

『国文学とコンピュータシンポジウム(第2回)』 見聞録

平成二年十二月十四日東京、地下鉄を降りた所で道に迷った。十数年以前に、数度訪れた国文研(国文学研究資料館)ではあるが全く見当が付かず、長い信号を待って、わざわざ大通りを越えてしまったようである。再び、地下鉄に戻り改札の前で地図を見たが、思い出せない。所在なく振り返ると、二人の男性が階段を上っていた。後ろ姿からして、シンポジウムに参加する風体に見え、後をつけることにした。

迷った動揺がおさまり、先行の二人を落ちついて観察してみた。一人は大型計算機センターの星野聡先生であることに気づいた。ともかく、足早に後を追い挨拶をした。もう一人は、大阪国際大学でタイ語のコンピュータ処理を研究されている柴山守先生であった。ずっと昔に話をしたことがある。迷ったことを星野先生にいうと、「国文研には、道をジグザグに行けば、たどり着きます」と仰った。当日の講演者に、事前に話をするのも、面はゆく感じられた。

会場には百名近くの人が集まっていた。見知った人も幾人かいた。参加者の中核、顔ぶれというもの、もう構成されている証であろう。この分野は成長しつつある。一年前の第一回の講演録では、「パソコン」の応用が柱の一つであったが、今回はその上位にあたる「ワークステーション」に話が移っている。空席を探していると、工学部の長尾真先生にぶつかった。「お話をうかがうのが楽しみです」と私。「いやなに、ボケたことしか話せへんですな」と、さらりと関西風に破顔一笑された。

今日は奇しくも、箱根を越えて京大から三人もの講演者が出ている。長尾、萩谷昌己(数理解析研究所)、星野の三先生である。この組み合わせは、シンポジウムが始まる前から、期待させるものがあつた。お三方は、この会では情報処理関係者として、捉えられていたが、各先生の論文や著書に目を通すと、すべて若き日に、哲学や文学、歴史分野に相当深く関与しておられたようで、今

回の参加もなるほどと、一人得心した。

九時すぎから夕刻五時すぎまで、ぎっしりつまったプログラムの中心は、国文研の先生方のものである。主調は古典本文データベースを構築することの考え方と、方法論である。一番の安永尚志先生と最後の北村啓子先生とが、旧岩波古典文学大系百巻のデータベース化に関する技術的な側面を受け持たれた。まず、古典の全文がデータベース化されることは、研究者や一般読書人にはかり知れない便宜を与える。しかもCD-ROM化に現れているように、その目的はパーソナルユースにある。全文データベースは利用者が自由にアクセスでき(標準化、データ流通)、その結果を、多様な研究の展開に応用できるようになっていなければならない。その成果は今、手で持てるCD-ROMに凝縮している。

竹取物語が縦表示のオンラインKWIC牽引としてパソコンに表示されるのを見て、あらためて成果の着実な進展を味わった。

一方同館の新井栄蔵先生は、大系の古今をノートパソコン上で自由に検索する様子を非常に興味深く話されていた。国歌大観番号と歌がリンクしていく部分に将来のひろがりを感じた。しかし、含羞のあるとまどいも隠されてはいなかった。国文学に限らず、人文科学にあつては原典そのものの正確な再現性というものが、急所となるようである。新井先生の研究生生活にあつては、写本を写真にとったいわゆる影印本ですら、別の、異本と考えなければならない学風の中で過ごしてこられたようである。

そのような厳密さと標準化、そして流通の狭間になつて、今後どのような方向性を持つのか、その見極めが、なかなか難しいのではないかと、私はきままに憶測した。

共立女子大学の内田保廣先生は、近世文学を題材に、データベースの普及というものに力点を置いておられた。何よりも、カードや付箋によって処理されてきたテキストが、データベース化され

た場合の扱い易さや研究の効率のよさを、多くの人に知って味わってもらわねばならない。ご自身のパソコンライフにからめて、そこらあたりのリアルな様子を強調されておられた。

