

IDS データと HDIC 原本画像・翻刻テキストとを 利用した古辞書の漢字字体研究について

— 『大広益会玉篇』を中心に—

李媛^{†1}

概要: 本研究は IDS データと平安時代漢字字書総合データベース HDIC に収録された古辞書掲出字原本画像・翻刻テキストを利用して、部品レベルで古辞書の字体状況を調査・考察するものである。近時、IDS を利用する漢字の部品レベルで写本・版本の字体の相違を考察する視点が生まれる。そのアプローチの研究成果として、守岡知彦氏は CHISE の多粒度漢字構造モデルに基づき、漢字字体規範データベース HNG に収録された標準文献の字体を整理した。古辞書は漢字字体の多様なバリエーションを網羅的に収録する。部品レベルで古辞書の漢字字体を整理することは新たな課題となっている。IDS は部品レベルで漢字を記述する方法であり、CHISE では、Unicode 統合漢字の約 7 万字の IDS データが公開されている。本研究はその IDS データを取り込み、HDIC の掲出字原本画像・翻刻テキストに基づいて、部品レベルで古辞書における漢字字体のバリエーションを考察する。具体的には、字体が翻刻する際に用いる康熙字典体に近い、版本である『大広益会玉篇』(宋本『玉篇』ともいう)を取り上げ、その調査の詳細を報告する。

キーワード: IDS, HDIC, 古辞書, 漢字字体, 宋本玉篇

Chinese Characters Glyphs study of Earlier Dictionaries based on IDS Data and HDIC : Songben Yupian

LI YUAN^{†1}

Abstract: This paper presents the results of investigation of Chinese characters glyphs at parts level. Concretely, with the method of using the IDS data and original pictures and inputted text included in HDIC. Recently, there is a new view point to study the glyphs of Chinese Characters at parts level by using the IDS data. As an approach, Tomohiko Morioka had tried to integrated HNG and the CHISE character ontology by using categorizing glyph-images based on Multiple Granularity Hanzi Structure Model. There are a large number of variations of Chinese characters glyph existed in the earlier dictionaries. How to sort out them is a new issue. IDS is a method to describe the characters at parts level. The IDS data around 70,000 characters are made publicly by CHISE. In this paper, Songben Yupian has been taken up for the investigation.

Keywords: IDS, HDIC, Earlier dictionaries, Chinese Characters glyphs, SongbenYupian

1. はじめに

本研究は、『大広益会玉篇』(陳彭年等により、宋・真宗の大中祥符 6 年に成立、以下宋本『玉篇』という)を中心に、IDS データと HDIC[a][1]に収録された古辞書掲出字原本画像・翻刻テキストを利用し、部品レベルで古辞書の字体状況を調査・考察するものである。守岡知彦氏による[2][3]は CHISE の多粒度漢字構造モデルに基づき、漢字字体規範データベース HNG に収録された標準文献の字体を整理した研究成果である。この研究成果に対して 本研究は、漢字部品レベルで、標準文献と性質の異なる古辞書について調査を行うものである。

2. 研究対象

2.1 漢字字体研究から見た古辞書掲出字

古辞書は、数多くの掲出字を収録する。漢字字体研究

において、HNG に収録された標準文献に比べて、次の二つの特徴が指摘できる。

掲出字は古辞書の説明対象としての骨組みであり、少数の重複字以外、ユニークな存在である。一方で、個々の掲出字そのもののバリエーションが僅少であるが、掲出字を網羅的に収録し、異体字も併記するため、漢字字体の多様性を備える。

同じ辞書、あるいは異なる辞書において、異なる掲出字の間に、同一漢字部品を持つものが多く存在する。漢字部品レベルでは、字体の同一性(単一パタン)と多様性(複数パタン)が観察できる。

次に、やや詳しく説明してみる。

について、例えば、宋本『玉篇』では、掲出字として「伴」が上巻二十五丁表の一箇所しか存せず、「伴」の字自体は唯一の存在である。しかしながら 宋本『玉篇』は 22,804

^{†1} 北海道大学大学院文学研究科博士課程
Hokkaido University, Graduate School of Letters

a) Integrated Database of Hanzi Dictionaries in Early Japan, 略称 HDIC.

