

統合漢字に申請された「殷周金文集成引得」図形文字の調査

鈴木俊哉^{†1}

昨年、7年間の作業を経た CJK 統合漢字拡張 C が ISO/IEC 10646:2003 の Amd. 5 として正式に発表された。拡張 C ははじめて原規格分離を適用せずに整理された漢字集合である。拡張 C への中華人民共和国からの申請は辞典類から収集された漢字が大半である。本発表では、その中で最大の収集元である「殷周金文集成引得」に由来する 1800 字程度の図形集合について、典拠を再調査した結果を報告する。拡張 C の統合作業中に提出された典拠確認資料を見ると、この収集は「殷周金文集成引得」の総画索引から、拡張 B までで符号化済みと思われるものを削って選定したと思われる。しかし、選定された漢字を本文中で確認すると、総画索引には出現するが、釈文では使われていないものも少なくないため、それらを全て申請することの妥当性には疑問がある。また、金石学の分野では「古文字字形表」や「金文編」など過去にいくつもの字書が出版されているが、「殷周金文集成引得」の特徴の一つに、見出し字の大半を宋体（風の図形）にしている点がある。金文字形に対する宋体（風の図形）が一意的に定まる、言い換えれば「殷周金文集成引得」で導入された図形文字が（この字書の外部でも）金文字形に対する識別子として機能するのであれば、他の金石学の字書類と整合する筈である。そこで、代表的な金石学字書である「金文編」と「殷周金文集成引得」の比較を行ない、この結果を報告する。また、これらの結果を踏まえ、「殷周金文集成引得」典拠情報の今後の管理方法について考察したい。

Investigation on Glyphs collected from "Index to Collection of Inscriptions of the Yin-Zhou Period" to submit to CJK Unified Ideographs

SUZUKI TOSHIYA^{†1}

After the long efforts during 7 years, finally ISO/IEC 10646:2008 has included CJK Unified Ideographs Extension C. This is the first Hanzi collection which is standardized after the expire of the source code separation rule which was introduced for existing regional character encodings. There are 2 large groups of the sources: Hanzi for personal names (especially from TCA) and Hanzi for post-kaisu palaeographic documents (especially from ROK).

In Ext. C, PRC submitted 366 glyphs taken from "Index to Collections of the Inscriptions in Yin-Zhou period" (殷周金文集成引得, I2CIYZ). The book is used to lookup Bronze objects (collected in "Collections of the Inscriptions in Yin-Zhou period" (殷周金文集成引得, CIYZ) including a specified Old Hanzi.

Most of 366 glyphs taken from I2CIYZ are suspected to be the glyphs invented only for the specification of Old Hanzi.

In this report, the submission by PRC and that by UTC are investigated for their original glyphs based on Bronze or Seal scripts (篆文, Zhuan Wen), and their original shapes (in references and evidences) and modernized shapes (in proposal and submission documents) are compared. The unification rules of CJK Unified Ideographs (ISO/IEC 10646 Annex S) are under revision process, but the discussion is based on the information interchange by the stabilized shapes of Hanzi current in use. The glyphs to specify Bronze script shape can be synthesized by different granularity. In fact, the identification rule by IRG Old Hanzi group for Oracle Bone script (甲骨文, Jia Gu Wen) is incompatible with ISO/IEC 10646 Annex S.

1. 背景

1.1 ISO/IEC 10646 における CJK 統合漢字

2008年12月に、ISO/IEC 10646:2003¹⁾の5個目の追補であるISO/IEC 10646:2003/Amd. 5:2008²⁾が正式に公開された。この追補には、2003年に公開された日中韓統合漢字拡張B以来、6年ぶりの大きな統合漢字の追加として日中韓統合漢字拡張Cが含まれている。統合漢字拡張Cへの採録作業が開始されたのは2000年5月であるので、約9年間にわたる長い作業を経た成果である。

1.2 拡張B以前と拡張C以降の差異

拡張Bまでは実際に情報交換に用いられているかどうかはともかく、各国の国内規格で符号化を完了したものを典拠に各国が申請し^{*1}、それらに対する統合作業を行なった。既に符号化されているということを重視し、拡張Bでも原規格分離の原則を適用した。

拡張C以降では原規格分離の原則は適用されない。これは、仮に拡張Bまでに含まれていない漢字が存在し、それが既に各国でISO/IEC 10646に先んじて符号化済であるとしても、その符号化作業はISO/IEC 10646の発表以降の作業であるから、原規格分離が必要となるのは

- 各国の国内規格化作業にミスがあって重複符号化した
- ISO/IEC 10646と相互運用できないような包摂基準で規格化した

という背景しか考えられないからである^{*2}。

実際、最初の拡張Cへの当初の申請は以下のようにほぼ全て印刷物を典拠とするもので

^{†1} 〒 739-8511 東広島市鏡山 1-4-2 広島大学大学院総合科学研究科
Faculty of Integrated Arts and Science, Hiroshima Univ., Kagamiyama 1-4-2, Higashi-Hiroshima-shi, 739-8511 Japan

