The University of Osaka
Institutional Knowledge Archive

Title	デジタル環境における出土文献の文字処理について
Author(s)	黒田, 秀教
Citation	中国研究集刊. 2004, 36, p. 89-93
Version Type	VoR
URL	https://doi.org/10.18910/60816
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

中国研究集刊 騰号（総36号）平成十六年十二月 89—93頁

デジタル環境における出土文献の文字処理について

黒田秀教

近年の中国思想史研究は、大きな変革期を迎えている。郭店楚墓竹簡を始めとする出土資料文献は、戦国秦漢思想史を塗り替えつつある。また電子技術の進歩によって、中国古代文献学にもデジタル化の波が押し寄せている。中国思想をデジタル環境で研究するに当たっては、文字処理が問題となる。デジタル環境で一般に用意されている文字は、あくまでも一般人の生活の上での利用を念頭に準備されており、中国古代文献上に頻繁に登場する特殊な文字は、用意されていなかった。これがために、研究者は必要とする文字を外字などとして自分で作成する必要があった。もっともこの問題は、Unicode の制定や今昔文字鏡の登場によって大きく改善されてきている。言うまでもなくこの二者を以てしても、中国古代文献上の文字を全てデジタル環境で表示できるわけではないが、研究者が自作する必要のある文字は大きく減少した。

ところが近年盛んである出土資料文献を取り扱う研究で、デジタル環境における文字処理の問題が再び大きく持ち上ることになった。出土資料上の文字は、楷書体の勃興より始まる現行字体とは大きく形を異にしている。研究の初歩段階で出土資料上の文字は現行字体に置き換えられているが、しかし中には、その文字が亡んでいるために、現行字体で代替不能の竹簡文字もある。この場合は、出土資料上の文字を楷書風に改めてデジタル環境で表示しなくてはならない。また現行字体に置き換えが可能であっても、極力出土資料上の文字に忠実な釈文を作成する場合には、やはり出土資料上の文字を楷書風に改めて電子媒体上で文字を表示しなくてはならない。更には文字の釈定が困難であり、そもそも楷書風に改めることのできないものも存在する。研究成果の発表手段の最

終形態は出版であるが、近年は出版物の版下作成もデジタル化されている。デジタル環境において出土資料の文字処理に手間取ると、研究成果の発表に遅延を招くことになりかねない。本稿は、このようなデジタル環境上での出土資料研究における文字処理についての一提言である。ただし、あくまでも当座をしのぐためのものであり、文字処理の問題に関する抜本的解決策ではないことをあらかじめ申し上げておく。また、Windows 環境を想定したものであることもお断り申し上げておく。

出土資料上の文字をデジタル環境で表示するための手法は、次の二つがあろう。一つめは、1「フォントで表示する」ことである。二つめは、2「画像で表示する」ことである。一つめの「フォントで表示する」手法は、更に二つに分けられる。一つめが 1-1「外字」としてフォント情報にすることで、二つめが 1-2「フォントセットを作成」してフォント情報にすることである。1-1「外字」としてフォント情報にするとは、独自に作成した文字を、「既存のフォント」の拡張領域に割り当てて表示することであり、windows 環境であれば、外字エディタを使用することで簡単に作成できる。1-2「フォントセットを作成」してフォント情報にするとは、「新規のフォント」を作成することである。2「画像で表示する」とは、出土資料上の文字を画像データとして読み込み、文書中にその画像を貼り付けて表示することである。次に、それぞれの利点と欠点とを列挙し、それぞれの作例を載せる。

1-1　「フォントで表示する（外字）」

利点

- Windows 標準機能で外字エディタが用意されているため、作成が容易である。
- デジタル環境で検索が可能である。

欠点

- 外字を作成したマシン以外での利用は、Microsoft 社の動作保証外となる。
- Windows 標準機能の外字エディタで作成すると BitmapFont になるため、粗いフォントになる。

・TrueTypeFont として作成する場合には、環境を整える必要がある。
・同時に二つの外字ファイルを使用することができないため、別のマシンで作成された外字ファイルを既に外字ファイルを持っているマシンで扱うときは、排他的利用となる。

