
重点領域研究「沖縄の歴史情報研究」に参加して

赤嶺 守：琉球大学法文学部

4年前に「琉球・沖縄の対外関係史」研究班代表の金城正篤先生から重点領域「沖縄の歴史情報研究」研究分担者の依頼のお話をもちかけられた時、歴史情報の電算化についてほとんど知識のないまま、当時は確か「一太郎 Ver.4.3」のワープロ機能をフル活用し、外字の多い論文の引用文献はパソコンボーイに外字を作ってもらい何とかクリアしていた私は、ひょっとしたら歴代宝案の編集作業に使えるかもと思い、気軽に「はい、はい」とお受けした。

金城班では『歴代宝案』の各文書に文書番号を付し、年代、文書形式、発信者、受信者、写本・影印本の現存状況などの歴史情報を提供するデータベース作成を中心に作業を進め、同時に関連史料を統一した形で情報化する基礎的作業として中国第一歴史档案馆所蔵の『清代中琉関係档案統編』『清代中琉関係档案選編』や『中山世譜』の全文テキストおよびデータベース作成の作業も始まった。

当初、全文テキストは入力作業を学生に依頼して外字や異体字など処理できない漢字は全部ゲタで処理するよう指示し、出来上がってきた本文テキストを自分で校正をかければよいだろうと脳天気にも構えていた。1ヶ月後、学生がもってきた外字・異体字のゲタだらけの本文テキスト初稿を見て「啞然・・・」とした。

想像以上に出た異体字をどう統一すればいいのか分からなかったし、「一太郎 Ver.4.3」の最大外字登録可能数をはるかに越える外字が出るであろうことも容易に推測できた。「どうしよう・・・」と台湾出身の女房がいった。パソコンボーイも匙をなげた。しょうがないので異体字に関しては「歴代宝案室」の異体字の統一方法を参照して、私よりは漢字につよい女房にすがりながら1字1字統一をかけていった。外字処理に関しては総括班にあずければ、これまた何とかなるだろうと他人任せで考えていた。

外字処理の問題は、総括班の主催する研究代表者会議でも何度も問題にされたが、『清代中琉関係档案統編』『清代中琉関係档案選編』『中山世譜』といった抬頭のある文書特有の書式を変えずに、文字化けせずにオンライン化による情報交換を可能にする方法を見つけれないまま、1年があっという間に過ぎた。2年目に入ったある日、パソコンの入力を覚えはじめていた女房が、「半角で何なの」ときいてきたので、「1文字の半分」と適当に答えていたら、しばらくして外字も半角2つのアルファベットや数字を組み合わせて記号化すれば原文書のスタイルを変えずに処理できるのではとやってきた。ひょっとしたらと思い、試しに入力して検索してみると外字の記号がうまく引っかかってきた。単純な発想に「赤嶺コード」と命名していただき、ちょっとおこがましい気持ちもあるが、しかし一時的な場つなげ的方法にせよ、アルファベットと数字を組み合わせ2バイトで表現する識別子を付加することによって、当初考えていた『清代中琉関係档案統編』『清代中琉関係档案選編』『中山世譜』といった抬頭のある漢文文書特有の書式を変えずに、文字化けせずにオンライン化による情報交換が可能になった。しかし、これはあくまでも本文テキストに校正をかけて完成させるまでの処理方法あるいは情報のオンライン化としては有効であるが、個人が実際にパソコンを利用して情報検索等をおこなう場合は外字フォントを作成し、それを Windows95 の外字エディタを使って自分のパソコンに文字移植をして本文テキストを Windows95 版のアプリケーション上で機動させるのがベストである。

『歴代宝案』関連史料の情報化に関しては、『琉球家譜』、冊封使の『使録』類、『明実録』や『清実録』の本文テキストやデータベースも完成しているので、この際、欲張って一気に『宝案』の档案関連資料を全てパソコン入力しようかという野望を、最近ほんのちょっとだけ抱きつつある。