^{*1} ただし中華人民共和国からの申請には符号化されていない、辞典類を典拠とする漢字が含まれていた。

^{*2} これは日中韓統合漢字の標準化の合意で、全用字系で原規格分離を適用しないという合意があるわけではない。

あった³⁾。例外的に、朝鮮民主主義人民共和国は2000年からIRGに参加したため⁴⁾、符号化文字集合KPS 10721-2000のうち拡張Bまでに符号化できていない339字が拡張Cに対して申請された^{5),6)}。

● 印刷物を典拠とするもの

- 漢語大詞典
- 辞海(1999年版)^{*1}
- 辞源(修訂版)
- 現代漢語詞典(修訂版)
- 新華字典(1998年版)^{*2}
- 古代漢語詞典(第1版)^{*3}
- 人民日報
- 高麗大蔵経
- 明文大字典

● 符号化文字・字形集合を典拠とするもの

- KPS 10721-2000

この後、2001年には台湾から18000字以上、韓国から23000字以上⁷⁾の漢字が申請され⁷⁾、整理作業がこれまでに膨大なものとなった⁸⁾。特に、申請された漢字が拡張Bまでに採録された漢字に統合されるかを記憶に頼って判別することは殆ど不可能と判断されたため、漢字入力方式や部首・画数整理によって精査することが合意された⁹⁾。しかし、2004年にはこの精査でも安定化が不十分であると判断され、日本NBからの提案で申請漢字を全てIDS分解して機械的に類似字形を拾い出して照合する作業方式に変更され¹⁰⁾、さらに当初の申請では不要であった用例^{*4}も提出が義務づけられた^{*5}。JIS X 0208:1997やJIS X 0213:2000の制定作業では前提であったものがようやく国際標準の作業にも反映されたとと言える。

しかし、拡張Cの標準化作業が8年の長期にわたり、また事後的に用例が必須とされたため、正式に公布された拡張C漢字の典拠は当初の典拠とはかなり異なるものとなっている。たとえば日本では当初は文字鏡研究会によって選定された1000文字程度の漢字を申請して

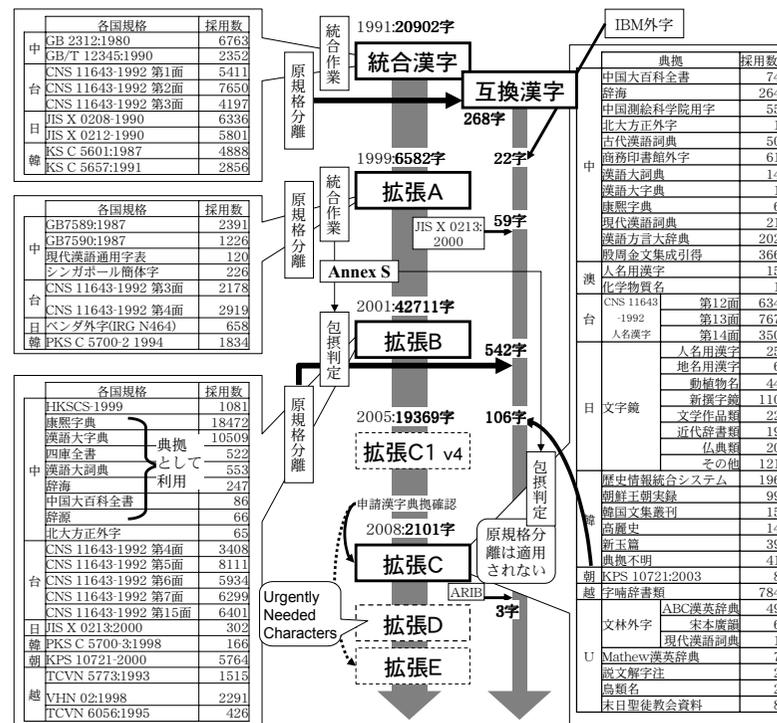


図1 日中韓統合漢字の拡張履歴

いたが、図2に示すように後に電子政府用の人名・地名漢字が追加された。また、典拠用例確認は文字鏡とは独立に行なわれたため、字形が文字鏡漢字から変更されたものも多い。

日中韓統合漢字の拡張Cまでの履歴を図1に示した。拡張Cに採録された漢字のうち、「殷周金文集成引得」を典拠とするものは366字あるが、これの用例確認はAmd. 5の投票に間に合わなかったため、図3に示すように、実際には典拠に無いものも含まれている。

1.3 字典を典拠とする漢字の申請

典拠・用例の不明なものを除外するという方針は、JIS漢字策定作業の教訓が国際標準に反映された成果と言える。しかし、JIS X 0213策定で意識されたような、この漢字集合は何を表現できるのか(何を表現するために何文字追加する必要があるのか)を明確にすべきであるという考え方は、まだ統合漢字の標準化では十分に浸透しているとは言い難い。JIS漢字に対してしばしば為される批判・要望の一つに、何らかの漢和辞典に含まれる漢字を全

