

REPORT
on the conference, held at the Villa Serbelloni,
Bellagio (Como), Italy, May 30—June 4, 1965 on
THE USE AND APPLICATION OF
COMPUTERS AND COMPUTING
TECHNIQUES IN IRANIAN STUDIES

BY

M. J. DRESDEN

(University of Pennsylvania)

The purpose of this report is to provide a memorandum to the participants in the conference and other interested persons. It does not intend in any way to be a precise, detailed and complete account of the papers which were read, the discussions which were held or the suggestions which were put forward. No stenographic or other accurate record was kept of the proceedings and only in one or two cases was a copy of a paper presented to the conference made available to the rapporteur. In the absence of such data, the informal character of the meeting becomes of necessity reflected in the style and contents of this report. It will also appear that more attention is paid to the philological and linguistic than to the essentially technical aspects of the proceedings. This, of course, finds its explanation in the deficient competence regarding the latter of the rapporteur.

1. Participants in the conference

Dr. P. K. Anklesaria—The M. F. Cama Athornan Institute, Andheri, Bombay 58,
—INDIA

Dr. J. P. Asmussen [University of Copenhagen]—Østerbrogade 50 IV, Copenhagen
Ø—DENMARK

Dr. F. Bernhard [University of Frankfurt]—Indogermanisches Seminar, University
of Frankfurt, Frankfurt am Main—GERMANY

REPORT

Dr. M. J. Dresden [University of Pennsylvania]—Bennett Hall, University of Pennsylvania, Philadelphia, Pa. 19104—U. S. A.

Dr. J. Duchesne-Guillemin [University of Liège]—84, avenue de l'Observatoire, Liège—BELGIUM

Dr. W. Eilers [University of Würzburg] —Theodor Körnerstrasse 6, Würzburg —GERMANY

Dr. R. Emmerick [University of London] —St. John's College, Cambridge —ENGLAND

Dr. R. N. Frye [Harvard University]—Center for Middle Eastern Studies, 1737 Cambridge Street, Cambridge, Mass.—U. S. A.

Dr. A. Guttman [IBM Switzerland] —Nueschelerstrasse 24, Zürich—SWITZERLAND

Dr. O. Hansen [Free University of Berlin]—Riemeisterstrasse 1, Berlin—Zehlendorf—GERMANY

M. F. Kanga—D/10 Cusrow Baug, Bombay, Fort—INDIA

Dr. O. Klima—Oriental Institute of the Academy of Sciences, Prague University, Prague—CZECHOSLOVAKIA

Dr. G. Lazard [University of Paris]—École des Langues Orientales Vivantes, 2, rue de Lille, Paris—FRANCE

Dr. W. Lentz [University of Hamburg]—Eppendorfer Landesstrasse 120, Hamburg —GERMANY

Dr. D. N. MacKenzie [University of London]—31 Lansdowne Road, London N. 10—ENGLAND

J. Moyne [Harvard University]—IBM Data Systems Division, 545 Technology Square, Cambridge, Mass.—U. S. A.

Dr. Philip H. Smith, Jr. [IBM]—IBM Research Division, Yorktown Heights, N. Y.—U. S. A.

Dr. Z. Telegdi [University of Budapest]—University of Budapest, Budapest—HUNGARY

Out of these 18 participants, 4 were from Germany and the U. S. each, 2 came from the U. K., 2 from India, and 1 from Belgium, Czechoslovakia, Denmark,

REPORT

France, Hungary and Switzerland each. More relevant is the fact that 3 (Guttman, Moyne, Smith) were specialists in the various techniques of using computers; the other 15 were professional scholars of Iranian languages and literatures.

Invitations to attend the conference had been sent to other persons. For personal or other reasons they had been unable to accept.