の掲出字が収録される。「伴」の傍の「半」を含む掲出字は、「半」「畔」「伴」などの21字がある。原本画像では「半」/「半」/「𠂔」との字体があることも確認できる。

は、古辞書において漢字部品レベルで、字体の同一性と多様性とを観察することである。同一性について、同じ辞書、あるいは異なる辞書の間、単一パターンだけ存在する部品がある。これらの部品の字体は、同一性があると言える。多様性について、部品によって、二種類の状況がある。一つは、同じ辞書の中バリエーションがあるものである。例えば、「糸」の字体は、高山寺本『篆隸万象名義』の中、「𦉳」「𦉴」が共存する。もう一つは、一つの辞書では同一パターンであるが、異なる辞書の間バリエーションが観察できる。例えば、「月」の字体は、高山寺本『篆隸万象名義』の中、「月」の単一パターンであるが、天治本『新撰字鏡』の中、「肉」に近い字体である。高山寺本『篆隸万象名義』と天治本『新撰字鏡』とは、「月」の多様性が確認できる。

さらに、HDIC に収録された古辞書は、古写本と古版本とがある。古写本は、字体のバリエーションが豊富であり、同時に誤写や誤脱を含むものである。原本『玉篇』、高山寺本『篆隸万象名義』、天治本『新撰字鏡』、図書寮本『類聚名義抄』、観智院本『類聚名義抄』はいずれも古写本である。古版本は、字体のバリエーションが少なくなり、写本より印刷が鮮明である。掲出字の解読が写本よりしやすい面がある。宋本『玉篇』は古版本である。HDIC の作業は、宋本『玉篇』の掲出字テキストデータベースを作成し、これを土台として古写本の辞書の作業を進める方針をとっている[4]。また、字体が翻刻する際に用いる康熙字典体も宋本『玉篇』に近い。そこで、本研究は古版本である宋本『玉篇』を取り上げ、その漢字字体の実態を考察する。そして、宋本『玉篇』データベースにおける掲出文字符号化の際の包摂の実態を整理する。

2.2 IDS データ

IDS は部品レベルで漢字を記述する方法である。Unicode では12個の漢字構成記述文字 IDC (Ideographic Description Character, U+2FF0~U+2FFB) を提供し、これらの IDC によって漢字部品の位置関係を示すことができる。CHISE では、Unicode 統合漢字の約7万字的 IDS データが公開されている[b]。IDS は、翻刻の際に、符号化できない難字を表現する手段である。一方で、大多数の漢字が IDS データによる記述があるため、IDS データによって、漢字の部品検索ができる。原本画像と併用することによって、部品の原本字体を確認の実現することが可能となっている。

2.3 HDIC 原本画像・翻刻テキスト

HDIC では、原本画像から一文字ごとに切り出した画像を対象としてデータベース化した。掲出字画像データベースは、漢字字体史研究に利用することを目的として作成しはじめたものである[4]。掲出字画像データベースを構築して、掲出字のテキスト入力効率化をはかる。そして、掲出字画像データベースによって、掲出字の字形の細部を観察できる。宋本玉篇データベースについて、HDIC のプロジェクトでは、2014年にその掲出字と注文とを入力済みである。その後、一次点検・校正の上、宋本『玉篇』のテキストを試行公開中[c]であり、筆者は担当者である。今回の調査は、試行公開中の掲出字テキストを利用したものである。宋本玉篇データベースの原本画像は宮内庁書陵部に所蔵する宋版テキストによった。

3. 漢字部品に関連する漢字字体の先行研究

漢字字体上の相異が問題になる漢字部品レベルのものが、今まで関連する各領域の研究がある。人文学では、[5][6][7]は、後漢から魏晉まで、魏晉から隋唐までの書体・字体変化の視点から、多くの実例を集め、『篆隸万象名義』の俗字の分析を行い、所用の俗字を共通する文字ごとに250種類に分類した。情報処理では、JISの包摂規準とUCSのIWDS (IRG Working Document Series) が挙げられる。

守岡知彦氏による[2]はHNGに収録された代表字形データへのIWDS-1[d]を適用した場合の包摂可能性を検討した。本研究はIDSデータを通して、部品レベルで宋本『玉篇』の掲出字の原本字体の同一性と多様性とを考察するものである。そして、宋本玉篇データベースにおける掲出文字符号化の暗黙的な包摂の実態を整理することによって、包摂規準を明確化させようとする。