*1 「辭海」の初版は1936年に発行されたもので、現在の中華人民共和国で流通しているものは簡化された版である。
 *2 新華字典は中華人民共和国成立以降に出版されたもので、簡化はごく初期に完了している。新規に追加した漢字が申請された。
 *3 古典漢籍の用語辞典であるが、簡化表記するために多数の簡体字が新たに造字されている。この造字された簡体字が申請された。
 *4 本稿ではIRGの中で言う"source"を典拠、"evidence"を用例と訳す。JIS X 0208:1997, JIS X 0213:2000の標準化作業においては、漢和辞典に記載があることは、典拠にはなるが、用例とは認めず、区別している。IRGの場合、各国の採録申請書をsourceと呼び、各国がどこからその漢字をひいてきたかをevidenceと呼ぶ。
 *5 また、日本NBから申請漢字に5%以上の重複や字形デザインミスが見つかった申請は拡張C1から拡張Dへ延期することも提案も為されたが、これは合意されなかった⁹⁾。

ISO/IEC 10646:2003/Amd.5:2008 (E)

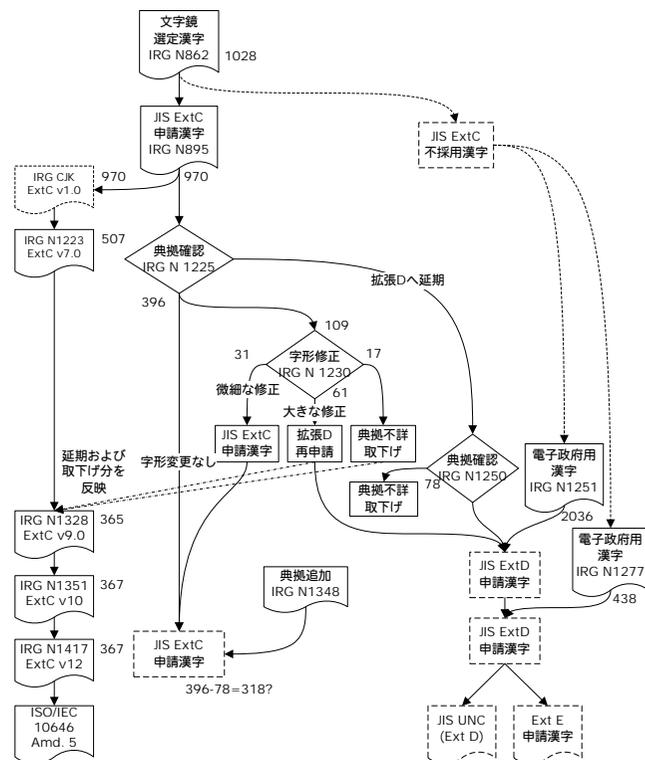


図2 拡張Cに申請された日本漢字の採録経過

て(相互に区別できるように)符号化してほしいというものがある。このような要望に対しては、漢和辞典は、既存の漢和辞典に含まれる難字・奇字は採録するものの、実際に地名・人名に用いられる漢字は漏れたままであるという反論があり¹¹⁾、JIS漢字の選定において字典の漢字を全て採録するという方針は今まで採られたことはない。

統合漢字拡張Bで既存の符号化文字集合からの採録が一段落したせいもあり、最近の申請では「ある字典に含まれる漢字を全て符号化する」という方針が目立っている。もともと、字喃(ベトナム申請)や方塊壮字¹²⁾(中国申請)など歴史的な漢字については、まず字典を選定し、その字典に含まれる漢字が全て符号化できるよう申請するという手法がとられていた。字喃や方塊壮字のようにもはや新たな造字がないと見られる場合は、原資料を直接あたらずに字典を典拠とすることの弊害は小さいと期待できる。

Ucode	C	J	K	U	V	Ucode	C	J	K	U	V
G	M	T				G	M	T			
2A700						𠄎					
						V4-4021					
						𠄎					
						G_ZJW00001					
										𠄎	
										K5-001C	

Index	Glyph	Meaning	Source	Location
zjw00001	𠄎		殷周金文集成引得 Index to Collection of Inscriptions of the Yin-Zhou Period	未查到。当刪。Not at this round of checking, to be withdrawn.

図3 ISO/IEC 10646:2003/Amd.5:2008の拡張C領域冒頭と、U+2A701の典拠確認結果²⁾。

一方、汎用的な字典を典拠・用例として採録申請しているものの、その字典の全ての見出し字を符号化するための申請なのか、使用頻度等によって絞り込んだ上での申請なのか、不明なものも多い。たとえば、拡張Bには、中国からの申請で康熙字典、辞海、漢語大字典などを典拠にして3万字近い漢字が申請されたが、拡張Cにもこれらの字典をもとに採録申請されており、これらの字典の全ての漢字が符号化されたのかどうか、また、符号化されていない漢字があるとすれば今後申請されるのか、不明なままである^{*1}。

中国の場合、当初は字典類から採取された簡体字が大半を占めていたが正式公布された拡張C漢字で最も多いのは「殷周金文集成引得」で用いられている金文由来の漢字366字である。しかし、これを典拠として申請された漢字は拡張Cまでで全て符号化されたわけではなく、拡張Eに積み残しとなっているものが1400字程度残っており、これで殷周金文集成引得の見出し字を全て符号化できるのかどうか不明でない。

