

BULLETIN OF THE EUROPEAN ASSOCIATION OF SINOLOGICAL LIBRARIANS

NUMBER 2

JULY 1986

BULLETIN OF THE EUROPEAN ASSOCIATION OF SINOLOGICAL LIBRARIANS

NUMBER 2

JULY 1986

CONTENTS:	FIFTH EASL CONFERENCE	3
	ANNOUNCEMENTS	4
	LIBRARIANS	6
	OBITUARY	7
	CHINESE COLLECTIONS IN EUROPE (2)	9
	Howard Nelson: PROBLEMS IN CHINESE BIBLIOGRAPHIC AUTOMATION	12
	Andrew Wang: OCLC'S CATALOGUING CAPACITY IN THE CHINESE, JAPANESE AND KOREAN LANGUAGES	32
	REVIEWS	36

Editing, preparation of camera-ready copy:

David Helliwell,  
Department of Oriental Books,  
Bodleian Library,  
Oxford OX1 3BG,  
Great Britain

Printing, distribution:

Johann-Michael Streffer,  
Ostasienabteilung,  
Staatsbibliothek Preussischer Kulturbesitz,  
Potsdamer Strasse 33,  
D-1000 Berlin 30,  
Germany (BRD)

Copyright 1986 by the European Association of Sinological Librarians

## FIFTH EASL CONFERENCE

summary of proceedings, by David Helliwell (Bodleian Library, Oxford).

The fifth EASL conference was held in the Staatsbibliothek Preussischer Kulturbesitz, Berlin, from 5th to 7th September, 1985.

The question of Chinese bibliographic automation was discussed, and it transpired that no European library had plans at present to automate the cataloguing of its Chinese holdings. It was announced that the Department of Oriental Manuscripts and Printed Books at the British Library was to have the use of the RLIN CJK system on an experimental basis for one year, beginning in 1986, and that its experience would be reported to EASL in due course. A representative of OCLC should be invited to demonstrate the OCLC CJK system at the next conference, when its suitability for use in European libraries could be assessed.

Johann-Michael Streffer explained the Federal Republic's "Zeitschriften Databank" (ZDB), and confirmed that it would be possible to augment it with non-German locations of Chinese periodicals. Members felt that this would be the best way of producing an updated union list of Chinese periodicals in European libraries, and undertook to participate subject to the following conditions: input would be in pinyin romanisation; it should be possible to obtain output of Chinese titles uniquely; output should be available in both microfiche and print; and output should be available regularly to participating libraries.

Sample copies of the inaugural issue of the bulletin were examined, and general satisfaction expressed with its form and content. Thanks were expressed to the Staatsbibliothek for sustaining the cost of printing and circulating the first issue. The second issue would be produced in spring, 1986, and the Staatsbibliothek would again sustain the cost of printing and circulation, but at the next conference the progress of the first two issues would be reviewed, and the possibility examined of charging for the third and subsequent issues.

A visit was made to the Asien-Afrika-Abteilung of the Deutsche Staatsbibliothek in East Berlin, and members were warmly received by Dr Karl Schubarth-Engelschall, the director, Frau Helga Keller, the Chinese librarian, and Herr Johann Dill, her assistant.

The next EASL conference would be held in conjunction with the EACS conference in Turin, in August, 1986.

## ANNOUNCEMENTS

The deferment of the two principal items of business for the sixth EASL conference until 1987 (i.e. the incorporation into the ZDB of non-German locations of Chinese periodicals and a demonstration of the OCLC CJK system) has left the agenda painfully thin. The SIXTH EASL CONFERENCE has therefore been reduced to one day, immediately before the EACS conference: SATURDAY 30th AUGUST, 1986. Members who have notified me of their intention to attend will soon receive details of time and place; those who have not and who wish to come should notify me immediately. David Helliwell, Department of Oriental Books, Bodleian Library, Oxford OX1 3BG, England.

The DEUTSCHE STAATSBIBLIOTHEK has microfilmed the 3-volume manuscript catalogue of its pre-1939 Chinese holdings in which several members expressed interest when we visited the library during the Fifth Conference. The catalogue, which is in the form of a handlist, was compiled between 1847 and 1939, and lists the Chinese books which had been acquired by the Preussische Staatsbibliothek up to the outbreak of the Second World War. The microfilm is valued at US\$80.00, and the library would prefer to receive payment in the form of exchange materials. Terms should be arranged with Dr Karl Schubarth-Engelschall, Direktor, Asien-Afrika-Abteilung, Deutsche Staatsbibliothek, Unter den Linden 8, DDR-1086 Berlin.

A new library has been formed in STOCKHOLM with the name OSTASIATISKA BIBLIOTEKET (THE FAR EASTERN LIBRARY). The library is in the Museum of Far Eastern Antiquities, and results from the merging of the Far Eastern language materials formerly in the following major Swedish collections: The Royal Library, Stockholm University Library, and the library of the Museum of Far Eastern Antiquities. The new library will be a major centre for research in Chinese, Japanese and Korean studies, and will also contain material in western languages. Visitors to the Library should enter by the southeast gate to the Museum. The new postal address is:

Ostasiatiska Biblioteket (The Far Eastern Library),  
Box 16358,  
S-103 27 Stockholm

Tel: 08-24 42 00 (Information & Reading Room)  
08-20 45 43 (Librarian, Lars Fredriksson)

The East Asian Studies department of PRINCETON UNIVERSITY is planning to publish a revived ASIA MAJOR, having acquired the rights and back stock from its former British publishers. The journal will publish scholarly articles on all aspects of Chinese culture and history, except those that are highly specialised or that require a large number of plates. It will include articles on Korea and Japan where these concern relations with China. Contributions should generally not exceed 50 printed pages (75 of typescript), and may include Chinese and Japanese script wherever necessary. It is planned to include periodic surveys of current scholarship, concentrating on current scholarship in China and Japan. Two substantial parts of approximately 240 pages will be published per annum. Contributions should be sent to Dr Howard L. Goodman, 211 Jones Hall, Princeton University, Princeton, New Jersey 08544, U.S.A.

The Tenth Annual Conference of the BRITISH ASSOCIATION OF CHINESE STUDIES will be held at Cambridge from 19th-21st September. Visitors are welcome, and further information is available from the secretary: David Chambers, Department of Extra-mural Studies, University of Bristol, Wills Building, Queens Road, Bristol.

LI ZHIZHONG, Acting Head of the Rare Book Section of the National Library of China, Beijing, has been invited to come to England for three weeks in September (probably the last three, but dates have yet to be confirmed). He will address the Tenth Annual Conference of the British Association of Chinese Studies in Cambridge and will spend the rest of his time examining the rare Chinese books in the British Library, SOAS Library, the Bodleian Library, and Cambridge University Library. The visit has been arranged with the help of the British Council and the Universities' China Committee. We should be very happy if a small group of our EASL colleagues would like to meet him during his stay, and further information is available from Frances Wood in the Chinese Section at the British Library.

## LIBRARIANS

CHARLES AYLMER has been appointed Assistant Under-Librarian in charge of the Chinese collections at CAMBRIDGE UNIVERSITY LIBRARY. In addition to the general collection of some 60,000 volumes, these comprise the Wade Collection of Chinese books, the Hopkins Collection of oracle bones, Chinese manuscripts, and printed ephemera such as the unique T'ai-p'ing materials. After graduating from Cambridge, Charles was an exchange scholar for two years at Peking University, where he studied philosophy and Chinese palaeography. Since 1980 he has taught Chinese in the Faculty of Oriental Studies at Cambridge, and has worked on a number of research projects, including computer-aided lexical analysis of Chinese texts, machine translation and lexicography. His publications include works on Chinese palaeography and computer-generated concordances to Chinese poetry; a concise Chinese-English dictionary of modern written usage is in preparation.

JOHN CAYLEY has joined Frances Wood and Beth McKillop in the Chinese section of the Department of Oriental Manuscripts and Printed Books in the BRITISH LIBRARY. John graduated from Durham University in 1978, and has subsequently been writing a thesis on parallelism in mediaeval Chinese prose. He has also worked on more modern material for the new East Asia Centre of Newcastle University. His other interests include the literary translation of Chinese poetry, and calligraphy.

Mrs YUNG-TZU WU, formerly assistant librarian at the Fung Ping Shan Library of the University of Hong Kong, has been appointed librarian of the SINOLOGISCH INSTITUUT at LEIDEN. Mrs Wu succeeds John Ma, who retired from this position in 1985. Apart from being responsible for the overall functioning of this library, Mrs Wu will maintain and expand the library's national and international relations, mainly with institutions in the Far East. She will also maintain the Leiden connexion with EASL.

## OBITUARY

WOLFGANG SEUBERLICH, 1906-1985

It was only a few days before EASL met in Berlin last year that a great colleague, sinologist and librarian died in Marburg. Wolfgang Seuberlich was formerly head of the Ostasienabteilung of the Staatsbibliothek Preussischer Kulturbesitz in Berlin (previously the Westdeutsche Bibliothek in Marburg). However, as an entirely new generation of European sinological librarians has now emerged, it is not surprising that quite a number of EASL members in Berlin did not know the man who had not only built one of the finest East Asian collections in Europe, but had also endeavoured to organise librarianship in our field beyond his immediate circle.

Yet these activities only occupied his later life. Although very little of his earlier career foreshadowed his achievements in our library during the fifties and sixties, these point in many ways to his younger years, when he accumulated unique experience and extraordinary knowledge.

Actually he was baptised with water from the Spree, something which no Berliner would not be proud of, but his parents soon left for their home country, the Baltic states of Russia. A visit to relatives in Manchuria in 1914 coincided with the outbreak of the First World War and made return impossible: history became fate for Germans in Russia, especially for those in the Far East.

Seuberlich turned "Landsmann" and took what education he could under the circumstances, primarily Russian in the pre-War tradition. He entered university in his then hometown of Harbin, where he studied law and occupied himself with learning Chinese. Very soon he started teaching Chinese, and after graduating he joined the customs service of Manchukuo while still continuing to teach at Harbin University. Having written his first scholarly articles in Russian, he soon turned to German, and published mostly in German periodicals on political science. He studied Japanese and became a specialist in Manchurian problems.

Only in 1937-1938 did he return to Germany, where he became a lecturer in Chinese at the famous Auslandshochschule in Berlin and took a course in sinology at Berlin University, finally receiving his Ph.D. in 1941. By then history was again his fate: national service, and Berlin in 1945. However, through his good knowledge of Russian he made his first contact with the Library: he acted as an interpreter and teacher of that language from May 1945 till June 1949 when he left Berlin for Mainz. But in 1952 he returned to the Library, then at Marburg, as the librarian responsible for Far Eastern languages.



