

## 科学研究費助成事業 研究成果報告書

平成 28 年 5 月 19 日現在

機関番号：10101

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25370506

研究課題名(和文)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

研究課題名(英文)A study of Hanzi dictionaries in early Japan using Unicode CJK characters database

## 研究代表者

池田 証壽 (IKEDA, Shoju)

北海道大学・大学院文学研究科・教授

研究者番号：20176093

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：本研究では、平安時代漢字字書総合データベース編纂委員会を設置し、『篆隸万象名義』、『新撰字鏡』、『類聚名義抄』(図書寮本、観智院本)のデータベース構築を進めた。また、このデータベースには、中国の漢字字書である、原本『玉篇』残巻、『大広益会玉篇』、高麗本『龍龕手鏡』の情報を含めている。漢字字書の掲出字は、コンピュータで入力処理が可能な文字と、原文から一文字ごとに切り出した画像とを対象としてデータベース化を進めた。発音・意味・字形を説明する注の部分は、コンピュータで入力処理が可能な文字を入力した。

研究成果の概要(英文)：The Editorial Committee of HDIC Database Project, was founded for the purpose of compiling an unified database of early Japanese dictionaries, named the Integrated Database of Hanzi Dictionaries in Early Japan (abbreviated as HDIC). The Database includes Tenrei Bansho meigi, Shinsen Jiky and Ruiju Myogi sho. We inputted the headwords and annotation text of these dictionaries, in order to construct an inquiry system. A sub-database of dictionaries, compiled in China, including, Yuanben Yupian, Songben Yupian and Gaoliben Longkan Shoujing, is also being independently constructed by the research group. The headwords are being processed in two ways. One is to input the headwords using the computer-readable characters, and the other is to cut a picture of each head word down from a photocopy of the original books. The annotation text, which refers to the pronunciation, meaning and variants of each headword, is inputted using computer-readable text.

研究分野：日本語学

キーワード：古字書 篆隸万象名義 新撰字鏡 類聚名義抄 玉篇 大広益会玉篇 龍龕手鏡 ユニコード

### 1. 研究開始当初の背景

近年における大規模漢字集合フォントによる多漢字文献のテキストデータベース化の進行は顕著であり、漢字字書研究の基本文献である『説文解字』(後漢・許慎)『広韻』(宋・陳彭年ら)『康熙字典』(清・張玉書ら)は Web 上で自在に検索可能となっているし、平安時代の漢字字書に強い影響を与えた『一切経音義』(唐・玄奘)も校訂テキストが Web 公開されている。これに対して日本の漢字字書のテキストデータベース化は遅れており、研究代表者が『篆隸万象名義』の全掲出字を大規模漢字集合フォントに対応したデータとして 2011 年に公開したのはその数少ない例である。平安時代漢字字書のテキストデータベース化が遅れているのは、第一に大規模漢字集合フォントの普及と利用の遅れ、第二に難字・僻字・異体字の多さによる解読上の困難による遅れ、第三に平安時代漢字字書が依拠する中国側の漢字字書のテキスト化の遅れが理由である。第一の理由は時間の問題であって、現行のコンピュータで標準的に利用可能なのは Unicode の基本漢字(20,902 字)と拡張漢字 A(6,582 字)であるが、これに加えて拡張漢字 B(42,111 字)が普及すれば、問題は解決する。第二と第三の理由は、本質的な問題点であって、迂遠なようであるが、中国側の漢字字書をテキストデータベース化し、それを参照・利用して日本側の漢字字書のテキストデータベース化に進むのが方法として確実である。平安時代漢字字書である『新撰字鏡』と『類聚名義抄』は、『玉篇』『切韻』『一切経音義』の三書を基本的な依拠文献としており、幸い『切韻』系韻書である『広韻』と『一切経音義』のテキストデータベースは一般に利用可能であり、『玉篇』のテキストデータベース化が残されている。本研究では、『玉篇』系字書として、原本『玉篇』残巻と宋本『玉篇』のデータベース化を行い、それに基づき、『篆隸万象名義』『新撰字鏡』『類聚名義抄』の 3 書を対象として大規模漢字集合フォントに対応したデータベースの構築を目指すに至ったものである。

### 2. 研究の目的

平安時代の漢字字書である『篆隸万象名義』(9 世紀初)、『新撰字鏡』(10 世紀初)、『類聚名義抄』(原本系・11 世紀初、改編本系・12 世紀後半)を取り上げて、大規模漢字集合フォント対応のデータベース構築を行って、字書編纂史研究と漢字字体史研究における、これらの漢字字書の位置を明確化することを目的とする。

漢字字書編纂史研究に関しては、『新撰字鏡』と『類聚名義抄』とにおいて、原本『玉篇』(原本『玉篇』残巻・『篆隸万象名義』)及び『大広益会玉篇』はどの程度まで対応するのか、所属部首の統合整理はどの程度までなされているのか、これらの漢字字書の全掲出字の相互リンクを構築することで解明す