4. 漢字部品リストの抽出

調査手順は、まずIDSデータによって、漢字を構成する部品を整理する。次に、抽出した漢字の部品を手がかりにして、HDICの掲出字原本画像・翻刻テキストに基づき、宋本玉篇における漢字字体のバリエーションを考察する。

CHISEで公開するUnicode統合漢字・基本多言語面のIDSデータを利用し、Excelで繰り返してマッチング作業を行う。この手順によって、次の534個のさらに分割できない基本漢字部品[e]を得られた。ただし、これらの部品の中では、簡体字による部品も混在し、選別する抽出作業が必要となる。これらの部品を対象として、宋本『玉篇』の掲出字の原本字体の調査を行う。

b) CHISE 漢字構造情報データ <http://kanji.zinbun.kyotou.ac.jp/projects/chise/ids/>
c) <http://hdic.jp>
d) List of UCV (Unifiable Component Variations) of Ideographs

e) なお今回の調査、「音」のような中間部品や符号化されていないCDP外字(台湾中央研究院)などを対象外とした。

1	2	3	4	5	6	7	8	9	10
夕	𠂇	イ	冂	几	リ	卜	𠂇	卜	𠂇
11	12	13	14	15	16	17	18	19	20
彡	氷	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
21	22	23	24	25	26	27	28	29	30
糸	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
31	32	33	34	35	36	37	38	39	40
𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
41	42	43	44	45	46	47	48	49	50
白	办	半	𠂇	𠂇	貝	貝	本	匕	比
51	52	53	54	55	56	57	58	59	60
必	術	畢	采	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
61	62	63	64	65	66	67	68	69	70
不	才	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
71	72	73	74	75	76	77	78	79	80
弗	𠂇	長	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
81	82	83	84	85	86	87	88	89	90
乘	乘	尺	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
91	92	93	94	95	96	97	98	99	100
豕	川	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
101	102	103	104	105	106	107	108	109	110
大	夕	イ	𠂇	丹	刀	、	电	𠂇	𠂇
111	112	113	114	115	116	117	118	119	120
丁	鼎	东	東	兜	斗	豆	鬥	盟	自
121	122	123	124	125	126	127	128	129	130
夕	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
131	132	133	134	135	136	137	138	139	140
飞	非	飛	丰	风	風	𠂇	𠂇	𠂇	𠂇
141	142	143	144	145	146	147	148	149	150
市	弗	𠂇	甫	父	阜	𠂇	甘	干	高
151	152	153	154	155	156	157	158	159	160
戈	革	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
161	162	163	164	165	166	167	168	169	170
骨	瓜	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
171	172	173	174	175	176	177	178	179	180
鬼	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
181	182	183	184	185	186	187	188	189	190
户	黄	黄	火	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
191	192	193	194	195	196	197	198	199	200
𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
201	202	203	204	205	206	207	208	209	210
見	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
211	212	213	214	215	216	217	218	219	220
金	金	井	𠂇	九	久	𠂇	𠂇	𠂇	𠂇
221	222	223	224	225	226	227	228	229	230
巨	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
231	232	233	234	235	236	237	238	239	240

来	來	𠂇	老	乐	了	未	力	立	吏
241	242	243	244	245	246	247	248	249	250
隶	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
251	252	253	254	255	256	257	258	259	260
马	馬	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
261	262	263	264	265	266	267	268	269	270
𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
271	272	273	274	275	276	277	278	279	280
内	内	年	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
281	282	283	284	285	286	287	288	289	290
牛	牛	女	𠂇	皮	片	𠂇	𠂇	𠂇	𠂇
291	292	293	294	295	296	297	298	299	300
平	支	七	齐	齊	千	𠂇	欠	且	青
301	302	303	304	305	306	307	308	309	310
青	丘	求	曲	犬	𠂇	𠂇	𠂇	𠂇	𠂇
311	312	313	314	315	316	317	318	319	320
日	内	肉	入	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
321	322	323	324	325	326	327	328	329	330
𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
331	332	333	334	335	336	337	338	339	340
十	𠂇	石	史	矢	豕	士	氏	𠂇	示
341	342	343	344	345	346	347	348	349	350
世	世	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
351	352	353	354	355	356	357	358	359	360
𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
361	362	363	364	365	366	367	368	369	370
巴	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
371	372	373	374	375	376	377	378	379	380
𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
381	382	383	384	385	386	387	388	389	390
亡	王	网	𠂇	韦	为	為	𠂇	𠂇	未
391	392	393	394	395	396	397	398	399	400
𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
401	402	403	404	405	406	407	408	409	410
戊	夕	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
411	412	413	414	415	416	417	418	419	420
𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
421	422	423	424	425	426	427	428	429	430
𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
431	432	433	434	435	436	437	438	439	440
𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
441	442	443	444	445	446	447	448	449	450
夷	臣	乙	巳	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
451	452	453	454	455	456	457	458	459	460
音	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
461	462	463	464	465	466	467	468	469	470
尤	由	西	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇

471	472	473	474	475	476	477	478	479	480
魚	予	羽	雨	禹	玉	聿	冂	日	甲
481	482	483	484	485	486	487	488	489	490
月	戊	龠	𠂔	戔	先	矢	由	乍	长
491	492	493	494	495	496	497	498	499	500
丈	爪	𠂔	兆	フ	永	之	久	止	𠂔
501	502	503	504	505	506	507	508	509	510
豸	中	重	舟	州	竹	𠂔	𠂔	佳	子
511	512	513	514	515	516	517	518	519	520
自	𠂔	ナ	𠂔	充	者	𠂔	マ	𠂔	𠂔
521	522	523	524	525	526	527	528	529	530
𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
531	532	533	534						
叟	既	直	眾						

HDIC の掲出字原本画像・翻刻テキストによって、次の漢字部品、計 96 個を選別し、漢字字体調査の対象外とする。これ以外 438 個の漢字部品の調査結果を次節で示す。

・簡体字による部品 (27 個)

46	75	84	113	135	142	169	195	198	200
貝	车	齿	东	风	弗	龟	夹	戈	见
202	210	235	251	276	332	364	372	386	424
纟	车	乐	马	鸟	𠂔	肅	头	为	𠂔
425	434	448	469	490					
严	页	义	鱼	长					

・原本字体に含まれない部品 (44 個)

17	33	35	69	71	97	110	127	131	139
正	龠	鬼	𠂔	弗	垂	𠂔	𠂔	𠂔	𠂔
150	212	217	224	243	244	245	262	274	278
高	金	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
279	294	310	343	345	350	352	365	385	398
𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
405	418	422	423	427	456	478	480	488	507
习	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
508	524	526	527						
𠂔	𠂔	𠂔	𠂔						

・原本字体に孤例である部品 (25 個)

1	4	13	56	80	98	167	178	223	231
夕	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
257	301	315	321	322	342	378	379	380	403
𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
406	407	420	484	522					
𠂔	𠂔	𠂔	𠂔	𠂔					

5. 漢字部品から見た字体のバリエーション

対象となる 438 個の部品を調査したところ、原本字体により、字体の同一性を示す単一パタンの部品と多様性を示す複数パタンの字体とに分けられる。次に、実例を示しながら説明する[7]。

