国文字研究与応用中心（王平 2005），『篆隷万象名義』は高野山大学密教文化研究所（大柴清圓 2008，2009，2011）で進行中の模様である。

本稿の筆者は『篆隷万象名義』の掲出字のユニコード対応版を公開している。

http://www.hucc.hokudai.ac.jp/～o16404/shikeda/kojisho.html

5　平安時代漢字字書全文テキストデータベース構築の方法

5.1　総合データベースによる全文テキストデータベース構築

平安時代漢字字書総合データベースの構築は，複数の漢字字書を同時にデータベース構築することで行う。この方法は無理があると批判されるかもしれない。

しかし,『篆隷万象名義』の本文を解読して全文テキスト化しようとすると，原本『玉篇』残巻･逸文と『大広益会玉篇』，さらに『広韻』との照合が必要である。『新撰字鏡』ではこれらに加えて玄応『一切経音義』との照合が必要となる。

図書寮本『類聚名義抄』は出典注記の漢字字書であり，出典文献との比較対照は容易であるが，観智院本『類聚名義抄』は，出典註記が省略されており，本文を解読するために，図書寮本『類聚名義抄』,『玉篇』,『広韻』,『一切経音義』との照合が必要である。また，観智院本には,『龍龕手鏡』か，それに類似する字書の影響があるから,『龍龕手鏡』との照合も必要となる。そこで『龍龕手鏡』データベースの構築も視野に入れている。

漢字字書の原文の正確な解読を効率的かつ正確に行うため，関連する漢字字書を総合したデータベース構築を行うものである。

5.2　テキストデータベース構築の実際

電子テキスト化に際して，原文本文の翻字と翻字本文の符号化とを区別する。

- 原文本文の漢字の翻字は，康熙字典体に改めることを原則とする。符号化の文字集合はユニコード基本多言語面の CJK 統合漢字 20,941 字と拡張 A 6,582 字を基本とし，追加多言語面の CJK 統合漢字拡張 B 42,111 字も用いる。
- 原文において字体・字形の差を特に問題とする場合は，可能な限り区別して翻字・符号化する（例「来來　二同」「為爲　二形作」)。
- ユニコードで符号化できない場合は，IDS（Ideographic　Description Sequence）による方法，すなわち漢字を組み合わせて別の漢字を表現する方法を用いる。
- IDS で表現できない場合と表現できてもあまりに複雑になる場合は，対応する康熙字典体にギリシア文字（β γ等）を付し区別する。対応する康熙字典体がない場合は■又は〓とする。
- 原文本文の解読・解釈を示すために，句読点を付す。ただし『類聚名義

図　平安時代漢字字書総合データベース（イメージ）

抄』を除く。

総合データベースは，現時点では File Maker Pro 12 を用いて構築している。掲出字画像は個別ファイルで維持し，掲出字と注文のテキストデータはエクセル形式や CSV 形式に変換が容易である。図に入力画面のイメージを示す。

6　漢字字体史研究と漢字字書編纂史研究

6.1　言語史研究への貢献

冒頭に述べたように，平安時代漢字字書総合データベースは，漢字字体史研究と漢字字書編纂史研究とに資することを目的とする。この点について略述して，和訓，字音を含めた言語史研究への貢献の可能性を述べる。

6.2　漢字字体史研究

石塚晴通（1999）によれば，中国では初唐（618～ 8 世紀初）に楷書の字体の基準が明確化し（初唐標準），その後，晩唐の開成石経（837 年）に至って

基準が大きく転換し，宋版の実践で普及定着する。一方，日本では奈良平安時代にかけて初唐標準が普及・定着し，開成標準の影響は軽微であり，江戸時代の出版文化に至って開成標準の字体の影響が顕著となる。

梁・顧野王『玉篇』は 543 年の成立，隋・陸法言『切韻』は 601 年の成立，唐・玄応『一切経音義』は 661 年頃の成立とされる。『玉篇』と『切韻』は字体の基準が明確化する前の段階，玄応『一切経音義』は字体の基準が初唐標準として明確化する段階の撰述である。『玉篇』撰述の約 60 年後に陸法言『切韻』が撰述，さらにその約 60 年後に玄応『一切経音義』が撰述される。これらに所用の字体には撰述当時の基準をそのまま残す部分と，字体の基準の変遷を反映する部分とがあると想像される。

これに対して，『新撰字鏡』は昌泰年間（898〜901 年）頃，『類聚名義抄』は原撰本が 1081 年以後 11 世紀初，改編本が 12 世紀後半頃の成立である。つまり『新撰字鏡』の成立年代はおよそ晩唐に重なり，『類聚名義抄』の成立年代は宋版が日本に輸入・利用される時期であって，開成標準がどの程度まで影響を及ぼしているかが問題となる。このような文献の成立・書写年代と所用字体との関係の問題は，池田証寿（2012）で「寂」を例に論じたところである。