Starting work on a non-existent collection - almost nothing of the large East Asian collection of the Preussische Staatsbibliothek had come to Marburg - he retired twenty years later from an East Asian department which employed sixteen members of staff, commanded considerable funding, and which had become a major European collection. These achievements speak for themselves.

Never having married, he was free to engage in many activities outside the Library: he frequently participated in conferences, many of which he organised himself. After retiring, he made several visits to China, and loved to reminisce about his adventures with his intimate circle of friends. A man of very regular habits, he always remained a "Landsmann": a German from Manchuria, with a unique love of his origins in that strange exile.

Johann-Michael Streffer (Staatsbibliothek Preussischer Kulturbesitz, Berlin).

## CHINESE COLLECTIONS IN EUROPE (2)

### A NOTE ON PUBLICATIONS FROM THE PRC IN THE STAATSBIBLIOTHEK PREUSSISCHER KULTURBESITZ, BERLIN

by Johann-Michael Streffer.

This note may be taken as a resume of 10 years' experience in book acquisition from China by the Staatsbibliothek in Berlin.

In the early seventies the resumption of publication in the PRC really took us by surprise. Although we had still been buying materials from Taiwan, Hong Kong, &c., we had to adapt to the very new kind of book market that sprang up after the Cultural Revolution under the Gang of Four. There were many pamphlets that we could buy, and many rumours about all the neibu material that we could not. As a kind of backup library for the Federal Republic, we decided on the basis of experience to concentrate on material that was freely available. Our main reason was that the hunt for valuable neibu material was very time consuming and usually ended in failure. Also, observation of readers' interests suggested that we should broaden the collection into many more "trivial" fields. Last but not least, we expected the practice of designating material as neibu to be abandoned.

This led to our placing a blanket order with Joint Publishing Company in Hong Kong in 1977. The specification of this order was very broad, but its relatively low cost meant that there were no financial obstacles to be overcome. It soon became apparent that the decision had been taken just in time to catch the train of the rapidly expanding publishing industry of the PRC in 1978-1979. By about 1980, only one problem seemed to remain: our supplier, though very efficient, was not really close to the sources of book production. So at the Frankfurt Book Fair in 1981 we established relations with a Beijing bookseller and transferred our blanket order - actually on the recommendation of Guoji Shudian! And book production still rose. I need not recount the official statistics of PRC publishing, but as far as our library is concerned, in 1985 we received about 12,000 books and 1,500 periodical titles, i.e. roughly 25-30% of actual production. This meant work, and even overtime in the library.

At an early stage we made many experiments in ways to speed up the processing of material, under the condition that standards should be maintained and proper acquisition and cataloguing guaranteed. For various reasons Library of Congress cards were of no help: there was no technical substitute for time-consuming original cataloguing. The problem was solved when the National Library of Beijing made an effort to

supply us with cards. It has therefore been possible since 1982 to develop a comparatively simple but effective procedure which enables us to maintain both quality and quantity without having to raise new manpower. We have even managed to improve the speed at which books reach the shelves, which is on average one year after publication for monographs, and six months after each completed year for periodicals.

Since our cataloguing is now derived from the National Library of Beijing, we have tied our acquisition policy as closely as possible to theirs and rely extensively on technical procedures, our guiding principle being cataloguing first, accession second. Cataloguing means joining a card with a book, but also checking the completeness and accuracy of the record. The first step makes use of booknumbers (i.e. the standard Chinese number, which is not as yet an ISBN). Incoming cards are filed according to these numbers, so that filing errors can only arise from incorrect numbering or human error, rather than any contradictory interpretation of the rules; the same applies to the joining of book and card. The second step, checking the completeness and accuracy of the record, is cataloguing proper. About 25% of the cards need completing or amending to suit our needs, mostly concerning the manner in which they are incorporated into our catalogue. Obviously this is the cornerstone of the process.

The process of accession, by which we register ownership of the book, involves: (1) identifying the book, and linking it to the catalogue entry with an accession number that also serves as a shelf number; (2) producing the catalogue entry itself; and (3) linking the entry with proof of proper invoicing and payment, i.e. adding the accession/shelf number to the invoice. We produce the catalogue entry from a pre-printed form which already contains most of our required standard information; to this is joined the Beijing card after the accession/shelf number has been added; it is then photocopied on to one card. The shelf number is then added to the book, which can then be shelved. The invoice is marked with the accession number of each item, and the price and invoice date of each item is noted in the accessions register.

The following steps then follow naturally: (1) copy the mastercard for as many entries as are required in the catalogue; (2) add headings; (3) check the accuracy of each record; (4) file; (5) check the filing. Regrettably, most of the librarian's skills have been taken out of this process: they are involved only in the process of checking. I am also saddened that the books have become mere bricks on a production line. Also, we have become very dependent on outside factors in the matter of book selection and cataloguing.

We keep a separate record of periodicals. Again, we have to cope with very large numbers: about 14,500 issues in 1985. The main problem is the assembling of single issues of the same periodical, which is done about every six months. This too is done by numbers, but this time by postal numbers: most titles have them, and errors are few. We try to keep close track of bibliographical changes: even very minor ones are noted, including errors in numbering. Three different registers are

used, all filed according to postal numbers, making for easy reference both to each other and to the periodical, and enabling us to work on all related files in one operation: (1) register of all issues received, plus a note of any bibliographical changes; (2) binding register; (3) file of mastercards - we catalogue periodicals ourselves. The only reason we keep these files separate is because of physical differences in size and format. The effectiveness of these tools in conjunction with the use of postal numbers went quite beyond our expectations: processing was speeded up, and at the same time bibliographic and physical control of periodicals were highly improved.

The procedures I have briefly described have actually meant drastic changes in almost all aspects of the work of the department, from selection right down to subject cataloguing, which did not exist before. As I mentioned above, the skill has been removed from much of our work, which is now reduced to mere routine in many areas. On the other hand, we have gained in speed and the amount of material acquired, especially periodicals. We have also been able to produce a Chinese subject catalogue, though this is not yet available to readers. And all without taking on extra staff. Furthermore, we have enhanced the bibliographic control of our holdings.

Whether all this has been worthwhile is still too early to say, though we now provide a better service. We can however be certain that it is only an intermediate stage: the next will be the computer. This will undoubtedly be in connexion with derived cataloguing, and I will even go so far as to say that this will be taken from our colleagues in Beijing.

## PROBLEMS IN CHINESE BIBLIOGRAPHIC AUTOMATION

by Howard Nelson (British Library).

### Introduction

This contribution to the EASL Bulletin arises from a series of seminars held in Scandinavia in 1983 and 1984: I am grateful to my hosts for giving me the impetus to get these thoughts down on paper. The article was written late in 1984, and while the general issues remain unchanged, some points of detail will be a little out of date. Only the section on RLIN has been deliberately revised in the light of subsequent developments, and I am grateful to my colleague Beth McKillop for her up-to-date account of the British Library's participation in the RLIN CLK programme. I shall try to deal generally with the question of library automation as it affects curators of collections of Chinese books. Most of the material will be intended to explain and clarify issues for sinological librarians who have little experience of computers; but if at the same time I am able to explain the problems of handling Chinese for the benefit of the systems analyst, so much the better.

Section 1 deals with some basic issues; it describes the context in which progress has been, and is being made; and indicates some of the basic dangers of adopting partial solutions. Section 2 gives more detailed consideration to the technical difficulties which lie in the way of the attempt to handle Chinese by computer, and to describe some of the solutions which have been reached; I shall deal in turn with the three areas into which computing is generally divided: input, processing, and output. In section 3, I want to turn away from the techniques involved in the machine handling of Chinese text, which are in themselves quite independent of any particular application, and concentrate more particularly on the library side. Finally, section 4 gives a brief account of the Chinese bibliographic systems which currently exist, and goes on to give some thought to the prospects for European cooperation in this field, and to draw the discussion as a whole to a conclusion.

### Section 1

#### Basic issues

The basic issues are these: (1) are there any advantages at all in automation in European sinological libraries? (2) if so, does it follow that the original Chinese script has to be included in the automated catalogue? (3) again if so, why is this difficult? In all that follows, it is taken for granted that a divided catalogue is nothing but a nuisance; that retrospective conversion of existing card catalogues is a practical impossibility; that automation is only worthwhile if it offers

advantages which outweigh the disadvantages of a divided catalogue; and that if the catalogue is to be divided at all, it should be done only once. It is better to stick with the present catalogue than to adopt a partial and temporary improvement.

If the computer is seen as no more than an up-to-date version of the card catalogue, it can probably be dismissed here and now as a waste of time, money, and effort. Most sinological libraries have found reasonably cost-effective ways of maintaining a local catalogue on cards, and the automation of purely local activities is unlikely to bring benefits commensurate with the costs. The advantages of automation must be seen in a wider context. Although we maintain reasonably efficient card catalogues, the fact remains that in so doing, we are duplicating effort on a very large scale. No survey has been made of the acquisition of Chinese books by European sinological libraries, but since the majority exist to support research in the humanities and social sciences, it is reasonable to suppose that the same books are being bought and catalogued by many different libraries. In so far as the computer offers the possibility of sharing data between libraries, two advantages immediately become evident: a single catalogue record can serve the needs of several libraries; and it may even become apparent that, with inter-library loan operating on a Europe-wide basis, fewer copies of the books themselves will be needed - provided the librarians know that they exist in other collections. Shared cataloguing and shared catalogue data are therefore the first advantages offered by automation: the second is the possibility of derived cataloguing. It would be foolish to pretend that derived cataloguing is a possibility at this moment, or that there will not be major problems of standardisation to be overcome when it does become a possibility: equally, it would be foolish to ignore the fact that the international use of MARC records produced by the various national libraries is as much a goal in the field of Chinese Japanese and Korean as it is elsewhere. While shared and derived cataloguing can be seen as two sides of the same coin, they may not be inseparable. In other words, the problems in deriving catalogue data from the country of origin may remain insuperable for some time to come; but this alone (provided the long-term goal is not lost sight of) may not be enough to negate the benefits of a shared catalogue between a number of European libraries.