2. 殷周金文集成引得の性格

「殷周金文集成引得」は、金文が鑄込まれている殷周青銅器の拓本の集成資料「殷周金文集成」に対する目録である。「殷周金文集成」の編纂作業とは独立に書かれていることには注意しなければならない。

「殷周金文集成」は基本的には拓本のみであり、金文が必ずしも判読し易い状態ではなく、また、その金文に対する釈文^{*2}も記載されていない。そのため、漢字の用例などを探す

*1 Unicode Technical Committeeが整備しているUnihanデータベースでは、拡張Bに含まれる康熙字典については管理番号が振られていない。

*2 古文書を読み易い文字に換字したり、文体を書き換えたもの。漢文の訓読文なども釈文である。金石学では一般に文字単位の換字を行ない、文体の書き換えまでは行なわない。

ことは難しく、この解決方法として「殷周金文集成引得」が書かれた。

金文を対象とする字書は「殷周金文集成引得」が最初ではなく、金石学の分野では過去に様々な字書が編纂されている。

- 説文古籀補¹³⁾ 呉大猷 (1883): 金文から石鼓文までを対象とする。
- 説文古籀補補¹⁴⁾ 丁佛言 (1924): 金文から石鼓文までを対象とする。
- 金文編¹⁵⁾ 容康 (1925): 金文を対象とする。
- 説文古籀三補¹⁶⁾ 強運開 (1933): 金文から石鼓文までを対象とする。
- 古文字類編¹⁷⁾ 高明 (1980): 甲骨文から石鼓文までを対象とする。
- 漢語古文字字形表¹⁸⁾ 徐中舒 (1981): 甲骨文から三国時代石経までを対象とする。

これらの伝統的な金石学字書と、近年発行されている「殷周金文集成引得」「金文集成」などには見出しの立て方や釈文の示し方に大きな違いが見られる。伝統的な金石学字書はまず説文解字による分類を引き写した上で、見出し字の説文小篆に対応する金文字形を列挙し、その拓本の出所を示す形式をとり、釈文は必ずしも示さない^{*1}。これに対し、「殷周金文集成引得」では数文字の拓本であっても全て釈文を示し、見出し字は説文小篆も含めほとんどを宋体^{*2} (全見出し字 4972 個に対し、金文図形のまま示すものは 280 個程度) に換字している。ただし、部首の立て方は説文に倣っており、康熙字典のような純粋な画数順にはなっていない。

次章で現代の字書の整理方法が金文の整理にそのまま適用できない問題点をまとめ、伝統的な金石学字書と、殷周金文集成引得の違いを整理する。

3. 金文字形の整理

金文の収集対象となる青銅器は BC1300 頃 (殷中期)~BC771(西周) と長期間にわたるため、この間の字形変化は無視できない^{*3}。また説文解字が書かれるまで (AD100) にも 900 年近い時間差があるため^{*4}、殷周金文集成引得の見出し項目数 (4972 項) は説文解字の見出し字 (9353 字) の半分程度であるが、説文に見えない漢字も多数含まれる^{*5}。

*1 金石学では説文に見えない文字も多数存在するので、それらを強引に小篆に対応づけるか、説文に無い見出し字を立てる際には様々な方法がとられる。大きな流れとして、古くは説文への附属書という体裁をとるため小篆への対応づけが行なわれていたが、近年では説文とは独立なものとして編纂するため新たに見出し字を立てるものが増えている。

*2 中国での日本の明朝体に相当する書体名で、日本で言う宋朝体とは異なる。現在の書体デザインは現代通用漢字字体表などによって規定されている。

*3 春秋戦国時代以降は陶器や簡帛の古漢字資料が増えるため、青銅器の字形が代表的なものとは言えなくなるが、漢代の貨幣までを金文に含める考え方もある。

*4 福田¹⁹⁾ は、金文字形と説文解字の籀文字形を比較し、説文の示す籀文は金文と同一または近似していると指摘しているが、比較に用いた金文資料が既に説文字形との対応付けが行なわれた金文編²⁰⁾ であるので、近似しているのは当然と言える。

*5 篆文になる前に使われなくなった文字と解釈され佚字などと呼ばれるが、小篆を経由せずに隷変したものは説文解字が無視している可能性も否定できない。

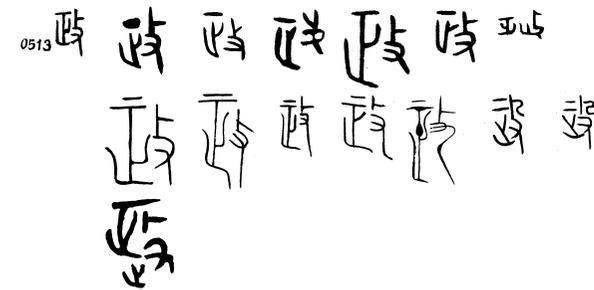


図 4 金文編²¹⁾ が「政」に分類した金文字形群

さて、現在使われる漢字字形は、漢代から唐代にかけて筆写体の書写能率を向上させるために筆画や字形が一度整理された隷書・楷書に基いている。現在の活字字形も楷書により整理された構造に大きな影響を受けており、字典の漢字分類もこれに強く依存している。同時に、漢字利用者が注目する図形的特徴もこの整理後の字形に最適化されている。金文はこの整理以前の文字であるから、字書を編纂する際に以下のような問題を考慮しなければならない。