2. Program

- May 31 9:30 R. N. Frye: Introduction
 10:30 Philip H. Smith, Jr.: "Concordance generators and dictionaries"
 4:30 A. Guttman: "Natural languages and artificial languages"
- June 1 9:30 F. Bernard: "Udānavarga concordance"
 4:30 continuation of same
- June 2 9:30 Constitution of several informal committees: 1. Pahlavi (MacKenzie, Eilers, Lentz, Asmussen); 2. Bibliography (MacKenzie, Lentz, Dresden); 3. Availability of computer facilities (Smith, Moyne); 4. Progress reports (Duchesne-Guillemin, Asmussen)
 10:30 John Moyne: "A Khotanese Saka concordance"
 4:30 discussion of several problems of organization
 5:00 O. Klima: "On the compilation of a Pahlavi dictionary"
 6:00 O. Hansen: "Oral explanation of the system of transliteration used in his textbook *Mittelpersisches Lesebuch* [Berlin, 1963]; a second volume containing a glossary and commentary is to follow"
 6:15 P. K. Anklesaria and M. F. Kanga: "Discussion of the urgent need for a Pahlavi dictionary"
- June 3 9:30 Z. Telegdi: "Remarks on the need for indices and the procedures to be followed"
 10:30 G. Lazard: "Remarks on the use of indices for Persian lexicography, grammar and literary studies"
 11:00 D. N. MacKenzie: "Remarks on the use of computers for the study of modern Iranian languages"

REPORT

- 4:30 Reports by the appointed committees (bibliography, work in progress, Pahlavi); adoption of a motion proposed by G. Lazard
- June 4 9:30 Philip H. Smith, Jr.: Report on available computer facilities
- 10:00 A. Guttman: Remarks on the organization of IBM and future computing equipment
- 2:00 Departure by bus to Gallarate for a visit to the Centro Automazione Analisi Linguistica
- 7:30 Arrival at the Milan station and dispersal of the group

So far for a more or less formal outline of the program.

3. Papers and Discussions

In a more detailed way the following selective notes seem pertinent.

The presentations by **Smith** and **Guttman** did much to acquaint the audience with the results that may be expected from and the problems connected with the use of computers for language study in general and for the making of concordances and indices in particular. One concrete example discussed by Smith was the project of Professor J. Raben (Queens College, Flushing, N. Y.) who used a computerized technique to determine the extent of Milton's artistic and intellectual influence on Shelley. (Raben discussed his work in "A computer-aided investigation of literary influence: Milton to Shelley," in *Literary Data Processing Conference Proceedings, September 9, 10, 11, 1964*, pp. 230-274; Smith contributed an article "A computer program to generate a text concordance" to the same volume (pp. 113-127) and wrote on the same topic, with K. F. Scharfenburg and R. D. Villani, in "A concordance generator" *IBM Systems Journal*, vol. 3, no. 1, 1964, pp. 104-111—*The Literary Data . . . Proceedings* volume is available from Jacques Duchesne-Guillemain, 84, ave. de l'Observatoire, Liège, Belgium.

A different approach is found in Sydney M. Lamb and Laura Gould, *Concordances from computers*, Mechanolinguistics Project University of California, Berkeley, California, 1964, with supplement of May, 1965)—**Guttman**, among other things, talked about such programming techniques as FORTRAN and COBOL. (Here again, to avoid duplication and save space, a reference to such programming manuals as Daniel D. McCracken, *A guide to FORTRAN programming* and *A guide*

REPORT

to *COBOL programming*, John Wiley and Sons, Inc., New York-London-Sydney, may be inserted).

The paper by **Bernhard** and the discussions it prompted in the morning and afternoon sessions of the second day did most to dispel the skepticism, if any such attitude was present, that may have existed in the minds of the Iranian philologists and linguists. Their question "What is the use of it all?" found an unambiguous answer. In a clear and convincing way Bernhard described the technique he had applied to the study of the approximately 1,000 Sanskrit fragments from Turfan that are part of the *Udānavarga* text. He outlined the technicalities of input, segmentation, sorting and output procedures by which he had succeeded in obtaining a word index, a reversed index and a verse index of these materials. Not only did the use of a computer result in a most remarkable saving of sheer time if compared with the traditional technique of "manual labor," it also provided clues to the solution of such problems as the proper location of fragments within the total text and the supplementation of broken fragments. In addition, the grammatical information which was collected made it possible to improve on F. Edgerton's *Buddhist Hybrid Grammar and Dictionary* (New Haven, 1953). Part of the concrete outcome of this project is Bernhard's *Udānavarga* I, Göttingen, Vandenhoeck and Ruprecht, 1965. Other projected research is a dictionary of all Turfan Sanskrit materials. Work on Tocharian is also anticipated. The response which Bernhard's presentation provoked led into a lively discussion of the feasibility of a similar technique to the Pahlavi materials.