こうした国文学オリエンテッドな方向とは異なる観点から、日本語（辞書）処理そのものに関して長尾先生の特別講演があった。まず OED（オックスフォード・イングリッシュ・ディクショナリー）第二版の CD-ROM 化の経過報告に、驚きを禁じ得なかった。二十巻二万二千頁、五十万単語、二百四十万件の引用文献の内容が、辞書の形をとりながらテキストの様に（語の用例が完全な引用文として見れる）格納されている様子は、その膨大さに圧倒される。そのような網羅性完璧性を踏まえた上で、これからの日本語の辞書は、文を一定の要素に区切る形態素解析を経て、文法情報や意味情報を、連想まで含めて扱えなければならない。もちろんこの順逆は不二のことである。多様な情報の関係性の中で、高度な形態素解析も可能となるのだから。そういう助けになる辞書という基礎があってこそ、テキストのデータベース化された結果を飛躍的な応用に導く。と、このような趣旨であった。

長尾先生のお話は、村上征勝先生（統計数理研究所）の「統計的手法による日蓮遺文の文体研究」に接したとき、さらに興味がつのった。日蓮上人の写本の真偽を計量的に判定した内容であった。動詞や名詞の使用の変化の中に、異なった文体がある。上人の佐渡流罪の数年前に現れた全体的変化とは異なる、別人の手になるものと推量される部分が、明瞭に区別できる様子が良く分かった。

さてここで、萩谷先生のお話は、抱腹絶倒のものであった。紙上ではとても再現できない。先生は数理解析研究所でソフトウェアの研究をなされている。もれ知る所では、その世界は芸術、いや魔術に近い世界らしい。あくまで伝統的な数学者ではない。その研究所で発生する悲喜こもごものお話であった。頭脳明晰な純粋数学者は、数学以外に頭を使うのを極端に嫌う、よって一般人の予測を裏切るものだが、数学者にとってコンピュータはほとんどの場合、無縁のモノらしい。しか

し便利なのは便利であり、それに気付いた先生方は少し触ってみる。それが奇想天外な利用であり、作業中に電源コードを引き抜くくらいまだましな方である。結局誰かが一人、人身御供（イケニエ）になって、一日中そういうことの後始末をしなくてはならない。パソコンの上位にあるワークステーションになると、その壮絶さは極みに達する。会場は爆笑に包まれたが、このアナロジーは、国文学と情報科学とに置き換えたとき、どうなるのか、私はいささか悩んだ。なお、先生が本論の「ワークステーション上の日本語処理」については、大部な冊子を配布されていることを、お断りしておく。「詳しくは、これを読んでおいてください」と。

その、ワークステーションを京大の大型計算機センターの研究室においたまま、会場のパソコンから検索実験をされたのが星野先生であった。五百キロを越えて、研究室の続日本紀本文を自由に操れることの意味を実地にかみしめてみた。ワークステーションがネットワークに適しているとはいっても、興味のある仕事が実際に目の前で動くのを見るのは、始めてであった。続日本紀という、市販ベースにはのりにくく、かつ幾人かの研究者にとっては必須のデータベースが、いとも容易に、パソコンさえあれば、利用できる体制にあるのである。少数の研究者グループによる共同研究が、現代のハイテクによって活性化していく可能性が感じられた。なお、先生は古典に記された風景や天文をもとに、コンピュータ上にそのイメージを再現する研究を続けておられる。これが、学内 LAN を通していろいろなところで見られるようになる日も、近いことであろう。

さてここまで紹介して、時間は四時前。以前から知りたかったことを小さな脳髓に全て詰め込み、私の神経は麻痺してしまっていた。残りのパネルディスカッションは夢うつつで聞き終えてしまった。終了は六時頃だった。懇親会は場違いなので、ただただ片隅でひっそりとサンドイッチなど食していたが、ちょっとした機縁で館長の小山弘志先生に挨拶するはめになった。先生は、盛会であっ

たことを心から喜んでおられるご様子であった。

私は早めに失礼し、夜道を旧知の国文研情報サービス室長にともなわれ、迷路を逆戻りした。新幹線の席についたときは、ほっとした。国文学とコンピュータ。ジグザグと進展し、やがて深い森

をぬけて峰の頂に立つのであろう。

附属図書館 図書館専門員 谷口敏夫
(平成三年四月)



編集後記

- 第100号をお届けいたします。
今回は通巻100号となりましたので、巻頭言を西田館長にお願いいたしました。
- 京大は図書館内に百年史史料室を設け、創立百年記念事業の役割を分担しています、また、近々のうちに全学の蔵書が500万冊に達します。これら区切りの良い数字は一つの通過点に過ぎないかも知れませんが、これを見守る側の者としては感がい深いものがあります。