- 画数の安定性
金文字形では、まだ線画整理が不十分なため、金石学での漢字同定では画数や線の断続をそれほど重視しない。そのため、宋体風の図形で見出し字を作り金文字形を指示しようとすれば伝統的な金石学字書よりも指示範囲が狭くなってしまう。たとえば、金文編²¹⁾ が「政」に分類した金文字形を図 4 に示すが、扁部を「正」で代表させて良いかどうかは自明ではない。
- 構成要素の増減
金文は部首としての整理が発展する段階の文字であるため、字形を構成する図形要素の何を省略するかはまだ安定していない。また、必ずしも古い字形が複雑で新しい字形が簡略化されているとも限らない。画数の場合と同様に、伝統的な金石学字書では見出し字が指示する図形的な範囲について、構成要素の増減をある程度許している。図 4 に示すように、小篆字形には無い構成要素を含む金文字形も同一字と整理する 경우가少なくない。しかし、活字字形で見出し字を示した場合に、このような対応づけを説文字形の知識がない利用者が暗黙に行なえるかどうかは疑問がある。
- 部首の相対関係
金文字形では、楷書のような扁旁冠脚への正規化傾向がまだ弱く、構成要素は現代の漢字に対応づけられるとしても、その相対的な位置関係が現在の組み合わせ方に対応づけられないものが多い。このような場合に、金文字形の相対関係を維持した字形で対応づ

(1)	(3)	(4)	(5)
0013 役、役	0218 顯、蠶	1737 盜、滌(漆)	4818 𠄎
0047 使、使	0232 顛、頭	16.10099 自乍(作)𠄎	12.7033 𠄎
0581 脊、脊	0276 妣、妣	(洗)盤	4819 𠄎(𠄎)
0866 兔、兔	0355 嬭、嬭	16.10290 蔡侯𠄎(申)	6.3663 𠄎(𠄎)
1217 釜、釜	0378 婦、婦	之尊𠄎(洗)𠄎	4820 𠄎(𠄎)
	0466 厄、輓	3059 洪(漆)	4821 𠄎
(2)	0800 殺、杀	16.10256 樊君嬰用自	
0142 𠄎(休)	1119 逋、逋	乍(作)𠄎(洗)也(𠄎)	
0155 𠄎(獸)	1233 𠄎、𠄎		
0167 𠄎、𠄎	1402 𠄎、𠄎		
0364 𠄎、𠄎	1411 𠄎、𠄎		

図7 殷周金文集成引得の項目見出し

- (1) 一点一画の増減や線画の断続を許している例
- (2) 部首の相対関係の違いを許している例
- (3) 構成図形の増減を許している例
- (4) 見出し字の重複使用を許している例
- (5) 金文字形を見出しに用いている例



図8 同一金文に対して金文編と殷周金文集成引得の見出しが異なる例

殷周金文集成引得は巻末に金文編の文字番号との対応表を収録しているが、漢字の同定は独自に行なっているため、金文編では同一字としているものを別字としたり、あるいは金文編では別字としているものを同字としたり、見出し字が金文編とは異なるものもある。図8には説文小篆と異なる金文字形に対してどのような見出し字を立てるかの違いの例を示した。

3.3 殷周金文集成引得を典拠とする申請の調査

殷周金文集成引得を典拠として申請された漢字 1852 字について、見出し字 1 個からなる項目で用いられるものは 733 字であった。このうち、468 文字は用例となる釈文が 1 例しか示されておらず、金文 1 文字を指示するのみに造字された図形と言える。次に、複数の見出し字を列挙する項目について考えると、国際標準への採録の観点からは列挙されている見出し字が本当に必要であるか(釈文に用いられているか)どうかにも注意が必要である。実際に調査すると、図9に示した 149 個は項目見出しに出現するが、その項目の釈文