鈴木慎吾（2005）は，王仁昫『切韻』の異体字注記の包括的検討から陸法言原本の再構を目指す研究であるが，『切韻』諸本の増補と成立・書写年代の観点を視野に入れている点で注目すべき成果である。たとえば，王仁昫『刊謬補闕切韻』（706 年頃）諸本のうち「切三」は「竒」を掲出字，「王二」「王三」は「奇」を掲出字とし，「王三」はさらに「通俗作竒」と注記する。また古い写本の P.3696 (2)は「竒」を掲出字としており，陸法言の原本は「竒」であった可能性が高いという。

6.3　漢字字書編纂史研究

従来の研究では，先行文献を引用した箇所を特定するために，掲出字の配列順序（字順）や注文内容の一致または類似を探し出すことが行われ，貞苅伊徳(1998)に見るように，大きな成果を挙げた。平安時代漢字字書総合デー

タベースはその作業を簡単に行うことができる。更に進めて，先行文献をどの程度まで採用したのか，つまり不採用の掲出字や注文がどの程度あるのかを調査することも容易になる。

たとえば『新撰字鏡』の『玉篇』引用部分には約 5,500 字の掲出字があるが，これを『玉篇』30 巻のどの巻から採用したかを集計してみると次の表のようになる。

この表によると，巻 15 の引用率が極端に低くなっている。巻 15 の総掲出字数 539 字中，『新撰字鏡』に引用が確実なのは 17 字，比率にして 3.2 パーセントに過ぎない。引用の内訳を見ると鹵部 17 字中 16 字，臥部 5 字中 1 字，計 17 字である。所収字数の多い部では，禾部 154 字，米部 83 字，网部 60 字があり，『新撰字鏡』では同じ部があるにも関わらず『玉篇』引用部分を欠いている。

池田証寿（1998）では貞苅伊徳の遺稿「新撰字鏡引用『玉篇』萬象名義・

表　『玉篇』30 巻から見た『新撰字鏡』の『玉篇』引用

巻	A引用数	B総数	引用率 A/B	巻 16	217	387	56.1%
巻 1	112	373	30.0%	巻 17	88	410	21.5%
巻 2	163	600	27.2%	巻 18	291	794	36.6%
巻 3	246	782	31.5%	巻 19	150	681	22.0%
巻 4	208	625	33.3%	巻 20	130	565	23.0%
巻 5	214	532	40.2%	巻 21	211	515	41.0%
巻 6	211	575	36.7%	巻 22	219	629	34.8%
巻 7	254	690	36.8%	巻 23	261	722	36.1%
巻 8	149	449	33.2%	巻 24	194	633	30.6%
巻 9	230	801	28.7%	巻 25	235	609	38.6%
巻 10	164	496	33.1%	巻 26	170	433	39.3%
巻 11	164	654	25.1%	巻 27	124	423	29.3%
巻 12	317	687	46.1%	巻 28	244	436	56.0%
巻 13	311	760	40.9%	巻 29	45	284	15.8%
巻 14	186	451	41.2%	巻 30	60	308	19.5%
巻 15	17	548	3.1%	合計	5,585	16,852	33.1%

宋本玉篇対照表」(400 字詰原稿用紙約 300 枚)を紹介し,「玉篇巻十五における引用の欠如」の指摘に言及した。「引用の欠如」は,正確には,鹵部と臥部を除くという条件付きであることが分かる。

6.4 和訓と字音

平安時代漢字字書総合データベースを言語史研究に活用するには,和訓と字音の研究に資する設計が必要である。

字音は,『広韻』情報の取り込みが課題となるが,これは文字同定・本文研究に必要な基礎作業である。

和訓は,義注(訓詁)との関係付けがやはり文字同定・本文研究に必要な基礎作業である。将来的には,築島裕(2007-2009)『訓点語彙集成』との関連付けが課題となろう。