1-2　「フォントで表示する（フォントセットを作成）」

利点

・独立したフォントセットなので、Windows 環境ならばほぼ全ての場合で利用可能である。
・作成に用いたマシン以外での使用も容易である。
・デジタル環境で検索が可能である。
・TrueTypeFont として作成すると、スケーラブル美しいものになる。

欠点

・作成のために環境を整える必要がある。

2　「画像で表示する」

利点

・出土文献上の文字を、そのままの形で表示できる。

欠点

・デジタル環境では検索が不可能である。
・画像を用意するために、環境を整える必要がある。
・画像として扱うため、文字に飾り付けすることや文書のレイアウト変更を行いにくい。
・文書ファイルのサイズが大きくなる。

以上をまとめるならば、まず、1「フォントで表示する」と 2「文字の写真で表示する」とを比較した場合、デジタル環境で扱う利便性は、「フォントで表示する」方に軍配があがろう。ただし 2「画像で表示する」は、出土資料の文字をそのままの形で表示できるので、書体などを問題にする文字学の分野ではこれを用いるべきである。

また釈定不能で出土資料上の文字をそのまま表示するしかない場合にも、用いられよう。ただし出土資料上の字形の再現性に厳密に拘らないのであれば、TrueTypeFont として作成する方法もある。なお画像として既に楷書風に改められた字体を取り込む場合もあるが、これはさして利点を見いだし難い。画像データの利点は、その文字をありのまま表示することだからである。そして 1「フォントで表示する」手段を用いる場合は、1-1「フォントで表示する（外字）」よりも、1-2「フォントで表示する（フォントセットを作成）」の方が望ましい。

現行字体（MS 明朝）	1-1	1-2	2
基 基 基	羽 羽 羽	羽 羽 羽	

1-1「フォントで表示する（外字）」の場合は利用に制約があり、研究に支障を来たす可能性がある。例えば研究者間でデジタル文書として研究内容を交換する時である。この時、送信先が独自の外字環境を構築していた場合、送信先はその研究内容を閲覧するために、自分の外字環境に代わって送信元の外字環境を導入しなければならない。1-2「フォントで表示する（フォントセットを作成）」であれば、デジタル文書とともに作成したフォントファイルを送信すれば、送信先はそのフォントをマシンにインストールするだけで閲覧が可能であり、且つそのフォントファイルが不用となれば、マシンからアンインストールするだけで良い。

ところで、1-2「フォントで表示する（フォントセットを作成）」場合に問題となるのは、どのコードにどの文字を当てはめるかである。出土資料上の文字を、対応する現行文字のコードに単に当てはめられるならば話は簡単であるが、実際には困難である。何故ならば、出土資料上の文字の釈読は研究者によって異なることもあり、また出土資料上の文字と現行文字とは、一対一の対応関係で無いものも存在するからである。更に事態を混乱させる要因は、出土資料上

の文字が現在進行形で増加していることにある。

そこで筆者は、文献ごとに分ける方法を提案したい。幸いにして出土資料上の文字についても、研究者が競うようにして逐字索引を作成している。これを用いれば、どの文献にどの文字が存在するかが容易に検索できるために、過去に作成したフォントを参照し易い。また取り扱う文献に必要とするフォントセットの選択も、自分の研究で扱う文献のフォントセットをインストールするだけであり簡便である。フォントファイルの整理も、フォントファイルにその文献名を名付けておけば便利であろう。

本稿では、デジタル環境における出土文献の文字処理について一つの提言を行った。始めにお断り申し上げたように、あくまでも当座の問題を回避するためのものであり、抜本的な解決にはつながらない。ここで筆者が提言した手法を用いても、研究者ごとの、或いは研究グループごとの、ローカルな環境が乱立することは避けられない。デジタル化が進んだ現在にあっては、デジタル環境における出土資料の文字表示手法について研究することも、重要な研究活動と言えよう。今後に求められることは、デジタル環境において出土資料上の文字を表示するための、統一的な指針を取り決めていくことであろう。本稿がそのための論議の先駆けとなれば幸いである。