The results on which **Moyne** reported in his paper on Khotanese, in the morning session of the third day, may be considered another eye opener. As in the case of the Turfan Sanskrit materials, a large amount of texts in Khotanese is at hand. Consequently, the demands on time, patience, accuracy and endurance involved in indexing by hand are very high. A relatively simple program makes it possible to instruct the computer to produce a concordance. Additional results as described for the *Udānavarga* fragments can also be obtained.

In the afternoon of the same day attention turned to Pahlavi. As was to be expected, **Klima's** paper as well as remarks made by **Hansen**, **Anklesaria**, **Kanga**

REPORT

and others showed that there is no lack of controversial and tricky problems. The most important of those lies in the nature of the Pahlavi script. Because of its deficiency, the script allows for a certain liberty when it comes to the interpretation of a single stroke or group of connected strokes. Another problem lies with the deficient character of many of the Pahlavi texts which have been published by means of Pahlavi types. As for the latter, it was generally agreed that the ideal basis of any attempt to make a concordance (or index) would have to be the manuscripts.

The problem of devising a satisfactory code as a guide for key-punching proved to be the most complex. And not without reason. In essence, it is the old crux of devising an objective transliteration system on a one-symbol-for-one symbol-only basis. It showed to have lost nothing of its intrinsic insolubility. Several solutions were proposed, discussed and rejected; none gained general acceptance. (Reference may be made to the study of Dr. Gernot Windfuhr of the Indogermanisches Seminar of the University of Kiel, June 1965, *Anweisung zur automatischen Sortierung des Buchpahlavi*, nach fachlicher Anleitung von Prof. Dr. W. Lentz, Hamburg, und mit technischen Hinweisen von cand. phil. Jürgensen, Kiel).

Another approach was put forward by which it was suggested to dissect the Pahlavi graphs and graphemes into an agreed number of basic strokes. The manuscript readings should then be objectively transposed to the punch-cards by means of specially designed keys or other such devices. The argument on these and related topics went back and forth for considerable time. It was continued in the afternoon of the next day.

In the morning of the fourth day, both **Telegdi** and **Lazard** gave exposés of the results which may be expected from the use of computers in the fields of Persian language and literature. It is not only for lexicography and grammar that such methods may be useful. Problems of authenticity and style might also be investigated and solved in a more complete and faster way than would otherwise be possible. Both lecturers were in agreement that a selective and experimental procedure has to be followed. Each project has to be carefully formulated. In the case of the Persian script, it seemed that a satisfactory code could be devised

REPORT

in a relatively simple way. After it was agreed that certain problems concerning other modern Iranian languages would also lend themselves to investigation by computing methods, a motion was unanimously adopted.

“The members of the international colloquium at the Villa Serbelloni [on June 3, 1965], having examined the possibilities of the application of modern techniques, used by electronic computers, to the study of the Persian language and Persian literature, consider that these techniques open a large and fertile field in this order of research.

They are of the opinion that they should proceed to experiments of a limited and selective nature in the immediate future.

They recommend that those who will be engaged in such research keep in touch with one another. For that purpose an information center will be established at the University of Liège, Belgium [c/o Professor J. Duchesne-Guillemin, Université de Liège, Liège, Belgique].”

In the afternoon, the bibliography committee reported on the existing card collection of Lentz in Hamburg, of Christensen in Copenhagen, of Dresden in Philadelphia and of Unvala in Bombay. All of these contain titles beyond the field of Iranian languages and linguistics. The Unvala collection is going to be published in the near future by the Parsi Punchayet in Bombay. Lentz made it clear that he would welcome publication of his collection after it has been checked for accuracy. No definite course of action was adopted. Asmussen, however, promised to look into a possible way of making these bibliographies available either in part or *in toto*.

As for reports on the progress of work in Iranian Studies, it was suggested that the desirability and possibility of regular statements in print would be taken up with J. de Menasce (Paris), who has previously shown interest in this matter, and with F. B. J. Kuiper (Leiden), one of the editors of the *Indo-Iranian Journal*.

