

コンピュータによる「白氏文集」総索引作成の諸問題

當山 日出夫

1. 「白氏文集」(中国の唐の時代の詩人、白楽天の作品集、全71巻)の総索引を、パーソナル・コンピュータによって作成する。その基本方針は、「文選索引」の形式にならった伝統的なスタイルの一字索引(単漢字のみの検索、熟語処理はしない)とする。これは、索引の実際の利用者である、中国・日本の文学・語学の研究者にとって、馴染み深いものであるとともに、簡便な利用が可能なものではないと考えるからである。

2. 「白氏文集」が、平安時代以降において我が国の文学作品に与えた影響ははかりしれない。そのため、「白氏文集索引」は、「白氏文集」そのものの研究に利用されることは当然であるが、同時に、日本・中国の比較文学的研究にも使えるものである必要がある。したがって、中国の唐時代以降(日本では平安時代以降)それぞれの時代にもっとも流布しよく読まれた代表的テキストを同時に検索でき、本文の異同も同時に見ることのできる校本索引を作ることが要求される。

3. 「白氏文集」という文献の量的な問題もさることながら、複数のテキストを同時に検索できる校本索引の作成は、コンピュータの利用によって初めて現実的に可能となる。従来の手作業による索引作成では、ほとんど不可能であったと言ってよい。

4. しかし、その過程として、実際に作成に携わる場合、もっとも大きな問題点として、複数のテキストを調査し読み比べて、本文の異同を調べる作業がある。基本となる本文の入力それ自体は、全体の労力からすれば微々たるものにすぎない。この校異をとっていく作業そのものは、機械化が不可能であり、写本・版本を直接比較し解読していく熟練した手作業によってのみなすうる仕事である。

5. コンピュータの利用によって、校本索引の作成が飛躍的に容易になり、確実なものが作成できるようになったが、その反面、校異を調べるといふ膨大な手作業を必要とする、ということでもある。

6. 成立の時代も違えば、それぞれに本文の質も異なっている複数のテキストにまたがって統一的な本文校訂をほどこなさなければならないので、漢字の字体の処理については、かなり強引な処置をとらざるをえなくなる。どれか固定した一つのテキストを対象とする場合と異なり、可能な限り厳密にある原本に忠実にという方針は採用することができない。むしろ逆に、いかにして全体として通行の字体に統一するか苦慮することになる。

7. テキスト・データベースは、それが単独で利用されるよりも、書物の体裁の本文なり索引なりと併用してこそ、その特性が発揮される。そのためには、将来的な課題として、テキストの原本の画像データとのかかわりが考慮されねばならないかもしれない。現在では、まだ、テキスト・データベースの現実的な利用方法そのものについて未知数の部分が多いのであるが、その利用を開拓していく時、画像等の周辺分野との協力関係が重要になってくるであろう。