A broadly similar answer applies to the question of whether the Chinese characters have to be included in a computerised catalogue record. Large and powerful computer cataloguing systems exist in plenty, and could have provided us with many of the benefits of automation some years ago, so long as we had been prepared to work in romanisation alone. It is common ground among sinologists and librarians that the romanised version of a Chinese text is intelligible to a certain point; and that given a limited context (e.g. a specific discipline, or a list of periodical titles) the level of ambiguity is sufficiently low to allow the user to reconstruct the original Chinese text with reasonable confidence. But the key word is ambiguity. No librarian in the humanities can ignore the fact that his vocabulary is virtually unlimited; and when it comes to Chinese names, from the whole span of Chinese history,

it is impossible to identify an individual from the romanised form of his name alone. (How many of us have not had to decline to search in a Chinese biographical dictionary for a name presented in the roman script alone?) In other words, adopting a limited solution - in this case accepting records in romanisation alone - is likely to create at least as many problems as it solves. While it would be possible to adopt a romanisation-only catalogue, and to create in addition a separate, manually maintained and photographically reproduced index of Chinese script data, this would be to disregard the long-term goal, the eventual possibility of using Chinese-script MARC records from the Far East. It seems worthwhile, at least for the time being, to hold out for a computer system which handles the Chinese script alongside the roman.

Finally, why is it difficult to handle Chinese characters by machine? It is only difficult in the context of predominantly roman-script libraries and computing systems. To some extent, we are paying the price for being first in the field: when information processing began in the west, no designer thought it necessary or worthwhile to go beyond the roman character set and the Arabic numerals. Even an extended roman set, with a wide range of diacritics, does not over-stretch the capacity of a set of 256 characters. Underlying this limitation is the fundamental architecture of the computer itself, which handles all data as a series of positive and negative electrical charges; mathematically, all data is represented in binary notation. The "bits" of positive and negative information are assembled into sequences of 8, called "bytes": the back of an envelope is sufficient to work out that there are 256 possible positive/negative sequences in a single byte. Such a limitation presents virtually no problem in the western world: but even elementary vocabularies in the Far East go far beyond this. 2-byte character representation would be grossly redundant in a predominantly western-language catalogue; but the capacity it offers for Far Eastern languages (i.e.  $256 \times 256 = 65,536$  possibilities) is a good deal more realistic. Seen from the other end, by Chinese and Japanese systems designers, 2-byte data-processing (not itself technically in the least difficult) is obviously necessary and justified, and the inclusion of the roman and other western character sets in a Chinese or Japanese set involves no more than a trivial extension of the basic capacity. Would that matters were so simple for us! I shall deal with the more detailed problems of handling Chinese characters by computer in Section 2. Here, it is sufficient to do no more than point to the problem of sheer size.

#### The context

Despite the difficulties, progress in Chinese bibliographic automation has been, and is being made. It is helpful to liken our present situation to that of the three worlds in the political and economic domain. In our field, the first world consists of the countries of the Far East, which have a self-evident need to create systems for bibliographic data processing in their own languages and versions of the Chinese/Japanese script, and relatively little interest in mutual cooperation. Japan and Taiwan already have effective systems in operation. The second world is North America, which has the resources, and the political interest, to

enhance existing bibliographic systems to encompass Chinese, Japanese, and Korean (increasingly known simply as "CJK") as well. This world has an overriding need to see the several countries of the Far East as a single entity, and to develop a system which takes an impartial view of all the incompatible standards emanating from the region. The Research Libraries Group (RLG), based at Stanford, California, was the first in the field to bring such an enhancement to fruition, and the On-line Computer Library Center (OCLC), at Dublin, Ohio, now offers an alternative. The rest of us - sinological libraries in Europe, Australia, and possibly those on the fringe of the Chinese world in Southeast Asia and Hong Kong - form the third world: we lack the impetus of the first, and the resources of the second. No single country in Europe could, or should, pursue an independent development in this field. Not only should we pool our knowledge and resources in Europe as a whole if we are to make any progress, but we should share what knowledge we gain with colleagues, especially in Australia, who find themselves in a situation very similar to our own.

#### The dangers of a partial solution

One or two points made above hint at the dangers inherent in taking too limited a view of what is a large and complex problem. The final danger is of being misled by the proponents of individual character-handling systems. Several exist already; some are being enthusiastically marketed; more will undoubtedly appear. Cataloguing itself requires a form of word-processing, certainly; but catalogue-production is a different story altogether. Whether we are thinking of an on-line access to a catalogue, or of batch output to paper or microfiche, the software required to turn raw MARC data - especially if it includes Chinese characters - is of a different order of magnitude from that required to drive a word-processing facility. Small-scale micro-computer systems exist to handle limited files of catalogue data, but in order effectively to fulfil the needs of European sinological libraries, the back-up of a mainframe computer is almost certainly required. The question we should ask the eager salesmen of Chinese word-processing systems, when they appear (as they will) on our library doorsteps, is: "Can your system also produce a catalogue? Can it sort and index very large files of highly complex data? And can it allow me and my colleagues in London, Paris, Berlin and elsewhere to maintain a shared database?" If the answer is no (as for the time being it must be), then what is on offer, however ingenious, is a partial solution, liable to create more problems than it appears to solve.

#### Section 2

##### Input

The problem here is that of assigning numerical values to each of the members of a character set - any character set - so that the text can be handled by the machine. A conventional typewriter keyboard is adequate



for the input of the roman script, and with the addition of an extra shift or two, it can cope with a wide range of diacritics and special characters if required. Chinese obviously presents a far greater problem. Nevertheless, the input of the Chinese script has attracted the initial and the greatest interest, and a wide range of ingenious solutions has been put forward. Indeed, in some ways the attention given to finding neat and economical ways of encoding the Chinese script has diverted attention from the other two areas of concern: even if processing is handled in a similarly economical way (which is not always the case), the need for an economical means of eventual output has all too frequently been disregarded altogether. As I pointed out at the end of Section 1, the danger of a partial solution is very great: a word-processing system which cannot produce a catalogue is useless in a library; similarly, an efficient and indeed a glamorous input system which cannot be linked to a cheap and reliable output system is a bad buy.

I shall make no attempt to describe all the various input systems which have been put forward, but limit myself to an account of the main types of approach that have been adopted. (Nor do I propose to go too deeply into the merits and demerits of each approach: this is a fascinating but largely subjective exercise which arouses deep passions among the various protagonists. The design of a better Chinese keyboard is something for sinological librarians to occupy their leisure moments with perhaps, but it should not, I think be done in library time.) Each approach, fundamentally, represents an attempt to encode the Chinese characters reasonably quickly and reasonably accurately. The word reasonably is important, because no solution as efficient as the roman script keyboard will ever be achieved for Chinese. Every system will have its disadvantages: each will represent a more or less satisfactory compromise between the various conflicting constraints. Nevertheless, progress has been made since the days of the Standard Telegraphic Code, which enabled telegrams to be sent in Chinese, but which required operators at each end (unless they became phenomenally skilled, as some of them are said to have done) to look up in a manual the code numbers for each character in a message. By and large, each system sets out to satisfy a group of requirements; each sets different priorities on speed, accuracy, ease of use, the need for trained personnel, and cost. Any choice one eventually makes will - or should - reflect one's own assessment of the priorities involved. Speed, for example, will matter very much to a commercial company in the Far East involved in Chinese typesetting by computer; but in a library context, it is unlikely to matter as much as other requirements.

Existing Chinese input methods can be divided into four (I mean here only those which, in the present state of the art, appear to have a chance of success): the whole character method; the character assembly method; the numerical analysis method; and the phonetic input method. (There are others, including automatic recognition of printed or handwritten characters; voice recognition; and even one device which links springs to a pen, and measures the varying tensions on the springs to identify the characters.)

The whole character method derives directly from the Chinese typewriter, in which each character has a unique position on the keyboard, (and in the case of the typewriter, a unique slug of metal which imprints the image). The size of the character set immediately demands a large keyboard: even if several - say up to 9 - characters share a single keytop, with a multiple shift system alongside, the physical size of the keyboard is always likely to present a problem. A compromise has to be reached between keeping the keyboard to a manageable size, and including all the characters the operator is likely to want. Considerations of keyboard layout, and of character frequency (obviously dependent on context), also come into play. One ingenious solution to the problem of size has been to lay out a grid of characters on a rotating drum, with a cursor able to move horizontally across the face of the drum. Electronic sensors read drum and cursor positions, giving a unique set of coordinates for each character. The prototype had some 4000 characters on the drum, arranged in pinyin order: it was fairly quick and easy to use, since it depended on knowledge which most western students of Chinese already have - namely the association of the character with its romanised form. Characters not on the drum were assigned theoretical coordinates, and had to be looked up in a code book (the same solution would have to be applied to characters not represented on the physical surface of any whole character keyboard).

Character assembly techniques range from stroke-by-stroke methods, up to the assembly of characters from varying numbers of components - and there are as many analyses of component elements as there are designers in the field. I describe here only two examples of this approach. In the first, two Australian researchers (note 1) began from the fact that all Chinese characters consist of combinations of five basic strokes, to which they gave the numbers 1 to 5. The code for each character is therefore the combination of numbers which reflect its stroke-writing sequence. However, this produces some very long and unwieldy strings. While remaining within the size and scope of a conventional roman keyboard, therefore, they have grouped around their basic-stroke keys a small number of standard elements which can be encoded to make their current claim for an average of 3.7 keystrokes per character.

At the other end of the same scale is the terminal developed in North America for the Research Libraries Group by the Transtech company (note 2). Based on a terminal originally developed in Taiwan, it is now extended to cope with Japanese and Korean as well as Chinese. The keyboard has 128 keys, the essence of the system being the recognition that characters are composed of elements rather than merely strokes. The elements chosen correspond with none of the traditional classifications, so a considerable amount of operator-training is obviously necessary, but again, economy and speed of operation are currently being claimed.

The numerical analysis method is most similar to the old 4-corner system, which as a dictionary look-up method must have baffled more students than this one. A group of scholars in Taiwan have devised the

"Three-corner coding method" (note 3), in which component shapes are assigned 2-digit numbers, from 01 to 99, allowing the characters to be defined in terms of the positions which the numbered elements occupy. However, the definition of the three corners of a Chinese character is a complex business, requiring an elaborate manual of instruction.