0255-0047.2	使	0422-0676.2	𠄎	0483-1139.2	𠄎	0564-1491.1	𠄎	0688-2094.1	𠄎	0765-2531.2	𠄎	0944-2932.1	𠄎	1074-3670.3	𠄎	1128-3887.2	𠄎	1340-4515.1	𠄎
z.jw00333		z.jw00231		z.jw01539		z.jw01499		z.jw00782		z.jw00741		z.jw00381		z.jw01516		z.jw00556		z.jw00041	
0262-0087.1	做	0429-0702.1	𠄎	0483-XXXX	𠄎	0564-1494.1	𠄎	0696-2138.1	𠄎	0785-2560.1	𠄎	0971-2956.3	𠄎	1075-3675.1	𠄎	1145-4098.1	𠄎	1389-4566.1	𠄎
z.jw00706		z.jw00391		z.jw009173		z.jw00803		z.jw00690		z.jw01425		z.jw01566		z.jw00588		z.jw01186		z.jw00279	
0279-0142.1	𠄎	0445-0831.2	𠄎	0483-XXXX	𠄎	0564-1556.2	𠄎	0700-2180.1	𠄎	0800-2570.1	𠄎	1008-XXXX	𠄎	1076-3692.1	𠄎	1145-4100.1	𠄎	1389-4566.2	𠄎
z.jw00211		z.jw00627		z.jw01207		z.jw00425		z.jw01310		z.jw01301		z.jw01226		z.jw01591		z.jw00180		z.jw00971	
0284-0198.1	𠄎	0445-0832.2	𠄎	0491-1184.1	𠄎	0586-1576.1	𠄎	0707-2233.1	𠄎	0855-2635.2	𠄎	1010-3131.2	𠄎	1077-3696.2	𠄎	1147-4118.2	𠄎	1389-4566.3	𠄎
z.jw01850		z.jw00430		z.jw01493		z.jw01081		z.jw01735		z.jw01218		z.jw01237		z.jw00928		z.jw01195		z.jw00170	
0336-0462.2	𠄎	0452-0873.2	𠄎	0497-1223.2	𠄎	0586-1577.1	𠄎	0711-2269.1	𠄎	0855-2637.2	𠄎	1010-3133.1	𠄎	1079-3723.2	𠄎	1148-4133.1	𠄎	1468-4665.2	𠄎
z.jw00710		z.jw01501		z.jw00206		z.jw01356		z.jw01743		z.jw01217		z.jw01236		z.jw00895		z.jw01393		z.jw00429	
0343-0468.2	𠄎	0452-0874.2	𠄎	0497-1225.1	𠄎	0590-1604.1	𠄎	0724-2336.1	𠄎	0860-2647.3	𠄎	1021-3199.2	𠄎	1088-3752.1	𠄎	1148-4137.1	𠄎	XXXX-XXXX	𠄎
z.jw00122		z.jw00995		z.jw00837		z.jw00790		z.jw01595		z.jw00253		z.jw00502		z.jw01884		z.jw01395		z.jw00901	
0347-0511.2	𠄎	0460-0947.1	𠄎	0499-1233.1	𠄎	0594-1645.1	𠄎	0732-2345.2	𠄎	0863-2657.1	𠄎	1023-3221.1	𠄎	1088-3760.1	𠄎	1173-4161.2	𠄎	XXXX-XXXX	𠄎
z.jw01833		z.jw00448		z.jw00006		z.jw01843		z.jw01616		z.jw00372		z.jw01824		z.jw01894		z.jw01403		XXXX-XXXX	𠄎
0347-0511.3	𠄎	0460-0947.3	𠄎	0513-1260.2	𠄎	0599-1674.1	𠄎	0740-XXXX	𠄎	0864-2666.5	𠄎	1034-3309.2	𠄎	1089-3769.1	𠄎	1179-4168.1	𠄎	XXXX-XXXX	𠄎
z.jw01827		z.jw00443		z.jw00064		z.jw01867		z.jw01640		z.jw01055		z.jw01386		z.jw01886		z.jw00663		XXXX-XXXX	𠄎
0347-0511.4	𠄎	0463-0984.2	𠄎	0518-1319.1	𠄎	0612-1705.2	𠄎	0750-2434.1	𠄎	0865-2667.1	𠄎	1038-3343.5	𠄎	1090-3774.1	𠄎	1184-4229.3	𠄎	XXXX-XXXX	𠄎
z.jw01547		z.jw01206		z.jw01570		z.jw01321		z.jw01485		z.jw01045		z.jw01789		z.jw01872		z.jw00993		XXXX-XXXX	𠄎
0359-0536.1	𠄎	0466-0990.2	𠄎	0523-1355.1	𠄎	0619-1752.4	𠄎	0756-2461.1	𠄎	0865-2676.2	𠄎	1041-3360.2	𠄎	1091-3786.2	𠄎	1246-4254.2	𠄎	XXXX-XXXX	𠄎
z.jw00688		z.jw00378		z.jw01030		z.jw00681		z.jw00670		z.jw00383		z.jw00556		z.jw01510		z.jw00671		XXXX-XXXX	𠄎
0418-0624.1	𠄎	0474-1072.6	𠄎	0524-1377.1	𠄎	0620-1755.1	𠄎	0758-2470.1	𠄎	0874-2730.2	𠄎	1043-3374.1	𠄎	1092-3791.1	𠄎	1259-4269.2	𠄎	XXXX-XXXX	𠄎
z.jw00138		z.jw00990		z.jw00351		z.jw00138		z.jw00064		z.jw00315		z.jw01772		z.jw01913		z.jw00839		XXXX-XXXX	𠄎
0419-0630.1	𠄎	0474-1072.8	𠄎	0527-1413.1	𠄎	0645-1918.1	𠄎	0758-2473.1	𠄎	0888-2770.1	𠄎	1043-3383.1	𠄎	1093-3820.1	𠄎	1289-4361.1	𠄎	XXXX-XXXX	𠄎
z.jw00148		z.jw00900		z.jw01036		z.jw00260		z.jw00677		z.jw00100		z.jw01778		z.jw01535		z.jw00029		XXXX-XXXX	𠄎
0419-0636.2	𠄎	0477-1089.2	𠄎	0533-1431.2	𠄎	0646-1926.1	𠄎	0759-2479.1	𠄎	0910-2905.1	𠄎	1059-3530.1	𠄎	1093-3821.1	𠄎	1296-4432.2	𠄎	XXXX-XXXX	𠄎
z.jw00151		z.jw00338		z.jw00731		z.jw00901		z.jw00680		z.jw00174		z.jw00973		z.jw01537		z.jw00051		XXXX-XXXX	𠄎
0419-0637.2	𠄎	0478-1107.2	𠄎	0546-1448.1	𠄎	0679-1979.2	𠄎	0759-2486.1	𠄎	0910-2908.2	𠄎	1067-3605.1	𠄎	1096-3839.1	𠄎	1300-4452.3	𠄎	XXXX-XXXX	𠄎
z.jw01860		z.jw01668		z.jw00441		z.jw00945		z.jw01814		z.jw00817		z.jw00525		z.jw01798		z.jw00794		XXXX-XXXX	𠄎
0420-0641.1	𠄎	0478-1107.4	𠄎	0546-1451.2	𠄎	0680-2016.2	𠄎	0760-2496.1	𠄎	0940-2924.2	𠄎	1068-3621.2	𠄎	1114-3951.1	𠄎	1309-4486.1	𠄎	XXXX-XXXX	𠄎
z.jw00787		z.jw01667		z.jw00157		z.jw01380		z.jw01907		z.jw00942		z.jw01519		z.jw01584		z.jw01478		XXXX-XXXX	𠄎
0422-0669.2	𠄎	0478-1107.6	𠄎	0555-1472.2	𠄎	0686-2035.2	𠄎	0764-2514.1	𠄎	0941-2926.1	𠄎	1074-3669.2	𠄎	1116-3976.2	𠄎	1309-4493.2	𠄎	XXXX-XXXX	𠄎
z.jw00786		z.jw01024		z.jw00167		z.jw00303		z.jw01428		z.jw01806		z.jw01524		z.jw01506		z.jw01391		XXXX-XXXX	𠄎