る。

次に、漢字字体史研究に関しては、『新撰字鏡』、『類聚名義抄』、原本『玉篇』(原本『玉篇』残巻・『篆隸万象名義』)、『大広益会玉篇』を相互に比較可能な掲出字の字体について、初唐の標準字体、開成石経の標準字体のいずれに一致または近似するかを解明する。

最後に、漢字字書編纂史と漢字字体史から見た中国側漢字字書の影響に関して、10 世紀初に成立の『新撰字鏡』には、中国唐代成立の漢字字書の影響、11 世紀初～12 世紀後半に成立の『類聚名義抄』には中国宋代の漢字字書の影響があることを漢字字書編纂史と漢字字体史の観点から実証的に解明する。

なお、データベースは最終的に公開することとし、「平安時代漢字字書総合データベース編纂委員会」を設置して所蔵者・所蔵機関の許諾を得るための手続きを進めることとした。

### 3. 研究の方法

研究の方法は、対象とする漢字字書の掲出字の画像データベースと、それらの掲出字のテキスト情報(漢字字形・所在・諸橋『大漢和辞典』番号・UCS 番号等)とを連結し、平安時代漢字字書を総合したデータベースを構築し、漢字字書編纂史と漢字字体史との観点から調査と分析を行った。

また、依拠するテキストに関しては、可能な限り原本調査を行い、鮮明な画像を入手するなどして、データベースの精度を高めることとした。

本研究課題の研究方法の特色は、平安時代漢字字書データの国際的共有、漢字字書編纂史の研究開発、漢字字体史の研究開発の 3 点にある。

まず、平安時代漢字字書データの国際的共有に関しては、『篆隸万象名義』については、すでにその UCS 対応版を公開していた(2011 年)が、『玉篇』系字書、『新撰字鏡』、『類聚名義抄』について同種のデータを公開すれば、国内外の漢字字書研究に貢献することが可能となる。

次に漢字字書編纂史の研究開発に関して、従前の漢字字書研究は、先行中国字書を利用・引用した箇所を特定することに着目していたが、本研究ではそれにとどまらず、直接に引用しなかった箇所も含めて検討することによって、総合的計量的に字書編纂の方法を分析することを目指した。

最後に漢字字体史研究における字書記述について、漢字字体の字書記述は、掲出字の字体と「正」「通」「俗」の注記によってなされるが、それぞれの字書が編纂された年代の漢字字体の規範と書写された年代の漢字字体の規範とが複雑に絡み合う。本研究は複数の漢字字書の掲出字画像データベースを連携してそれらの漢字字体を包括的に比較調査する点に大きな特色がある。

#### 4. 研究成果

当初の計画とおり、次の漢字字書について大規模漢字集合フォント対応のデータベースを構築した。を付したものは、掲出字画像データベースと全文本文テキストデータベースの入力が完了したものである。は掲出字画像データベースが完成し、全文本文テキストデータベースを構築中のものである。

原本『玉篇』(梁、顧野王撰、543年成、現存2,086字)

『大広益会玉篇』(宋、陳彭年等撰、1013年成、約2万2千字、宋本玉篇)

『篆隸万象名義』(高山寺本、空海撰、9世紀初、約1万6千字)

『新撰字鏡』(天治本、昌住撰、10世紀初、約2万1千字)

『類聚名義抄』(原本系、図書寮本、11世紀初、約3千6百項目)

『類聚名義抄』(改編本系、観智院本、12世紀後半、約3万2千項目)

これらのデータベースを利用した研究成果は次のとおりである。

雑誌論文 と中文版の において、平安時代漢字字書総合データベースの概要を述べた。雑誌論文 においては、収録する漢字字書に関する諸本・複製本・先行研究・原本調査報告・データ構造を詳細に述べた。

字書編纂史研究に関しては、雑誌論文 において、Unicode の対応状況を解説し、応用研究として『篆隸万象名義』と『新撰字鏡』とを対照して収録範囲を検討し、常用平易な掲出字を不採用とすることが目立つことを明らかにした。学会発表 とそれに基づく図書 では、中国仏典音義の受容方法を通して日本古字書の『新撰字鏡』と『類聚名義抄』の共通点と相違点を論じた。

漢字字体史研究に関しては、学会発表 とにおいて発表したが、近刊の『漢字字体史研究 二』(2016年秋、勉誠出版)に掲載の予定である。『篆隸万象名義』『新撰字鏡』の2書と『類聚名義抄』とにおいては、初唐の標準字体と開成石経の標準字体の受容に大きな相違が認められるという見通しを述べた。

また、近年のオープンアクセスの進行を受けて、前述のように一部をインターネットに公開しているが、雑誌論文 、学会発表

などでこのプロジェクトの内容を広く社会に発信した。

構築したデータベースは、「平安時代漢字字書総合データベース」とし、英語名をIntegrated Database of Hanzi Dictionaries in Early Japan、略称をHDICとして広くインターネットに公開することとした。