Finally, the phonetic input method, which is the latest arrival on the scene. Although at first sight the least promising, because it approaches the Chinese characters in their most ambiguous aspect, it nevertheless offers some of the most exciting prospects for the future. This is because in order to overcome the ambiguity, the machine itself, rather than the operator, has to be made to do as much of the disambiguation work as possible. Contrast the whole character approach, where the whole burden of character-identification falls on the operator; contrast the character-assembly and character-analysis techniques, where much mental work and skill is, once again required from the operator. The phonetic approach, at its best, requires the operator to do no more work than input what he already knows, leaving the machine to work out which character is intended. In a sense, the machine is programmed to do as far as possible what anyone listening to a Chinese speaker does: work out from the context what is meant. At its simplest, the input of a single romanised syllable displays on the screen a menu of characters, from which the operator selects the one he intends; at its most advanced the machine is loaded with a dictionary of terms (i.e. combinations of characters) against which it assesses the romanised words as they are input. Where the romanised term input can represent more than one Chinese-character term, these terms are ranked (within the machine) in a hierarchy of frequency or probability. The most likely term is automatically displayed; if it is wrong, the operator scrolls through the menu of terms until the right term is found. The pattern of development here is characteristic of that seen in other computer applications. In the early days, machine costs were avoided at the expense of a greater burden on the human operator; increasingly, the burden is transferred to the machine, where it quite properly belongs.

The question is frequently raised as to the value of the Chinese tones in the phonetic input method. It has been found in practice that context is a far more powerful tool for the disambiguation of Chinese utterances; and also that while westerners can convert rapidly and confidently to romanisation alone, they are generally far less accurate in assigning the tones. This is a consideration particularly relevant in library work, in which the vocabularily regularly used tends to be far removed from that of everyday speech.

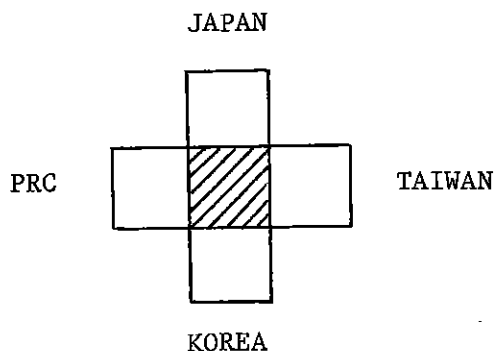
There is a host of Chinese input systems already available, and no doubt many more will appear in the immediate future. The choice appears bewildering, especially as each new system is offered in equally enthusiastic terms. Each, however, will suit different combinations of abilities and requirements. All that is really essential in putting any total system together, however, is that each element be independent of the rest. The input system chosen should be capable of being replaced by a better one, or a later-generation model, without the need to re-

structure the system as a whole.

### Processing

In Section 1, I outlined the principal difficulty in handling Chinese by computer, namely that of the sheer size of the character set. Here I go into more detail about the nature of the problem, and outline some of the approaches that have been made to its solution.

Although I have written so far as if Chinese were the only language of concern, it is time now to broaden the discussion, and to include the other two Far Eastern languages which use scripts based on Chinese. While from the point of view of the librarian only one of these three languages is likely to be a daily preoccupation, from the point of view of the systems analyst the three languages present different versions of the same problem: economy dictates that effort expended at this stage should be directed towards the problem as a whole, and not just to one part of it. A diagram drawn by Mr Takahashi Tokutaro of the National Diet Library (NDL), Tokyo, (though for a slightly different purpose) illustrates the point:



Each of the countries in this diagram uses for its own purposes a portion of the total, which is represented by the cross as a whole. It is obviously sensible for any country outside the Far East to see the complete picture, and not to adopt the linguistic viewpoint of any one country: to attempt a solution to the whole problem, rather than to one part of it. However, the impetus to produce standard codes for the internal representation of the script has, not suprisingly, come from the countries of the Far East themselves. Equally unsurprisingly, they have devoted their efforts to their own view of the character set, having a good deal more to gain from the solution to internal needs than from overwhelming evidence of international goodwill. One of the earliest sets produced was the Japanese Industrial Standard of January 1978, which in due course was extended by the NDL (in the form of the NDL-70 set) to cover their bibliographic needs. But the Japanese view of the character set is a good deal more limited than the Chinese, and up to this point at least, the NDL-70 list has not been extended to allow the NDL to process records for Chinese books.

Two rival sets of character codes exist for China: the set for Taiwan is called the "Chinese Character Code for Information Interchange" (note 4). It has been published in parts, starting in April 1980. The second edition was published in 2 volumes in November and December 1982. Broadly speaking, the set first arranges the 4,807 most common characters in radical/stroke order, and then follows them with a further 16,197 less common characters, and some 10,000 variant forms.

The PRC set is called "Code for Chinese Graphic Character Set for Information Interchange" (note 5). Only the first part has been published so far (in 1981); a second part is promised for late 1984. There are 3,755 common characters first, arranged in pinyin order, and some 6,000 further characters in (PRC) radical order.

Both Chinese sets have a similar structure: the characters are arranged in a notional framework of 94 planes, each of which has 94 sectors of 94 positions, giving a potential set of over 800,000 characters. 3 bytes are needed for this structure, the first identifying the plane, the second the section, and the third the position. The proponents of each set aspire to international standard status, and both stick scrupulously to existing ISO standards for information exchange (ISO 646 and 2022).

The Korean standard is called the "Korean Information Processing System Code".

It is obviously a matter of regret that in an area which need hardly be contentious, international cooperation has not been possible, and in particular that political rivalry between the two Chinas has been allowed to lead to duplication of effort on a vast scale, and to the production of mutually incompatible schemes. However, there was no point in waiting for such rivalries to sort themselves out, when in 1981/82 the Research Libraries Group found itself faced with the need to adopt a comprehensive character code for its own CJK enhancement. Boldly, they decided to accept the world as it is, and set out to design a system which would cope with all four national character sets. The result is the RLIN East Asian character Code (note 6) - in the words of one commentator, a de facto international standard. Briefly, RLIN took the largest set then available, compared it with next largest and added any characters not already found, and so on until they had built up a file containing every separate character in the four standards, with links to their codes in the standards of origin. The result is a superset, against which records from each of the countries concerned can henceforth be read, translated, and incorporated into the RLIN database. It is a very considerable achievement.

### Output

Output is the area where, in my view, most work still needs to be done. It is not surprising, though it has to be acknowledged as a major fault in systems design, that the quality and cheapness of the end product has been lost sight of in the search for solutions to the input and processing problems. Writing as a librarian, I take the view that Computer

Output Microform (COM) represents one of the greatest single advances in catalogue production we have seen. It is very cheap, very economical of space, very secure, easily amended, and reasonably easy to use. COM recorders with the definition required for Chinese exist, and once again, the technology exists to produce Chinese on COM; but it can so far only be done by adopting a different, and vastly more expensive, printing technique. Each character has to be separately drawn, and it is apparently impossible to combine standard roman-script techniques with graphic techniques, with the result that the cost of mixed-script COM production is based on the cost of the most expensive item in each run.

One alternative is output to a VDU, i.e. on-line access to a catalogue file, perhaps with a local printer attached to the screen. High quality graphics both on screen and printer certainly exist, but I think we should question very seriously whether on-line access alone (such as is offered by RLIN) is suitable for the sort of work we do. However skilled cataloguers and other library staff may become at working in on-line mode, it is likely to be a long time before readers are similarly adept. Catalogue production, it must be stressed, is quite as important as information retrieval.

Laser printers exist, and probably represent, for the time being, the strongest possibility for Chinese catalogue production; but so far, no system available in the west has linked Chinese catalogue files to a laser printer. It has been interesting to observe the development of the RLIN/LC system in this regard. RLIN's philosophy is on-line access only, whereas LC remains committed to the production of its traditional catalogue cards. For some time after LC had agreed to the link with RLIN, it had found no solution to this dilemma other than double cataloguing. The solution arrived at by LC at the time of writing is as follows: standard cards in romanisation only are ordered from Stanford (RLIN having departed thus far from its on-line philosophy), and to these are pinned their original character versions, obtained via the printer attached to the local terminals. The two versions are then sent for conventional printing in Tokyo.

One further aspect of the output problem is the storage of the character shapes themselves. I have said enough to make it clear that the input problem is a matter of assigning numerical values to the characters as economically as possible; and processing ceases to be a problem once the values are assigned, since the nature of the script remains transparent to the machine until the very last moment, when the numbers have to be decoded again. In the early days, much laborious work was done to draw each character on to a grid, which gave a series of coordinates for either dot- or vector-matrix reproduction. The storage space required for these "drawing instructions" was considerable. Nowadays, techniques have been developed which "digitise" character shapes automatically; they are used particularly in the typesetting field, and produce very high quality results. Microchip technology has advanced to the point where the vast quantity of data thus generated can be stored in ROM (read-only memory) on no more than two or three chips.

To summarise up to this point: not only have all the separate problems in the machine handling of Chinese been solved, but there is a ferment of activity in progress. New developments are constantly being made, and I am quite sure that we shall see advances in CJK data processing in the next five years quite as great as those we have seen in the past five.

### Section 3

What particular or additional problems arise in the attempt to apply the technology of machine handling of Chinese text to bibliographic work?

#### Strategic problems

First of all there is the strategic problem, particularly acute because we work in a minority area, of cost effectiveness. Should we, as sinological librarians, be seeking to enhance an existing bibliographic system, or should we look for an independent system catering precisely for the needs of Chinese, Japanese and Korean collections? To an extent this is an open question, but a first impression, when faced with the sheer computing power needed to turn raw MARC data into a catalogue, whether for on- or off-line access, must surely be that it would be foolish to think in terms of recreating all the necessary software. Certainly the present projects in the western world rely on existing software, to which CJK enhancements have been, or are being, added. On the other hand, experiences in Europe so far suggest that the needs of Chinese libraries, and users of Far Eastern books (indeed books written in any non-roman script), are accorded a very low priority, and are constantly being pushed to the back of all possible queues for systems analysts' and programmers' time. Almost any other enhancement to the existing system seems more important than CJK.

#### Definition of needs

A second strategic problem for sinological librarians is the decision as to what exactly our needs are. If a European CJK database is to be set up, do we want on- or off-line access? If it makes no sense - as I believe to be the case - for a single library to invest in a CJK system of any sort, how are we to share data quickly and effectively without an on-line view of the database? If however, we cannot afford the telecommunications costs - and in considering Europe as a whole we are talking about a fairly large potential area - how are we going to distribute catalogue data quickly enough to be of any practical use?

#### Information retrieval

Thirdly, what do we want in terms of information retrieval? It is fundamental that the Chinese characters have to be included in the record, but do we also need to find records by means of the characters?