図9 拡張 C, D に申請されたものうち、典拠項の見出しに出現するがその釈文では使用されていない文字の一覧。各図形の左上に典拠ページ番号・見出し項目番号・見出し字番号、左下に申請字の管理番号を示した。

で使用されていない。

4. 考 察

一般的な字書における見出し字は、図形文字の同定(この見出し字と同字と見て良いかどうか)が判断できることが重視される。伝統的な金文字書では、古漢字の見本字形として広く通用している説文解字の字形との字形差によって同定されることを期待し、見出し字を説文字形とした。また、小篆に該当するものが無い場合は、古漢字ではなく楷書によって見出しを立て、説文に見えないことを明確にするという編纂方針がとられる。現在では古文・籀文に相当すると思われる古漢字資料が多数存在するが^{*1}、それらを見出し字に採用せず楷

*1 古文・籀文という分類はあくまでも説文の中の分類であって、考古学的な実体に対応づけて良いのかは不明であるが、古文は戦国時代の秦以外の国で用いられた古漢字、籀文は戦国時代の秦で用いられた石鼓文に対応づけることが多い。

書の見出しを立てることは、楷書の造字は許されるが、小篆に類似した字形の造字は(それが典拠のあるものであっても)忌避されていると考えられる。

前節で整理したように、殷周金文集成引得は伝統的な金石学の字書の形式とは異なり、可能な限り見出し字を宋体にするべく大量の外字を作成した。この中には、伝統的な金石学の字書で導入された楷書字形も含まれるが、金文編での楷書見出し字が600字程度しかないのに対し、殷周金文集成を典拠に申請された漢字が1800字を超えることを考えれば、大半は見出し字を宋体にするために説文小篆や金文から作り直したと考えられる。

ISO/IEC 10646において、甲骨文・金文・篆文等を、漢字と別個に符号化することが合意されたのが2003年であるから²²⁾、これらの漢字が拡張Cへ申請された2002年の段階では、古漢字を指示するための宋体図形を漢字の一群として申請することは現実的な選択だったと言える。しかし、古漢字を漢字と別個に符号化することが決定した現在では、これらの宋体図形を漢字の一群として申請することが妥当かどうかは疑問である。

5. まとめ・今後の課題

ISO/IEC 10646の漢字領域拡張の動向として、既存の符号化文字集合を典拠とした申請が日中台韓で一段落したため、今後は古文書や印刷物を典拠とした申請に移っていくと考えられる。JIS漢字の策定では字典を典拠とすることの問題点は認識されているが、ISO/IEC 10646の策定においてはこの認識は広く共有されているとは言いがたく、今後も字典を典拠とする申請は多いと予想される。