参考文献

池田証寿(1982)「玄応音義と新撰字鏡」,『国語学』130,pp.1-18
池田証寿(1984)「新撰字鏡玄応引用部分の字順について」,『国語国文研究』71,pp.40-58
池田証寿(1998)「貞苅伊徳著「新撰字鏡の研究」解説」,『新撰字鏡の研究』汲古書院,pp.323-341
池田証寿(2005)「図書寮本類聚名義抄の翻字本文及び注解の作成に関する基礎的研究」北海道大学,平成 14 年度-平成 16 年度科学研究費補助金基盤研究(C)(2)研究成果報告書
池田証寿(2012)「漢字字体の実用例と字書記述 ──「寂」の場合」,石塚晴通編『漢字字体史研究』,勉誠出版,pp.259-282
石塚晴通(1999)「漢字字体の日本的標準」,『国語と国文学』76-6,東京大学国語国文学会,pp.88-96
石塚晴通(2012)『漢字字体史研究』勉誠出版(石塚 1999,池田 2012 を収録)
井野口孝(1978)「新撰字鏡「玉篇群」の反切用字」,『文学史研究』17-18,pp.49-62
上田正(1973)『切韻残巻諸本補正』,東京大学東洋文化研究所附属東洋学文献センター
上田正(1981)「新撰字鏡の切韻部分について」,『国語学』130,pp.1-18
上田正(1984)『切韻逸文の研究』汲古書院
上田正(1986)『玉篇反切総覧』神戸:上田正
内田賢徳(2005)『上代日本語表現と訓詁』塙書房

大柴清圓(2008)「『篆隷萬象名義』における俗字の研究(1)　後漢の隷変字から魏晋の草書の楷書化まで」,『高野山大学密教文化研究所紀要』21，pp.140-89
大柴清圓(2009)「『篆隷萬象名義』における俗字の研究(2)　魏晋から隋唐までの階書の俗字」,『高野山大学密教文化研究所紀要』22，pp.90-33
大柴清圓(2011)「『篆隷萬象名義』における俗字の研究(3)　付録・『篆隷萬象名義』俗字表」,『高野山大学密教文化研究所紀要』24，pp.124-79
大槻信(2002)「古辞書と和訓 ── 新撰字鏡〈臨時雑要字〉──」『訓点語と訓点資料』108，pp.16-34
京都大学文学部国語学国文学研究室（1958)『新撰字鏡国語索引』臨川書店
貞苅伊徳(1959)「新撰字鏡の解剖〔要旨〕── その出典を尋ねて ──」,『訓点語と訓点資料』12，pp.53-76
貞苅伊徳(1960)「新撰字鏡の解剖〔要旨〕付表（上)」,『訓点語と訓点資料』14，pp.63-82
貞苅伊徳(1961)「新撰字鏡の解剖〔要旨〕付表（下)」,『訓点語と訓点資料』15，pp.1-28
貞苅伊徳(1983a)「「新撰字鏡」〈臨時雑要字〉と「漢語抄」」,『国語と国文学』60-1，pp.44-62
貞苅伊徳(1983b)「観智院本類聚名義抄の形成に関する考察　その１　字順をめぐる問題」，第 48 回訓点語学会研究発表会発表資料
貞苅伊徳(1998)『新撰字鏡の研究』汲古書院（貞苅 1960，1961，1983a，1983b を収録）
鈴木慎吾(2005)「王仁昫切韻の異体字注記について」,『開篇』24，pp.14-36
田島毓堂(1990)「類聚名義抄の注釈的研究 ── 電算機利用による ──」名古屋大学，平成元年度科学研究費補助金（総合研究Ａ）研究成果報告書
田村夏紀(2000)「観智院本『類聚名義抄』と『龍龕手鏡』の漢字項目の類似性」,『訓点語と訓点資料』105，pp.40-52
築島裕(1950)「類聚名義抄の倭訓の源流について」,『国語と国文学』27-7，pp.42-52
築島裕(2007-2009)『訓点語彙集成』汲古書院
長島豐太郎(1958-1959)『古字書索引（上・下)』日本古典全集刊行会
吉田金彦(1954 a)「図書寮本類聚名義抄出典攷（上)」,『訓点語と訓点資料』２，pp.61-80
吉田金彦(1954 b)「図書寮本類聚名義抄出典攷(中)」,『訓点語と訓点資料』3，pp.203-236
吉田金彦(1955)「図書寮本類聚名義抄出典攷（下一)」,『訓点語と訓点資料』５，pp.73-98
吉田金彦(1958)「観智院本類聚名義抄の参照文献」,『藝林』9-3
吉田金彦(2013)『古辞書と国語』臨川書店（吉田 1954a，1954b，1955，1958 を収録）
山田孝雄(1916)『新撰字鏡攷異并索引』六合館（西東書房 1933 年改版）
胡吉宣(1989)『玉篇校釈』上海：上海古籍出版社
呂浩(2007)『篆隷万象名義校釈』上海：學林出版社
王平(2005)「《说文》《玉篇》《万象名义》联合检索系统″的开发：从原本《玉篇》到宋本《玉篇》」,『中国文字研究』第 6 辑，华东师范大学中国文字研究与应用中心编

（付記） 本稿は，第 108 回訓点語学会研究発表会(2013 年 5 月 26 日，京都大学)における口頭発表に基づく。日本学術振興会科学研究費補助金基盤研究（C）「大規模漢字集合フォント対応のデータベース構築による平安時代漢字字書研究」（課題番号 25370506）の成果の一部である。