All our card files depend on romanisation as the ordering and locating principle: the characters are there only to confirm the identity of the item already located by other means. Is such a system satisfactory for the age of automation? Too many automated systems have been no more than computerised card catalogues, and we surely don't want to fall into that trap. On the other hand, the Chinese themselves acknowledge that romanisation is a more effective means of ordering large files of characters (witness the printed catalogue cards of the National Library of China): why should we pay the extra cost in solving a problem the Chinese themselves do not perceive?

Bear in mind when thinking about this problem that if and when MARC records become available in the Far East, they will be written in characters, with little or no romanisation built in. The weight given to data in the original script in any European system must surely take into account the most obvious source of that data.

#### Word division

Fourth, word division. Whether or not one also approaches the files via the original characters, access by romanisation is, in the western context, essential. We cannot, as librarians, fudge or evade the issue of Chinese word division, even in a card catalogue, since standard filing rules take account of divisions into "words". (For the sake of this argument, I am treating letter-by-letter filing as non-standard.) You can solve the problem in a card catalogue by separating all the Chinese syllables (again, as on the National Library cards); although the characters are written separately, they are perceived by readers, western and Chinese alike, as grouping together to form semantic units similar to "words". There would be no difficulty if even the Chinese themselves agreed on the nature of the Chinese "word": but they don't, and it is my own belief that the current attempt to produce a set of rules for Chinese word-formation (being pursued both in Beijing, and by AFNOR in Paris) is doomed to failure, at least in so far as the rules, to be effective, will have to be simple, brief, and easy to remember.

Computers file, and computers index, and they do what they are told. You cannot expect a computer to produce Daqing or Huangming from its index if you ask it for Da Qing or Huang Ming, still less to file them together if it has been programmed to file word-by-word.

The approach adopted by the Research Libraries Group has been to produce its own set of guidelines for "Chinese syllable aggregation". Using the Wade-Giles system, which separates all syllables, the guidelines nevertheless introduce into both the vernacular and the romanised versions of the data a link between syllables. This serves to reduce the index to manageable proportions, so that "words" similar to western-language words, rather than single syllables (which are generally meaningless in modern Chinese in any case), are indexed. Unfortunately, it has proved impossible to reduce the RLG guidelines to anything less than a document some 10 pages long; to say the least, this is a lot for busy cataloguers to master, let alone readers, who have to apply the same rules in order



to retrieve Chinese material from the database.

#### Character set maintenance

Fifth, the need to keep the standard character set up to date. In a limited environment, it is at least theoretically possible to stay within a predetermined vocabulary; but this is not the case in library work. There will always be a need to cope with new characters, and provision must be made not only to add new characters to the repertoire, but for mechanism to ensure that all users of the system have the same updated vocabulary at the same time. The bigger the system gets, the greater the management problem involved: it is a problem no greater than other aspects of authority control, but it does imply the need for staff, over and above cataloguing and reader-service staff, to exercise the necessary control.

#### MARC format

Sixth, MARC format, and how to adapt it for mixed-script data. The various formats came into being a number of years ago, when the world was a simpler place. They are now groaning under the burden that is being placed upon them to cope with an ever wider range of material (music, maps, manuscripts, audio-visual materials), and those of us dealing with non-roman scripts meet with a very cold response when our special problems are raised. As always, we have the problem of cost-effectiveness: there is too little evident return to make the necessary investment of effort worthwhile. What, however, do we need? It is one of those problems, again, which depends on the ultimate use one wants to make of data: do we want or really need to process the original character part of the data? If we are content to automate the card catalogue - in other words rely on the romanised part of the record for all location and access work, and use the characters only to confirm the identity of material we have already found by means of the romanisation - then there is little actual need to alter any MARC format: the necessary characters can be held in the notes area, which already exists, and the problem is almost confined to its technical aspects. I say almost, because the MARC alteration required would be a provision for escape codes (i.e. coded signals to the computer of changes between character sets): this sounds simple enough, and in purely technical terms, it is. But since MARC exists to facilitate the universal sharing of bibliographic data, it follows that no change made in favour of one interest-group should unduly inconvenience the rest. The addition of escape codes implies a category of data which only some users need: it would probably be felt necessary for users of MARC records to suppress data introduced by an escape unless they could make use of it, and this inevitably costs money which non-CJK libraries would quite reasonably object to spending.

However, the principal disadvantage in adopting a basically roman entry in an otherwise standard MARC record is that it would make it impossible to incorporate character-based data from the countries of origin, thereby defeating one of the fundamental purposes of MARC.

The alternative, then, is to hold the original characters in the standard MARC field - which of course implies the possibility that they will have to be indexed - and to store the romanised version, which we can safely assume will still be required, in parallel or mirror-image fields. UK MARC, at least, does not at present allow certain basic fields to be repeated: books do not generally have two titles after all. The more you think about it, the more complicated it gets. I shall outline, for what it is worth, the solution adopted by RLIN. For their own internal purposes, they have allowed the repetition of basic fields, so that the vernacular and roman versions follow each other, in identical field numbers, down the record. RLIN users with CJK terminals will be able to see all parts of the record; others, who may still need to retrieve CJK records, will only see the romanised fields. For communication purposes, the record is translated back into a more standard format, and the vernacular data is transferred to repeats of a private field, with automatically supplied links to the roman fields with which it was originally associated. This technique allows the MARC format to be used without requiring major changes at the national level.

Both Japanese and Chinese MARC formats exist (the Chinese comes from Taiwan): each is based closely on UNIMARC, and each contains provision for the representation of some data in romanised form, generally subfields of the main field are used. Incidentally, escape codes are not used, because they are unnecessary: the roman script represents such a small addition to the Chinese or Japanese that it is simply incorporated - a nice insight into the way things might have been had computers been invented by users of almost any script other than the roman. But while the contents of Japanese and Chinese MARC are perfectly adequate for internal purposes, such records will not in themselves provide the mirror-image data which is likely to be needed in western libraries, even though they will provide a very sound basis.

#### Section 4

Having covered the technical problems in handling Chinese by computer, and some of the solutions that have been arrived at so far, and having pointed to some of the special problems which arise in the attempt to apply technology to the particular circumstances of work in a sino-logical library, I turn now to a brief account of the Chinese bibliographic systems which currently exist.

#### Cambridge University Library

The system I begin with is not strictly speaking an existing bibliographic system, since it has been abandoned; but since it was the first attempt in the western world, and since it attracted a good deal of attention at the time, it is worth a brief mention here. In retrospect, the story is not one of success, and has greatest value as a cautionary tale. In 1978, the British Library made a research grant to Cambridge

University Library to develop and assess a mixed-script bibliographic system, based on the whole-character drum device I described earlier, which had been invented shortly before in Cambridge. The project, though a brave attempt, was understaffed and too dependent on a single individual: even had he not left the library after only a year's work on the project, I wonder now whether he would have been able, virtually alone and with a wide range of other responsibilities, to tackle the whole range of problems I have already described. Ways were sought of continuing the work within the British Library after the formal withdrawal of the grant from Cambridge, and money being as tight then as it is now, the new design of the project involved paring it down to its barest possible minimum: a useful, possibly instructive, but in the end depressing and uncreative exercise. What we were basically trying to do, with insufficient funds and no full-time staff, was write the software necessary for a specialised peripheral (i.e. a Chinese character input device; microcomputer with disc drives holding the character drawing instructions and local software; VDU and printer), which would create data for processing on the British Library's mainframe computer, and then translate the resulting catalogue tapes back into the Chinese script for eventual output. No money at all was available for work on the main system, and we were forced to seek methods which overloaded the peripheral. The project was finally cancelled in the spring of 1982: I conclude that it was underfinanced and understaffed, and that it is not worth embarking on a similar project - indeed it would be positively dangerous to do so - unless adequate resources of staff, money and time were available and guaranteed from the outset. Equally, enormous developments have taken place since 1978, and anyone beginning afresh now would have infinitely greater experience and technical accomplishments to draw on.

#### National Diet Library, Japan

In Japan, the National Diet Library has been moving steadily into automation since 1971. Beginning with an index to the Diet debates, they eventually moved on by 1977 to the completion of a monograph-cataloguing system. Over 200,000 records have been written since then, with the database growing at some 60,000 entries per year. Also recently completed is a retrospective conversion exercise which takes the database back to include 1969 imprints. The Japanese MARC format, like the Chinese closely based on UNIMARC, was finalised in 1980, and distribution of Japan MARC tapes began in the spring of 1981. The National Diet Library expresses itself willing to exchange tapes with other agencies: they already incorporate the LC tapes into their own database.

Characters are input by means of a large keyboard, which has some 2,600 whole characters available. Characters not on the keyboard are entered either by numerical code, or by means of a combination of their "pattern" (40 geometric shapes have been identified) and their component elements (which have required the addition of 139 additional elements, not in themselves independent characters).

On-line input, editing, and information retrieval are part of the

system, as is output either to a line-printer or to catalogue cards. Interestingly, there is no mention in the literature currently available to me of COM output.

### China

Taiwan, similarly, has made great progress, and in a shorter time. The work on the character code which I have already described has been pursued in parallel with the development of the Chinese MARC format, the writing of an ISBD-based cataloguing code, and the installation and programming of the necessary equipment at the National Central Library in Taipei. A database is now in existence, on-line access exists, and I have no doubt that MARC tapes will shortly become available.

In passing, no similar achievements can yet be reported for the PRC, although a good deal of work is going on in addition to that on the character code, which has already been described. There is little doubt that an automated cataloguing system will be developed in Peking in the next year or two. Chinese librarians on recent visits to the west have been overwhelmingly interested in information retrieval systems, and the needs of China's developing industry for a scientific information service need no elaboration from me. Resources are scarce, however, and I suspect that Chinese librarians and information scientists are having to take their place in a queue for equipment and trained personnel.

### Research Libraries Group

Turning back to the western world, the Research Libraries Group (RLG), with headquarters at Stanford, California, supports the only CJK cataloguing system currently in existence. RLG is a cooperative library-development association which includes among its members most of the major academic libraries in the USA. One of RLG's activities is the maintenance of a shared on-line catalogue, called the Research Libraries Information Network (RLIN). Generous funds were made available some three or four years ago for the enhancement of the RLIN system to make it able to handle data in the Chinese script. Enhancement is the key word here: the CJK system is part of RLIN as a whole, and not separable from it. The CJK programs depend entirely on the processing software already driving RLIN, so there is no question of anyone being able to purchase the CJK system on its own. The system is also entirely on-line: it is no part of RLIN's policy to offer off-line input, or any sort of batch input. Use of the system is via RLIN's own terminals, and in the case of CJK participants, via the specially developed Transtech 4-language terminal, which has recently had a good deal of publicity.