この一つの典型として、本稿では、日中韓統合漢字拡張Cおよび拡張Dに提案された、「殷周金文集成引得」に由来する図形文字について調査し、以下のことを明らかにした。

- 申請資料が典拠資料のページ数しか示しておらず、用例確認に漏れが生じている。
- 見出しに見えるが釈文で使用されていないものが149字あり、これらについて「殷周金文集成引得」を用例と言えるか疑問がある。
- 金文字形を指示するために楷書・活字風の図形を導入することは殷周金文集成引得にはじまるものではないが、殷周金文集成引得が導入した図形が既存の金文字書と異なるものがあり、漢字として扱える程度に安定した字形と判断できない。

これを解決するためには以下の課題がある。

- (1) 金文字形の包摂規準
- (2) 前項により定まる図形的な範囲を指示するための宋体字形(群)
 - (a) 既存の金文字書で導入された宋体図形をどうするか
 - (b) 金文字形から宋体字形をひくか、宋体字形から金文字形をひくか
- (3) 前項の宋体字形が現在の通用漢字の包摂範囲と整合するかどうか

課題1の解決は、金文の標準符号化と整合性を持つべきである。また、課題2-aの解決のために金文字書の網羅的な調査が必要で、その調査対象についても金文の標準符号化と矛盾が生じないようにしなければならない。しかし、本稿執筆時点では古漢字の標準化グループは

甲骨文の標準化を行なっている段階で、金文に対する作業には着手できていない。以上の状況を考えると、殷周金文集成引得を典拠とする漢字の採録は審議延期の可能性も含めて検討すべきであろう。

また、一般に字典類を典拠とする申請の問題点として、申請字数が膨大になりがちであることが挙げられる。拡張Bまでに既存の符号化文字集合に含まれているため、今後典拠として用いられる字典は巨大なものか、あるいは特殊な用途のものになると予想される。特に難字・奇字については、他の字典を参照しているだけで実際の用例を欠く字書も少なくないため、字書を典拠・用例として良いかどうか判断規準の明確化が必要である。字書しか用例が無い場合の申請動機としては「ある字典の漢字を全て符号化したい」というものが考えられる。標準符号でこのような動機を考慮するとすれば、典拠となる字典が全て符号化できることを規格利用者に明確にすべきである。

参 考 文 献

- 1) ISO/IEC JTC1/SC2: *ISO Standards: Information Technology - ISO/IEC 10646:2003, Universal Multiple-Octet Coded Character Set (UCS)*, ISO (2003).
- 2) ISO/IEC JTC1/SC2: *ISO Standards: Information Technology - ISO/IEC 10646:2003/Amd.5:2008, Universal Multiple-Octet Coded Character Set (UCS)*, ISO (2008).
- 3) Zhoucai, Z.: *ISO/IEC JTC1/SC2/WG2/IRG IRG N1743: Summary of Extension-C Submissions* (2000).
- 4) ISO/IEC JTC1/SC2/WG2: *ISO/IEC JTC1/SC2/WG2 IRG N2204: Resolutions of Beijing meeting 38* (2000).
- 5) The committee for Standardization of D.P.R. of Korea: *ISO/IEC JTC1/SC2/WG2/IRG IRG N727: Proposal to add ideographs of D.P.R. of Korea* (2000).
- 6) The committee for Standardization of D.P.R. of Korea: *ISO/IEC JTC1/SC2/WG2/IRG IRG N731: DPR proposal to add to CJK C Form, Chart* (2000).
- 7) Sato, T.: *ISO/IEC JTC1/SC2/WG2/IRG IRG N830: Ext.-C ad-hoc group report* (2001).
- 8) ISO/IEC JTC1/SC2/WG2/IRG: *ISO/IEC JTC1/SC2/WG2/IRG IRG N907: CJK Extension C1 Version 1.0* (2002).
- 9) Ext C1 Editorial Group: *ISO/IEC JTC1/SC2/WG2/IRG IRG N922: Report from the Ext C1 Editorial Group* (2002).
- 10) Sekiguchi, M.: *ISO/IEC JTC1/SC2/WG2/IRG IRG N1100: Proposal to Reschedule the Ext-C1 Development Plan* (2004).
- 11) 芝野耕司: 増補改訂 JIS 漢字字典, 日本規格協会 (2002).
- 12) 広西壮族自治区少数民族古籍整理出版規劃領導小組: 古壮字典, 広西民族出版社 (1989).

- 13) 吳 大澂：説文古籀補，中華書局 (1988).
- 14) 丁 佛言：説文古籀補補，中華書局 (1988).
- 15) 容 康：金文編，貽安堂 (1925).
- 16) 強 運開：説文古籀三補，中華書局 (1986).
- 17) 高 明：古文字類篇，中華書局 (1980).
- 18) 徐 中舒：漢語古文字字形表，四川辭書出版社 (1981).
- 19) 福田襄之介：中國字書史の研究，明治書院 (1979).
- 20) 容 康：金文編，中國科學院考古研究所 (1959).
- 21) 容 康，張 振林，馬 國權：金文編，中華書局 (1985).
- 22) Guoying, L. and Bishop, T.: *ISO/IEC JTC1/SC2/WG2/IRG IRG N1014: Draft Agreement on Old Hanzi Encoding* (2004).