Pahlavi was once more discussed. Two views came to stand out. The first, proposed and explained by **Lentz**, wanted to go about the matter in three steps: (1) “transposition” by which the Pahlavi symbols (or graphs) would be rendered by certain, previously agreed upon, symbols from the available keyboard for the

REPORT

purpose of coding and programming, in other words a sort of revised and improved Bartholomae system. This first step would lead to (2) transliteration, more or less in the manner of Nyberg's system, and to (3) transcription in the fashion of de Menasce and others. The second, put forward by MacKenzie, proposed to use the basic stroke approach by means of an improved keyboard. Naturally, no final agreement was reached. Yet, several of those present decided in favor of making individual attempts and trying several possible techniques on selected and limited sets of materials. In addition, they would keep each other informed of their results. It does not seem unlikely that further technical advances will facilitate the finding of a solution.

In the morning of the fifth and last day, **Smith** presented a list of computer facilities available for research. Computers are available at the Tata Institute in Bombay, the universities of Cambridge, London, Paris, Pennsylvania (Philadelphia), Prague, Budapest, in Darmstadt, Cambridge (U. S. A.), Yorktown Heights (New York); in Liège a computing center for research in the field of classical studies is already in operation: LASLA (Laboratoire d'analyse statistique des langues anciennes) under the direction of Prof. Louis Delatte.

In the concluding session, **Guttman** talked about the organization of the IBM Company and the assistance which can be given by its branch offices. In addition, he mentioned the fact that more advanced equipment is being designed and manufactured. A revised keyboard, with a much larger (one hundred or more) number of keys as compared to the 46 keys presently in use, will be available within the next one or two years. This more developed machine will simplify both coding and programming procedures.

In the afternoon, the group visited the Centro Automazione Analisi Linguistica in Gallarate where Father R. Busa S. J. explained the work being done by him and his collaborators on the Corpus Thomisticum and the Dead Sea Scrolls.

4. Results

In principle, the usefulness of computerized techniques for the study of Iranian languages and literatures seems beyond doubt. The case of the Khotanese concordance is both convincing and impressive and proves the point in a clear way.

REPORT

The Udānavarga study is additional proof, if any such evidence were needed, how much can be gained from a well conceived and clearly formulated use of the computer as a mechanical tool. Its capacity tends far beyond the resources of human memory and endurance as tools of data collecting. In cases where such labor can be saved the use of the computer is clearly indicated.

Organization and place of meeting

This report would yet be more incomplete if it did not include mention of the sponsors of the meeting and of the hospitality provided at the Villa Serbelloni. The warmest thanks of the participants go to the International Business Machine Company, to the Rockefeller Foundation and to Mr. and Mrs. J. Marshall, our hosts at the Villa. During our stay “humiles et sordidas curas aliis mandavimus et nos ipsos in alto isto pinguique secessu studiis asseruimus,” to modify Plinius’ phrase. It is too early to say whether “effinximus aliquid et excussimus quod sit perpetuo nostrum.” If we have, even in part, we should not forget our debt.

付 記

1964年秋以来私が師事してきたペンシルヴェニア大学の M. J. ドレスデン教授は、その堅実な学風をもって知られる著名なイラン学者である。特に Khotanese の分野では、謙遜な教授自らがほぼ完全なカードを所有していると言われるとおり、実質的に第一人者として、今後の活躍が期待されている。その学識の程は Khotanese の Jātaka-stava 訳註 (1955), Reader in modern Persian (1958), 発行間近い Dēnkart 新写本の出版, Handbuch der Orientalistik 中の Survey of the history of Iranian Studies (これは原稿未揃いのため未刊) に遺憾なく示されている。私個人としては、教授の学識そのものによりも、その真しな研究者としての生活態度と、その献身的な教師としての使命観に、より感銘深いものがある。私の場合、多くクラスは一対一で、自分の脇にすわらせて、Reichelt や Gershevitch を “Here!” と言っては、その太い指で押えつけるようにして説明されることで進行し、また雑談に花の咲くことも多かった。

このレポートは全く非公式なもので、教授が個人の意志で作成し、出席者と若干の友人に配布されたものである。それが単に新しいイラン学界の動向を伝えるものとして貴重かつ興味深いのみでなく、コンピューターの無制限かつ完べきな記憶力を用いた全く新しい方法論のゆえに、その分野外の諸氏の関心にも値いするものと思い、教授に本誌への掲載をお願いしたところ、快諾を得たものである。 (1966年5月24日 上岡弘二)