RLIN made a considerable coup when the Library of Congress - not a full member of the corporation - announced that it would be cataloguing its CJK material into the RLIN system as soon as it was ready. The first terminals were installed in the summer of 1983, and the first record was input on 12th September that year.

What does RLIN development mean for the rest of us? There are three

possibilities: we can, either individually or collectively, join RLIN; we can wait until the RLIN software is made portable, and then license it and run it collectively in Europe as a parallel operation to the North American one; or we could invest in the Transtech terminals, and write a brand-new catalogue-support software package just for CJK use. I report here the outcome of a feasibility study conducted by the British Library during 1984 into the first of these possibilities. Telecommunications costs are not taken into account in the discussion which follows: we assume that costs will be incurred and met chiefly for non-CJK work. The cost of the dedicated transatlantic communications link is currently (1986/87) around £60,000 p.a.: this is met by the library primarily in support of a separate on-line catalogue, the Eighteenth Century Short-title Catalogue.

#### Financial considerations

Costs would be in the order of US\$50,000 for the purchase of the 4-terminal cluster, and US\$15,000 annually in maintenance charges.

#### Operational considerations

Disadvantages are mainly twofold: since RLIN carries out system maintenance and loading during Californian night time, the main RLIN database (which contains the CJK records) is unavailable for searching or input until 1300 London time. Since RLIN is an on-line system, with hard-copy output proving extremely laborious and time-consuming to obtain, the prospects of using the RLIN database as a current catalogue with access thus restricted are, for the time being, slight. If RLIN were to make the system available from say 1000 London time, the issue would take on a different complexion, and this has not been ruled out. Secondly, a major problem for the British Library in incorporating romanised RLIN CJK records into the BLAISE/LOCAS system is that the library's standard romanisation for Chinese is pinyin, making RLIN's Wade-Giles records impossible to incorporate without large-scale modifications.

Advantages clearly lie in reducing cataloguing costs as a result of using derived records. Preliminary experiments using the romanised-only records suggest that a very large proportion of the Chinese material awaiting cataloguing in the British Library has already been entered into the RLIN system, and that original cataloguing could be considerably reduced if the problems of access and output outlined above could be solved. For unfulfilled reader requests, RLIN offers a useful searching aid. Future developments planned by RLIN include the mounting of Japan MARC and a capacity for Hebrew and Arabic original script - facilities which no other single agency can envisage offering in the foreseeable future.

While recognising that the fundamental importance of the restricted catalogue access and problems with output make the present RLIN CJK system unattractive to the British Library's CJK sections at present, the exciting prospect of shared cataloguing and access to the collec-

tions of well-stocked American libraries for the purpose of obtaining photocopies or loans meant that when the Research Libraries Group generously offered to lend the British Library one CJK cluster, two terminals and a printer for one year, the library accepted, and its CJK staff are at the time of writing preparing to be trained in the use of the terminals. This will be the first time that a European East Asian library has participated in an original-script, on-line cataloguing system, and the British Library hopes that practical experimentation will be of benefit to the library and its readers alike.

### OCLC

Since this article was first drafted in 1984, OCLC's CJK cataloguing capability has become operational. Andrew H. Wang's paper on the current state of this project is therefore reprinted in its entirety immediately following my article.

### Prospects for European cooperation

If the eight British libraries with collections of Chinese books are understaffed and underfinanced, and yet are each doing original cataloguing for every Chinese and Japanese book they acquire, the same sorry picture is true of Europe as a whole. Recent discussions among European curators of sinological libraries have revealed no initiative towards automation, and indeed the same pattern of small poorly-staffed collections, each of which is attempting to do a very similar job. If one of the largest and most advanced libraries of Europe could not on its own afford the investment for the automation of Chinese, the same is almost bound to be true of other libraries. Collectively, we are caught in a form of poverty trap, where annual budgets for Chinese books are sometimes smaller than the cost of a single computer terminal. If progress is to be made therefore, it should probably be conceived on a European scale, no longer on the basis of a single country - let alone a single library.

I suggested earlier that we belong to a bibliographic third world, whose need can be broadly described as follows: we badly need to be able to take advantage of central cataloguing services, and in varying degrees we need union catalogues, which would enable us better to exploit the resources we have, and possibly to avoid some of the unnecessary duplication which goes on at present; we need to support, or have access to, a large CJK potential requirements file; we need a means of recording locations; and we need a cheap, off-line form of output. Insofar as libraries in Europe, Australia, and elsewhere have common needs, I should like to hope that from now on we may explore common solutions.

That was how the international picture appeared to me at the end of 1982, and I think the picture is broadly unchanged today. What has changed is the range of options that are open to us. In talking about RLIN, I described the results of our feasibility study into one of the three available options, namely that of joining the cooperative. The other two options have not been systematically explored, but preliminary

soundings suggest that licensing RLIN software to run a parallel operation in Europe has the virtue of a certain predictability: we should know in advance what we were letting ourselves in for; but it is certain to be expensive, especially if participation is limited to CJK libraries alone. Rewriting RLIN software from scratch, in order to exploit independently purchased Transtech (or other) terminals looks like a vastly expensive and time consuming exercise. Here, we should learn from the Cambridge exercise, where we began from a piece of specialised hardware, and set out to write the software to drive it, with sad consequences.

On the whole, then, despite some high hopes at the beginning, RLIN appears to offer little hope of progress for European sinological libraries. Nevertheless, I do not believe that there is reason to be unduly pessimistic. There is a ferment of activity on all aspects of the Chinese automation problem, and I am reasonably confident that a library application, at a price we can afford, will appear in the foreseeable future. The arrival of the OCLC CJK package introduces a new element, and it is worth watching closely; it certainly appears to be considerably cheaper.

Another possibility is that tailor-made library systems will come on to the market in the Far East: there is certainly a marketing potential in the bibliographic third world which I described just now.

Were the options limited to RLIN, I would have suggested that a feasibility study on the European scale might now be justified; but given the results of our own, and the current fluidity of the market, we should continue to wait until we can take advantage of other people's mistakes.

### Conclusion

This, then, leads into my conclusion. Despite a considerable eagerness to see progress made - I can hardly bear to think of the effort we are currently wasting on cataloguing Chinese books, every single one of which has been or is being catalogued elsewhere - and despite my impatience with the current state of sinological affairs in relation to the opportunities open to our western-language colleagues, I feel that we have more to gain at this moment from waiting than from rushing ahead. Meanwhile, we are going to be hearing more and more of developments in the technological field: more and more salesmen and sales literature are going to be reaching us, promising more and more speed and efficiency in Chinese word-processing. It will all be true, but my advice is: don't be taken in by a word of it. The computing power needed for catalogue production or information retrieval is vastly greater than that needed for word-processing or text analysis, and the key questions to put to any salesman who comes your way are: can it handle MARC, and can it produce a catalogue? If it can't, then it is of no use to a librarian.

## Notes

1. Shui-yin Lo and Zwi Barnea: Information processing of Chinese characters. In International cooperation in Chinese bibliographical automation (abbreviated hereafter to ICCBA), Canberra, 1983.
2. Russell G. Fischer: The CJK terminal, RGL and Transtech's achievement. In Library Hi-tech, Spring 1984, pp27-34.
3. Hu Li-ren: Zhong wen san jiao bian hao fa xun lian shou ce (Training manual for the three-corner coding method). Taipei, 1979.
4. Zhong wen zi xun jiao huan ma (Chinese character code for information interchange). 2nd. edn. Taipei, 1982. (2 vols).
5. Code of Chinese graphic character set for information interchange, primary set, GB 2312-80. Peking, 1981.
6. John Haeger: Overlapping standards, the RLIN East Asian character code. In ICCBA, pp.54-62.



## OCLC'S CATALOGUING CAPABILITY IN THE CHINESE, JAPANESE AND KOREAN LANGUAGES

by Andrew Wang (Manager, Online Systems Products & Services Department, OCLC Online Computer Library Center).

In the past decade or so, while more and more libraries in North America have participated enthusiastically in library automation activities, librarians of East Asian language materials have remained observers of this revolutionary event in librarianship. Many, if not all cataloguers of East Asian language materials in North America are still using the most primitive technique, namely handwriting, for the preparation of catalogue cards and other library records. It is not that the needs of East Asian language librarians have been ignored by automation planners; rather, it is that Chinese characters, though artistic in calligraphy and rich in meaning, are difficult to process by computers that can readily handle the Roman alphabet and Arabic numerals. This explains the gap between automation of Roman alphabet and non-Roman alphabet materials that exists in the OCLC Online System.

However, fifteen years after the OCLC Online System became operational in 1971, OCLC will close this gap by offering a cataloguing capability in Chinese, Japanese and Korean (CJK) characters in July, 1986. This is the first step OCLC will take towards the fulfillment of its commitment to the automation of non-Roman alphabet languages.

### OCLC's approach

OCLC has established four criteria for the automation of CJK and other non-Roman scripts:

- (1) The software must run on OCLC equipment. Users should not be required to purchase a different set of hardware for each non-Roman script. The CJK Workstation should also be capable of cataloguing in Roman alphabet languages, as well as performing word processing functions in both CJK and Roman-alphabet languages.
- (2) The software must be user-friendly. Extensive training should not be required.
- (3) The price must be reasonable and affordable.
- (4) The software must be designed in a modular fashion. The addition of other non-Roman scripts in the future should be possible without causing major difficulties.

### OCLC CJK packages

OCLC's CJK software consists of three packages: (1) CJK Online Cataloging Package; (2) CJK Card Production Package; (3) CJK Word Processing Package. Users may lease any or all of these packages to meet their needs.