このうち、基本資料となる『大広益会玉篇』は全文テキストデータベースを、『篆隸万象名義』は掲出字テキストデータベースをイン

ターネット公開している(<http://hdic.jp>)。公開は、研究代表者が委員長となって「平安時代漢字字書総合データベース編纂委員会」を設置し、所蔵者・所蔵機関からの許諾を踏まえて行うこととした。

「平安時代漢字字書総合データベース」の概要を日本語・英語・中国語で紹介するとともに、関連する研究資源(公開画像や研究論文)へのリンクを設定するなど広く社会に研究成果の発信を行っている。

データベース公開に関しては、特に国内資料の場合、所蔵者の許諾の問題が大きく、現状では『篆隸万象名義』の公開にとどまっている。『新撰字鏡』と『類聚名義抄』(図書寮本)は所蔵者(宮内庁書陵部)の許諾を得ており、公開に向けて点検を進めている。『類聚名義抄』(観智院本)は入力自体が完了しておらず、当面は内部利用に限定して構築を進めざるを得ない。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 5件)

池田 証壽、李 媛、申 雄哲、賈 智、齋木 正直、平安時代漢字字書のリレーションシップ、日本語の研究、査読有、12(2)、2016、2016、pp.68-75

劉 冠偉、李 媛、池田 証壽、平安時代漢字字書総合データベースの拡張と和訓対応、情報処理学会研究報告人文科学とコンピュータ研究会報告、査読無、2015-CH-106(4)、2015、pp.1-8、

<http://ci.nii.ac.jp/naid/11000989954>  
6

池田 証壽、平安時代漢字字書総合データベースの構築、高山寺典籍文書総合調査団研究報告論集、査読無、平成26年度、2015、pp.86-78、

<http://hdl.handle.net/2115/58469>

池田 証壽、平安時代漢字字書総合データベース：現状と課題 2014 夏、漢デジ2014 デジタル翻刻の未来、査読無、京都大学人文科学研究所附属東アジア人文情報学研究センター、2014、pp.3-43、<http://hdl.handle.net/2115/59207>

池田 証壽、平安時代漢字字書総合データベースの構築、北海道大学文学研究科紀要、査読無、142号、2014、pp.79-90、<http://hdl.handle.net/2115/54954>

[学会発表](計 7件)

池田 証壽、平安時代漢字字書総合デー

データベース構築の方法と課題 『類聚名義抄』を中心にして、シンポジウム「字体と漢字情報 HNG 公開十周年記念」、2015年11月21-22日、国立国語研究所（立川）

池田 証壽、データベースを利用した日本語の歴史的研究 古辞書と訓点資料、NINJAL セミナー「データが主導する日本語研究」(主催：国立国語研究所・国立台湾大学日本語文学系・国立台湾大学文学院日本研究中心) 2015年9月15日、台北(台湾)

池田 証壽、佛經音義與日本古字書、第九屆漢文佛典語言學國際學術研討會暨第三屆佛經音義國際學術研討會、2015年8月25-26日、北海道大学(札幌)

池田 証壽、漢字字体史の資料と方法 初唐の宮廷写経と日本の古辞書、第111回訓点語学会研究発表会、2014年11月2日、東京大学(東京都文京区)

池田 証壽、李 媛、申 雄哲、賈 智、斎木 正直、平安時代漢字字書のリレーションシップ、日本語学会2014年度秋季大会、2014年10月19日、北海道大学(札幌)

池田 証壽、平安時代漢字字書総合データベース 現状と課題 2014夏、公開シンポジウム「漢デジ2014：デジタル翻刻の未来」、2014年8月5日、北海道大学(札幌)

池田 証壽、平安時代漢字字書総合データベースの構築、第108回訓点語学会研究発表会、2013年5月26日、京都大学(京都)

〔図書〕(計 3件)

池田 証壽(共著) 現代図書、『大広益会玉篇』データベースの構築と利用 『篆隸万象名義』『新撰字鏡』『大宋重修広韻』との対応 (加藤重広・佐藤知己編『情報科学と言語研究』) 2016、pp.65-84

池田 証壽(共著) 上海辞書出版社、佛經音義與日本古字書(徐時儀・梁虹曉・松江崇編『佛經音義研究：第三屆佛經音義研究國際學術研討會論文集』) 2016、pp.36-42

池田 証壽(共著) 上海辞書出版社、宋版辞書在院政、鎌倉時代寺院社会的流通及影響 以《類聚名義抄》為例(石塚晴通・唐焯編『敦煌學・日本學 續編』) 2013、pp.257-269

〔産業財産権〕

出願状況(計 0件)

名称：  
発明者：  
権利者：  
種類：  
番号：

出願年月日：  
国内外の別：

取得状況(計 0件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
取得年月日：  
国内外の別：

〔その他〕  
ホームページ等  
<http://hdic.jp>

6. 研究組織

(1) 研究代表者

池田 証壽 (IKEDA, Shoju)  
北海道大学・大学院文学研究科・教授  
研究者番号：20176093

(2) 研究分担者

( )

研究者番号：

(3) 連携研究者

( )

研究者番号：