

# 中国語図書WINEデータ入力現状

高木 理久夫（資料管理課）

## 1. 2001年～2005年までの進捗状況

2001年10月より、新規受入中国語図書のWINE書誌データ作成を開始した。それと併行して書架に配架済みの中国語図書についても、その書誌データを遡及入力していく作業も始めた。当初、データ入力作業は、図書館利用者用の端末機の文字コード環境が整っていなかったため、未だ試験入力の段階であった。Unicode対応のWINEシステム構築のため、この当時学術情報課は日夜、試行錯誤を繰り返していた。

2003年11月、蔵書検索において「多言語版WINE」が正式に公開され、中国語、ハングル等、日本語のコンピュータ環境において、いわゆる特殊文字を有する書誌データを利用者が検索する環境が整った。入力データはそのままWINE OPAC上に反映されるようになり、入力作業から「試験」という文字がなくなった。

2004年4月より、新規受入分中国語図書について、中央図書館の委託を請け、紀伊國屋書店がWINE書誌データの作成を開始した。これで入力体制が整った。『ふみくら』第70号（2003.1）に記しておいた課題が一応クリアされた。

2005年4月、図書館開館以来、続けられてきたカード目録作成を停止する。中国語図書の目録データは、今後WINEで維持されることとなる。

## 2. 入力データの現状と今後の遡及入力について

2005年10月までにWINEに入力された中央図書館所蔵中国語図書データ数は、次のとおりである。

書誌件数 38,900件  
（2002年10月時点 = 5,842件）  
図書冊数 105,108冊  
（2002年10月時点 = 10,276冊）

さらに、中央館所蔵の中国語雑誌約1,000タイトルの入力が完了している。

現在、中央館地下1階研究書庫に配架されてい

る旧分類中国語図書（イ、ロ、ハ・・・の請求番号がついているもの）の遡及入力を推進しており、入力が完了した分類部分は次のとおりである。

ロ（哲学）、ハ（宗教）、ホ（語学）、ヘ（文学）、ト（教育）、チ（芸術）、リ（歴史）、ヌ（伝記）、ル（地理）、ヲ（社会）、カ（政治）、ヨ（経済）、タ（財政）、ツ（交通）

残りは、イ（総類）他であり、よって2005年10月末現在、中国語図書に関して未入力データ図書冊数は残り約1万2千冊あまりとなった。

予定としては、地下1階旧分類部分については2006年6月をめどに遡及入力を完了させる。さらに引き続き、下記のスケジュールで中央館中国語図書の遡及入力を進行させるつもりである。

- (1) レファレンス図書コーナー配架中国語図書
- (2) 旧語学教育研究所移管本中の中国語図書
- (3) 入江、福島、澤田文庫等の中国語図書
- (4) 新分類中国語図書で未入力のもの

## 3. WINEで検索する際の留意点

中国語図書を検索する場合、「多言語版WINE」の利用を推奨している。これにより中国語表記が画面上にほぼ正確に表示される。しかし、検索等に次のようなことを留意する必要がある。

### (1) ピン音検索について

WINEで中国語図書を検索する場合、漢字とピン音両方で検索できる。ピン音検索の場合、固有名詞等は、スペースを空けずに検索語を入力する。

例1) 書名検索で「日本文化入門」と入力する場合

riben wen hua ru men

× ri ben wen hua ru men

例2) 「中華人民共和国」、「中華民国」

zhonghua renmin gongheguo

× zhong hua ren min gong he guo

zhonghua minguo

× zhong hua min guo

例3)「毛沢東」(人名の場合)

mao zedong \*姓と名の間にスペースを入れる。

× mao ze dong, maozedong

(2) 著者名漢字検索について

著者名検索において、漢字で検索する場合、姓と名の間にスペースを入れない。

例4)「毛沢東」と入力する場合

毛沢東 ×毛 沢東

ただし、これは中国語図書の検索の場合であって、日本語図書の著者名漢字検索の場合は、姓と名の間にスペースを入れる。

(3) 文字コードの不統合について

「多言語版WINE」はUnicode (UTF-8)対応になっており、複数の言語コードを統一的に縦覧することができるようになってきている。すなわちGB、BIG 5、Shift-JIS、それぞれ字形がちがっても表す意味が同じものは、簡体字、繁体字、当用漢字で検索しても、原則として同じ検索結果が、得られるはずである。しかし、どういうわけかその原則が適用されていない文字が、ままある。たとえば、「台」と「臺」、「勞」と「勞」は、同じ文字とは認識されておらず、「台湾」と「臺灣」、「労働」と「劳动」では検索結果が違ってきってしまう。

このような齟齬を防ぐためには、中国語図書を検索していて求める図書が見つからない場合、漢字検索ばかりでなく、ピン音検索も利用していた

だきたい。また、できればお手持ちのパソコンに中国語ワープロソフトを搭載し、中国語の文字入力環境を整えていただければ、漢字の不統合に対処できるかと思う。未だ完全な文字統合システムではないということをご了解いただきたい。

(4) 大部の叢書類について

『四庫全書』、『統修四庫全書』、『四庫全書存目叢書』、『叢書集成』、『百部叢書』等、大部の叢書類については、何千、何万にもおよぶ膨大な収載書について、その一つひとつの書誌を作成していない。現在の中央図書館の目録作成体制では、時間と労力があまりにかかりすぎるためである。それぞれの叢書の目録を利用していきたい。

4. 今後の展望

中国語図書の遡及入力を開始した当初、データ入力作業に未だ習熟していなかったため、完了までには最低でも10年はかかるだろうと悲壮な覚悟をしていた。しかし、システムの特徴を把握するにつれ、効率的に入力作業をおこなうことが可能となり、現在では月平均1千冊ペースで入力をこなしている。何事もなく、このまま順調に作業が進めば、2006年10月頃には、遡及入力作業はほぼ完了するだろう。

未だ文字コードに関しては問題があるが、これからの技術の進化が解決を促していくことと思う。

[ 図 ]  
多言語版WINEにおいて書名「零陵県志」と入力した場合の検索結果。  
『ふみくら』70号7頁(2003年1月発行)の画面と比較していただきたい。