- (1) CJK Online Cataloging Package enables users to interact with the OCLC's Online Cataloging Subsystem. This package supports the

- Chinese, Japanese and Korean scripts, as well as the Roman alphabet and Arabic numeric characters. In other words, users will be able to input and retrieve bibliographic records in CJK and/or Roman scripts into or from OCLC's Online Union Catalogue.
- (2) CJK Card Production Package can produce standard catalogue cards in CJK characters and the Roman alphabet. Bibliographic records catalogued through the OCLC Online System can be stored in the diskettes that reside in the CJK Workstation. Card production is controlled at each workstation in the library by an operator who may activate the card production function at the end of the working day to let the microcomputer direct card production activities after office hours. Using a Toshiba P351 printer, characters will be printed in 24x24 dot matrix.
  - (3) CJK Word Processing Package is a stand-alone package: it will not, nor does it need to interact with the OCLC Online System. This package combines three word-processing packages of Chinese, Japanese and Korean languages into one, and has all the features usually found in a well-designed word processor. It has, for instance, file creating, editing, deleting, merging, and saving functions. This package runs on the OCLC CJK Workstation, or an IBM PC if it contains a CJK Character Generating Board and a graphics board. This package can be used for printing bibliographies, correspondence, class notes, and so on.

#### Features of the CJK Online Cataloging Package

- (1) It runs on the OCLC CJK Workstation, which is a modified, general-purpose IBM personal computer (IBM PC). It does not require a dedicated workstation designed exclusively for CJK use. One CJK Workstation may be used for cataloguing English language or other Roman alphabet materials part of the time and CJK language materials at other times. CJK characters will display in 16x16 dot matrix on the screen.
- (2) It uses OCLC's standard English language keyboard. Therefore if one is familiar with the standard English language keyboard, no additional training in keyboard operation is necessary.
- (3) It provides five input methods; one is based on character composition and four on pronunciation:
  - (a) Tsang-chieh, a character-based input method that can be used for generating Chinese characters in both full and simplified forms, Japanese kanji and Korean hancha.
  - (b) Pinyin, a pronunciation-based input method that can be used for generating Chinese characters in both full and simplified forms.
  - (c) Wade-Giles, also a pronunciation-based method that can be used for generating Chinese characters in both full and simplified forms.
  - (d) Modified Hepburn, a pronunciation-based input method that can be used for generating Japanese kanji, katakana and hiragana.
  - (e) McCune-Reischauer, a pronunciation-based input method that can be used for generating Korean hancha and hangul.
- (4) For the purpose of retrieval, 15 search keys are provided as follows:

- (a) Numeric search keys:
  - LCCN
  - ISBN
  - ISSN
  - CODEN
  - OCLC Control Number
  - Government Document Number
  - Music Publisher Number
- (b) Derived search keys based on romanisation:
  - Title
  - Name/Title
  - Personal Name
  - Corporate Name
- (c) Derived search keys based on CJK characters:
  - Title
  - Name/Title
  - Personal Name
  - Corporate Name

In order to serve libraries world-wide, OCLC intends to provide romanised fields in records that contain CJK characters.

- (5) It adopts the RLIN East Asian Character Code (REACC), which consists of the following interchange codes:
  - (a) CCCII (Chinese Character Code for Information Exchange)
  - (b) CCGSII (Code of Chinese Graphic Character Set for Information Interchange)
  - (c) JIS (Japanese Industrial Standard)
  - (d) KIPS (Korean Information Processing System)

In building its machine-readable REACC, OCLC has made use of the CCCII tapes obtained from Taiwan.

#### Building a CJK database

Hardware, software, and a bibliographic database are three indispensable components of library automation. We have addressed some of the issues related to hardware and software. How does OCLC plan to build a CJK database? CJK records to be added to OCLC's database will come from four sources:

- (1) Library of Congress.
- (2) Research Libraries Information Network (RLIN).
- (3) National Central Library (Taiwan), and other libraries in the Asian/Pacific region.
- (4) Input or enhancement by OCLC member libraries.

#### Summary of Costs

The following costs are based on my recommendations to the OCLC Pricing Committee on 16th January 1986, and are subject to the Committee's approval:

- (1) Non-recurrent charges:
  - (a) CJK Workstation (includes first annual CJK System Support and Online Cataloging Package lease fee): US\$12,400
  - (b) Toshiba P351 printer, including parallel interface,

tractor feed, and parallel printer cable:	US\$1,400
(2) Recurrent annual charges:	
(a) CJK System Support and Online Cataloging Package lease fee (i.e. second and subsequent years, for each group of up to 5 units):	US\$2,000
(b) CJK Card Production Package lease fee:	US\$750
(c) CJK Word Processing Package lease fee:	US\$500
(d) OCLC system service fee:	US\$318
(e) Telecommunication costs (estimated):	US\$1,776
(f) CJK Workstation maintenance fee (estimated):	US\$912
(g) OCLC Modem, including maintenance:	US\$780
(h) FTU/FUP charges:	(as for Roman alphabet records)

Note:

- (1) CJK software lease fees include new versions of software if and when issued, without additional charges.
- (2) There is a discount of US\$1,000 on the second and US\$2,000 on the third and subsequent CJK workstations if purchased by the same institution within any 12-month period.

## REVIEWS

LIU (Guojun) & ZHENG (Rusi), tr. by ZHOU (Yicheng): The story of Chinese books. Foreign languages Press, Beijing, 1985. vi, 125pp. ISBN 0-8351-1304-3.

by Timothy Barrett (Faculty of Oriental Studies, Cambridge).

Some members of the Association may be lucky enough to possess Story of the Chinese book, the English translation of Liu Guojun's original Zhongguo shu de gushi, which was published by the Foreign Languages Press of Peking (as they then termed it) in 1958. This was a superb little introduction to the topic - something a sinological librarian could safely recommend to a maiden aunt to explain what the profession is all about without creating too much alarm and confusion. The appearance of a new, enlarged version of this survey can only be welcomed, even if the process of revision has not been entirely beneficial. There is, as the authors note, a great deal more information packed into the Story as retold, but although this is in some ways welcome (we are treated, for example, to a number of new photographs and illustrations), the throng of additional book titles quoted, always with English translations and inevitably with some fairly inept ones at times, in itself makes for a far less easy read. The much more cramped format of the new, narrower pages also renders the book physically less appealing than its predecessor, and the updating has not entirely eliminated some doubtful statements (such as the acceptance of a faked manuscript of the History of the Three Kingdoms as genuine on p.iii) nor added all relevant new information - no mention of the Korean printed dharani found in 1966 in the account of the origins of printing, for instance.

No matter: this work, after all, costs about as much as a pint of beer in Britain today, a fraction of the price of D.C.Twitchett's little volume reviewed in the last Bulletin. It is therefore still the ideal gift, if not to maiden aunts then at least to other librarians. The publication statistics on pp.119-122 in particular, already somewhat outdated as they are, should be underlined for the attention of colleagues in order to emphasize the vast scale of the problems with which the sinological librarian has to struggle. This is obviously not the ultimate reference work in the field, but it might prove to be a handy means of winning a few friends: something which most Chinese collections need rather badly.

MA (John T.): Chinese collections in western Europe, a survey of their technical and readers' service. Inter Documentation Company, Zug, 1985. (Bibliotheca asiatica, 18.) 90pp. ISBN 3-85750-064-6.

by Lutz Bieg (Berlin).

This long awaited "simple survey of Chinese collections in Western Europe" (p.1), contains short descriptions of the Chinese holdings in 60 libraries (not 56, as on p.3) in Austria (1), Belgium (3), Denmark (5), France (8), Federal Republic of Germany (15), Great Britain (11), Italy (5), Norway (2), Netherlands (1), Sweden (7), Switzerland (1), and the USSR (1). The author's "brief report" (p.1), whose contents range from the mere address (p.34) to nearly two pages of text (pp.75-76), is intended to serve as an introduction for researchers such as myself to libraries with which they are unfamiliar but which may contain materials relevant to their research.

Some limitations of this survey may be passed over briefly. Any user will immediately detect missing countries (Czechoslovakia, German Democratic Republic, Poland, Portugal, Spain, &c.) and places (Heidelberg, Münster, Tübingen, Würzburg). For West German libraries at least there is a broader and more detailed survey by Klaus A. Pretzell: Topographie asienskundlicher Schrifttumssammlungen in der Bundesrepublik Deutschland und Berlin (West), (Hamburg, 1978; reviewed by H. Walravens in Monumenta serica 34, 1979/80, 560-561). The descriptions of the various collections are rather loose and unindexed: apart from the "Table of Contents" at the beginning of the book, there is only an index of librarians at the end (pp.89-90).

On the matter of published catalogues, which the author discusses in the introduction (p.6), it should be pointed out that the catalogue of Chinese, Japanese and Korean books in the Bavarian State Library is not complete as suggested on p.45: by March 1986 only 3 of the proposed 10 volumes had been published. Also there is now a catalogue of the Far Eastern holdings of the Staatsbibliothek in Berlin: Staatsbibliothek Preussischer Kulturbesitz, Berlin, Katalog der Ostasienabteilung, hrsg. von Rainer Krempien, (Osnabrück, 1983-1985, 19 vols.). The Zeitschriftenkatalog der Ostasienabteilung (p.29) was updated in two supplements which appeared in 1981 and 1983. Werner Meissner's Katalog der Arbeitsstelle Politik Chinas und Ostasiens des Fachbereichs Politische Wissenschaften der Freien Universität Berlin cited on p.32 is in no way connected with the Bibliothek der Sinologischen Abteilung, the faculty under whose heading it appears, but describes the most important materials on Asia and China in the library of the Fachbereich Politische Wissenschaften, and is a useful tool for research into Chinese politics and contemporary China. The second volume of David Helliwell's Catalogue of the old Chinese books in the Bodleian Library appeared in 1985, and is a catalogue of the books sold to the library by Alexander Wylie. About one third of the contents of Paul Pelliot's unpublished Inventaire sommaire (p.71) has been made accessible through

Engels' an Marx 1844-83; 2c: Briefe an Dritte im Alphabet der Adressaten; 4: s.o.

In diesen Auflistungen sind - mit schöner Präzision gezählt - 5717 Uebersetzungen von 517 Werken (Briefe nicht gezählt?) aufgeführt. Und man wird die Ergebnisse im einzelnen nicht hoch genug schätzen, wenn man selbst einmal versucht hat, wie hier, auch Bruchstücke in Uebersetzung zu verzeichnen. Abgesehen einmal vom Suchfleiss steckt ein hohes Mass an Mühe allein in den Zuordnungen solcher Bruchstücke zum richtigen Werk. Erleichtert wurde diese Arbeit allerdings durch die weit fortgeschrittene chinesische Gesamtausgabe der MEW, zu deren Umkreis dieses Verzeichnis zu rechnen ist.