5.1 単一パターン

(1) 31

IDS	Entry	IMG
𠂔イ司	伺	
𠂔イ童	僮	
𠂔イ二	仁	
𠂔イ鹿	儻	
𠂔イ丕	伾	

(2) 61

IDS	Entry	IMG
𠂔月リ	別	
𠂔崩リ	蒯	
𠂔井リ	荆	
𠂔肖リ	削	
𠂔奇リ	劓	

一部の原本字体しか示さないが、「イ」と「リ」の字体は、宋本『玉篇』ではすべて異同がない。原本字体のレベルでは、単一パターンであることが確認できる。

単一パターンであることが確認できた漢字部品は次の 237 個がある。宋本『玉篇』において、これらの部品は、字体の同一性が確認できた。

f) 例示する「イ」「リ」「巴」「甘」等の部品をもちいる掲出字の翻字は、[8]に収録された宋本『玉篇』の翻刻テキストと一致する。

3	6	9	10	14	15	16	20	21	22
イ	リ	卜	才	生	彡	王	𠂔	糸	𠂔
25	27	28	29	30	31	32	36	39	45
月	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
50	52	53	55	57	58	59	64	65	67
比	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
68	72	73	81	82	83	85	86	87	88
冊	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
89	91	92	94	100	102	103	104	108	109
虫	豕	川	舛	寸	歹	イ	𠂔	电	𠂔
111	112	114	116	117	120	122	124	129	130
丁	鼎	東	斗	豆	白	𠂔	而	𠂔	方
132	133	134	136	137	145	149	152	154	155
非	飛	丰	風	缶	父	干	革	艮	豕
156	157	160	161	162	166	168	173	180	181
更	工	谷	骨	瓜	𠂔	𠂔	果	戶	戶
184	185	187	188	194	197	201	203	209	211
火	𠂔	及	𠂔	𠂔	甲	見	糸	斤	金
215	218	225	226	229	232	233	234	236	237
九	韭	夬	𠂔	𠂔	來	𠂔	老	了	耒
239	240	246	247	248	249	250	252	254	258
立	吏	龍	鹵	鹿	率	麻	馬	毛	門
260	263	264	265	268	269	270	272	273	277
糸	民	𠂔	𠂔	木	目	乃	内	年	鳥
281	282	283	285	286	287	290	291	298	300
牛	牛	女	皮	片	𠂔	𠂔	平	欠	青
304	305	306	311	312	316	317	318	319	325
曲	犬	隹	日	肉	三	彡	色	山	申
326	327	328	330	331	333	334	336	337	338
申	身	生	尸	十	石	史	豕	士	氏
339	340	341	344	346	347	351	353	356	359
ネ	示	世	事	手	百	疋	黍	東	水
361	362	367	374	375	381	382	384	387	389
巳	𠂔	夂	土	屯	亡	王	口	為	爲
393	394	395	396	397	399	409	411	413	417
我	巫	无	毋	𠂔	五	下	乡	心	熏
419	421	426	428	430	431	432	433	435	436
牙	𠂔	言	央	𠂔	𠂔	也	业	曳	頁
437	439	441	443	444	445	446	447	450	451
一	衣	夷	乙	巳	𠂔	𠂔	𠂔	𠂔	𠂔
453	455	460	466	470	473	475	479	483	489
彳	尹	尤	于	禺	羽	禹	日	𠂔	𠂔
491	496	497	499	500	501	502	503	505	509
丈	承	之	止	𠂔	𠂔	中	重	州	隹
511	512	513	514	515	516	517	519	520	521
自	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
523	525	528	529	530	532	533			

𠂔 宙 具 冗 刃 既 直

5.2 複数パタン

(3) 38 巴

IDS	Entry	IMG
𠂔月巴	肥	
𠂔止巴	𠂔	
𠂔立巴	𠂔	
𠂔戸巴	𠂔	
𠂔木巴	𠂔	

(4) 148 甘

IDS	Entry	IMG
𠂔𠂔甘	𠂔	
𠂔甘覃	𠂔	
𠂔甘兼	𠂔	
𠂔𠂔甘	𠂔	
𠂔干甘	𠂔	

「巴」と「甘」の字体は、宋本『玉篇』ではそれぞれ二つのパタンが観察される。原本字体レベルでは、複数パタンであることが確認できた。

複数パタンであることが確認できた漢字部品は次の 201 個がある。宋本『玉篇』において、これらの部品は、字体の多様性が確認できる。

2	5	7	8	11	12	18	19	23	24
𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
26	37	38	40	41	43	44	47	48	49
𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔

51	54	60	61	62	63	66	70	74	76
必	采	卜	不	才	卅	冊	甬	厂	車
77	78	79	90	93	95	96	99	101	105
少	臣	辰	丑	𠂔	串	垂	束	大	丹
106	107	115	118	119	121	123	125	126	128
刀	、	兜	鬥	盟	夕	儿	耳	二	女
138	140	141	143	144	146	147	148	151	153
缶	、	市	由	甫	阜	丐	甘	戈	鬲
158	159	163	164	165	170	171	172	174	175
弓	井	凵	乖	卅	龜	鬼	冂	禾	黑
176	177	179	182	183	186	189	190	191	192
乎	虍	互	黄	黄	冂	几	己	ㄣ	互
193	196	199	204	205	206	207	208	213	214
无	夾	束	角	子	走	丰	巾	井	冂
216	219	220	221	222	227	228	230	238	241
久	白	白	巨	丿	ㄣ	口	ㄣ	力	隶
242	253	255	256	259	261	266	267	271	275
、	叵	矛	𠂔	米	丐	末	母	内	廿
280	284	288	289	292	293	295	296	297	299
聿	月	水	冂	支	七	齊	千	冂	且
302	303	307	308	309	313	314	320	323	324
丘	求	丹	冉	人	肉	入	夾	上	舌
329	335	348	349	354	355	357	358	360	363
尸	矢	首	几	鼠	丨	爽	水	厶	四
366	368	369	370	371	373	376	377	383	388
肅	少	田	酉	冂	凸	瓦	丸	网	韋
390	391	392	400	401	402	404	408	410	412
未	、	文	兀	戊	夕	西	丁	巳	小
414	415	416	429	438	440	442	449	452	454
辛	玄	穴	羊	、	、	、	、	、	、
457	458	459	461	462	463	464	465	467	468
、	永	用	尤	由	酉	𠂔	又	史	與
471	472	474	476	477	481	482	485	486	487
魚	予	雨	玉	聿	月	戊	戈	先	矢
492	493	494	495	498	504	506	510	518	531
爪	、	兆	、	、	舟	竹	子	マ	叟
534									
粟									

6. おわりに

本研究は、CHISE で公開する Unicode 統合漢字・基本多言語面の IDS データを利用し、さらに分割できない基本漢字部品リストを得られた。簡体字による部品、宋本『玉篇』に含まれない部品、宋本『玉篇』では孤例である部品以外、438 個のものを対象にし、部品検索によって、HDIC に収録された宋本玉篇の原本画像を利用し、宋本『玉篇』の漢字字体の実態を考察した。

438 個の基本部品のうち、237 個の部品に対応する漢字字体は宋本『玉篇』において、単一パターンであり、字体の同一性が確認できた。201 個の部品に対応する漢字字体は宋本『玉篇』ではバリエーションが観察され、字体の多様性が確認できた。

さらに、これらバリエーションが観察される 201 個の部品に対応する宋本『玉篇』の字体を整理し、宋本玉篇データベースにおける掲出文字符号化の包摂の実際を整理した。

今後、中間部品と外字とも取り込み、それによって、宋本玉篇データベースにおける掲出文字符号化の包摂の実際を考察する必要がある。さらに、IWDS と JIS 包摂規準とを取り上げ、これらの包摂規準と宋本玉篇データベースの掲出文字符号化の包摂の実際とを比較する。そして、IWDS と JIS 包摂規準を宋本『玉篇』の掲出文字符号化の作業に適用とする場合、どのように拡張すべきかを検討したい。

参考文献

- [1] 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. 2014, vol. 142, pp. 79-90.
- [2] 守岡知彦. 長安宮廷写経の漢字字体と包摂規準 - HNG と CHISE の統合を通じて -. 東洋学へのコンピュータ利用 第 27 回研究セミナー. 2016, pp. 15-48.
- [3] 守岡知彦. 多粒度漢字構造モデルに基づく字形整理の試み - 漢字字体規範データベースの CHISE への収録を通じて -. 人文科学とコンピュータシンポジウム 2015 論文集. 情報処理学会. 2015, pp. 1-8.
- [4] 池田証壽. 『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. 情報科学と言語研究. 現代図書. 2016, pp. 65-84.
- [5] 大柴清圓. 『篆隸万象名義』における俗字の研究(1) 後漢の隸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. 2008, vol. 21, pp. 49-100.
- [6] 大柴清圓. 『篆隸万象名義』における俗字の研究(2) 魏晋から隋唐までの楷書の俗字. 高野山大学密教文化研究所紀要. 2009, vol. 22, pp. 227-284.
- [7] 大柴清圓. 『篆隸万象名義』における俗字の研究(3) 附録. 『篆隸万象名義』俗字表. 高野山大学密教文化研究所紀要. 2011, vol. 24, pp. 93-138.
- [8] 古佚三書上元本玉篇/韻/小學鉤沉三編. 四川出版集團. 四川辭書出版社. 2013, pp. 1011-1562.