Vor allem löste sich damit eines der Hauptproblem ganz in einem chinesischen Sinne, es lagen einheitliche Sachtitel für jedes Werk in chinesischer Sprache vor. Der Bezug auf die chinesische MEGA einerseits und ihrer deutschen, russischen und japanischen Pendants andererseits bestimmt dann auch folgerichtig jeden einzelnen Eintrag zusammen mit kurzen Angaben zur Entstehung und der Ursprungssprache des Textes selbst.

Die Verzeichnung der Uebersetzungen geschieht dann jedoch - auf Kosten der Uebersichtlichkeit - in der Form fortlaufenden Textes, scheinbar ohne feststehende Regeln und - auch dies bedauerlich - nur in seltenen Fällen mit einem Hinweis auf die Ausgangssprache des Uebersetzers. Die seltenen Bemerkungen zum Umfang der Uebersetzung gegenüber dem Original gehen in einem laufenden Text dann natürlich unter. I.d.R. begnügte man sich mit der Angabe des Uebersetzers, des Uebersetzungstitels (soweit gegeben), der Fundstelle in knappester Form sowie etwaiger wiederholter Publikation wie die Uebernahme in MEGA (chin.).

Für den Bibliographen ist dies traurig, beschwerlich dann aber die Tatsache, dass auf den fast 1200 Seiten kein einziger Titel oder Name in der Originalsprache gegeben wird (2 Ausnahmen wurden gefunden!). Auch für chinesische Kollegen wird die Verknüpfung Beibeier - Bebel vielleicht noch angehen, aber andere? Unter bibliographischen Aspekten wird der Rückgriff auf die jeweilige MEGA (deutsch usw.) unvermeidlich sein. Wer bei seiner Recherche gar von einer der Originalsprachen ausgeht, steht vor dem schier unlösbaren Problem, den entsprechenden Eintrag zu finden. Aber wir wissen, dieses Erfordernis gehört noch nicht unbedingt zum Handwerk chinesischer Kollegen.

Wer sich trotz dieser leidigen Auslassungen in dieses Buch vertieft, wird für seine Mühen jedoch reichlich entlohnt und wäre es auch nur durch die Rezeptionsgeschichte der einzelnen Werke der Klassiker, die sich mit einem Blick erfassen lässt; ja, ein grossartiger Anknüpfungspunkt für weitere Arbeit ist, denn schon allein dabei ergeben sich sowohl bekannte aber eben auch neue Einblicke, die jedem Geistesgeschichtlicher wohl anstehen.

Auffallend schon die Zahl verschiedener Uebersetzungen desselben Werkes, nicht selten über einen kurzen Zeitraum publiziert. Allein die Kritik

des Gothaer Programms wurde zwischen 1922-30 in fünf verschiedenen Uebersetzungen vorgelegt, der Anti-Dühring 1930 einmal vollständig, einmal mit einem ersten Bande. Und selbstverständlich verwundert es vor dem historischen Hintergrund der Wirkungsgeschichte des Marxismus in China keineswegs, wenn auch immer wieder die gleichen Uebersetzer auftreten.

Sollten sie nicht voneinander gewusst haben? Unglaublich!

Cheng Hsueh-chia in einem kleinen Aufsatz (Issues and studies 4.1967, S.10) zum Thema konstatiert lediglich ein " ... the later translators are better ... but still left much to be desired." Man ziehe den umgekehrten Vergleich zum westlichen Sinologen und wird auf Nahe- liegendes stossen: Immerhin war nicht nur eine völlig neue Ideologie zu erwerben, anders als für Sozialdemokraten wie Bebel war für den chinesischen Sozialisten eine völlig neue Terminologie zu entwickeln, und was für eine!

Vor diesem Hintergrund allerdings treten die bibliographischen Mängel des Werkes weit hinter seinen Wert als Rüstzeug ideologiegeschichtlicher Forschung zurück.

Quanguo baokan dianying wenzhang mulu suoyin, 1949-1979, Beijing tushuguan shehui kexue cankaozu Zhongguo dianyingjia xiehui dianyingshi yanjiubu hobian. Zhongguo dianying chubanshe, Beijing, 1983. 10, 879pp.

von Horst Schäfer (Berlin).

Diese Bibliographie mit ihren Listen von etwa 30,000 Artikeln zum Thema Film aus den Jahren 1949-1979 darf zurecht als die Ergänzung zur "Entwicklungsgeschichte des Chinesischen Films", die 1963 vom Chinesischen Filmverband herausgegeben wurde, bezeichnet werden. Der Umfang dieses Buches mit seinen über 1.2 Millionen Schriftzeichen auf 879 Seiten zeigt, welchen Stellenwert der Film in der Volksrepublik China seit 1949 im Kulturellen und politischen Leben sowohl auf Seiten der Filmschaffenden als auch auf der Seite der Rezipienten und Kritiker besitzt.

Aus 62 Tageszeitungen und 108 Zeitschriften des gesamten Landes wurden über einen Zeitraum von 30 Jahren alle Artikel zum Thema Film gesammelt und in diesem Buch inhaltlich in drei Hauptabschnitten zusammengefasst: (A) Der chinesische Film (S.1-570); (B) Der ausländische Film (S.571-763); und (C) Filmtechnik (S.764-837). Auf den letzten Seiten des Buches findet man Artikel zu bestimmten Kritikkampagnen, die die chinesische Filmwelt betrafen.



Alle Hauptabschnitte gliedern sich durch Untertitel, die mit chinesischen Ziffern bezeichnet werden, und diese wiederum in noch speziellere Titel, die durch arabische Ziffern geordnet sind. Diesen kleinsten Unterteilungen sind die Artikel in der Reihenfolge, dass zuerst die Tageszeitungs-, danach die Zeitschriftenartikel jeweils chronologisch aufgelistet werden, zugeordnet. Alle Artikel sind mit vollem Titel, Verfasser, Zeitungs- bzw. Zeitschriftentname, Datum (bei Zeitungen in chinesischer Schreibweise Jahr:Monat:Tag; bei Zeitschriften durch Jahreszahl:Nummer der Ausgabe) und Seitenzahl, die durch einen Schrägstrich von den Zeitdaten getrennt wurde, versehen. Die Seitenangabe bezieht sich immer auf den Anfang des Artikels und lässt keinen Schluss auf seine gesamte Länge zu.

Der erste Hauptteil "Der chinesische Film" gliedert sich in neun Untertitel, die folgendermassen bezeichnet wurden: (a) Verwaltung der Filmwirtschaft; (b) Produktions- und Organisationseinheiten der Filmwirtschaft; (c) Uebersicht der Filmschaffenden (unter dieser Rubrik werden nach allgemeinen Artikeln zur Arbeit der Filmschaffenden im folgenden Artikel, die in ihrem Titel den Namen einer bestimmten Person beinhalten, nach dem Familiennamen geordnet; die Reihenfolge wird durch Strichzahl, der Reihe der Radikale und schliesslich den o.g. allgemeinen Kriterien, die diesem Buch zugrunde liegen, bestimmt); (d) Geschichte des chinesischen Films; (e) Theorie und Forschung; (f) Filmkritiken (auch in dieser Rubrik, die sich in Spielfilme, Opernverfilmungen, Bühnenkunstfilme, Animationsfilme, Dokumentarfilme und Wissenschaftliche Erziehungsfilme untergliedert, gelten die Ordnungskriterien, die schon bei den Artikeln über einzelne Personen genannt wurden); (g) Filmliteratur; (h) Vertrieb und Projektion; (i) Internationaler Filmaustausch.

Der zweite Teil "Der ausländische Film" enthält sechs Untertitel: (a) Die ausländische Filmwirtschaft (die Reihenfolge der Erdteile und Länder entspricht den Kriterien, die die Chinesische Nationalbibliothek für ihre Arbeit bestimmt hat); (b) Ausländische Filmschaffende; (c) Ausländische Filmtheorie und -forschung; (e) Ausländische Filme und Filmkritiken; (f) Ausländische Filmliteratur; (g) Internationaler Filmaustausch.

Der dritte Teil "Filmtechnik" unterteilt sich in folgende sechzehn Untertitel: (a) Allgemeine Kritiken; (b) Filmwissen; (c) Kameratechnik; (d) Spezialeffekte und Ausstattung; (e) Filmvertonung; (f) Montage; (g) Filmmaterial, Entwicklung und lichtempfindliche Chemie; (h) Projektion; (i) Studio- und Kinodesign; (k) Spezialfilme; (l) Zeichentrickfilme und Untertitelung; (m) Filmindustrie; (n) Fernsehen; (o) Technologische Innovationen; (p) Filmtechnische Entwicklungen; (q) Verschiedenes.

Im Anhang des Buches findet man Artikel zu folgenden Themen: (a) Anti-Rechts-Bewegung und Rehabilitationsberichte aus den Jahren 1957-58, die die Filmszene in Beijing, Shanghai, Changchun und im besonderen folgende Personen betrafen: Zhong Dianfei, Chen Renbing, Wu Yin, Wu Zuguang, Lü Ban, Sha Meng, Wu Yonggang, Shi Hui, Ding Ling, Chen Ming u.a.; (b) Kritik an dem Buch "Sammlung der Filmthesen" aus den Jahren 1965-66; (c)

Kritik an "Bringt neue Monologe", 1963-70; (d) Kritik an der "Entwicklungsgeschichte des chinesischen Films" von 1966-68 und (e) Kritiken aus der Zeit der Kulturrevolution an 42 Filmen, die z.T. in dieser Zeit entstanden sind oder zum Opfer neuen Kulturpolitik seit 1966 wurden.

Dieses Buch schliesst endlich die Lücke, die seit Jahren zwischen der "Entwicklungsgeschichte des chinesischen Films" und dem seit 1981 regelmässig erscheinenden "Jahrbuch des chinesischen Films", das das Filmschaffen der Volksrepublik China seit 1979 detailliert dokumentiert, aufklaffte. Die Hinweise auf 30,000 Artikel zum Filmschaffen dieser ersten 30 Jahre seit Gründung der Volksrepublik China sind nicht nur eine Goldgrube für einheimische und ausländische Beobachter dieses Massenkommunikationsmittels und dessen Einsatz im kulturpolitischen Leben seit 1949, sondern bedeuten auch eine Meisterleistung chinesischer Bibliothekare und der Angestellten des Chinesischen Filmverbands, die mit der Herausgabe dieser Bibliographie eine Basis geschaffen haben, die weltweit die Auseinandersetzung mit diesem Zweig neuester chinesischer Kultur vorantreiben wird.