Pergamon Titles of Interest

Journals

Computer Languages
Computers & Data Bases
Computers & Education
Computer & Graphics
Computers & Human Concern
Computers & the Humanities
Computers & Urban Society

Books

Author-Subject Index to *Computers and the Humanities*

Computer-Assisted Research IN THE *Humanities*

A Directory of Scholars Active

Edited by Joseph Raben
Queens College
City University of New York

PERGAMON PRESS
New York / Toronto / Oxford / Sydney / Frankfurt / Paris

Pergamon Press Offices:

U.S.A.	Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford, New York 10523, U.S.A.
U.K.	Pergamon Press Ltd., Headington Hill Hall, Oxford OX3, OBW, England
CANADA	Pergamon of Canada, Ltd., 207 Queen's Quay West, Toronto 1, Canada
AUSTRALIA	Pergamon Press (Aust) Pty. Ltd., 19a Boundary Street, Rushcutters Bay, N.S.W. 2011, Australia
FRANCE	Pergamon Press SARL, 24 rue des Ecoles, 75240 Paris, Cedex 05, France
WEST GERMANY	Pergamon Press GmbH, 6242 Kronberg/Taunus, Frankfurt-am-Main, West Germany

Library of Congress Cataloging in Publication Data
Main entry under title:

Computer-assisted research in the humanities.

Includes index.
1. Humanities -- Data processing -- Directories.
I. Raben, Joseph.
AS8.C63 001.3'028'54 75-16447
ISBN 0-08-019870-8

Printed in the United States of America

Preface

This book is a response to a frequently expressed need for a single reference work describing the great variety of computer-assisted research in the humanities and the related social sciences. It is based on data collected semiannually between November 1966 and May 1972 and published in *Computer and the Humanities*. For this compilation, all original contributors were queried twice with the intent of obtaining the latest information on their projects. In the majority of entries, therefore, the information can be considered essentially current. Projects for which no new data could be obtained were reprinted verbatim from their appearance in the journal, but with the date of original publication indicated in brackets.

In a few instances contributors asked that, for any of several reasons, their entries be deleted. These requests were naturally honored, but for the sake of the book's completeness, the number and title of the project have been given, again with a reference to the date of original publication. When we have reason to think that the address given is no longer valid, but is a starting point for any attempt at communication with a project director, we have marked it with an asterisk.

Appropriately, this book has been produced by computer-driven photocomposition. Because the earlier entries in our data bank had not been made machine-readable, it was necessary to edit them for stylistic consistency, keyboard them, and edit them several times on a video display terminal. This task has been performed with professional and personal devotion by Marguerite Raben. Her task was facilitated by the unflagging cooperation of Laurence H. Heimrath, Solomon Gongola, Jonathan Marz, and Robin Warren. In other ways, the book has benefited from the efforts of Veena V. Paralkar and Lucie Tamassy. The index is based on one generated by Ellis Mount. The original idea of producing this compendium came from Theodore C. Hines.

Education

E1. Analysis of Language Teaching Texts

Principal investigator: Michael Mepham, Assistant Professor, Dept. of Linguistique, U. Laval, Cité Universitaire, Quebec 10, Canada. *Associates:* William F. Mackey, Lorne Laforge, Jean-Guy Savard.

Scope: Automatic processing of language teaching texts for the analysis of linguistic content and organization. *Method:* Application of dictionary lookup, concordances, and parsing schemes.

Type of computer: 1) and 2) IBM 360/50. *Size of storage:* 1) 512k, 2) 256k. *Language:* 1) PL/I e, 2) APL. *Disks or tapes:* 1) 9 2314 disks, 6 9-track tapes; 2) 9 2314, 30 2250 disks.

References: "L'Analyse des méthodes l'enseignement des langues à l'aide de l'ordinateur," *Sciences de l'Education pour l'ère nouvelle* 394 (July-December 1968). [*CHum,* May 1970]

E2. DOVACK, a FORTRAN Package to Support Individualization of Initial and Remedial Reading Instruction by the Language Experience Approach

Principal investigator: Florine L. Way. DOVACK Consulting Services, P.O.Box 3205, Tallahassee, FL 32303. *Associates:* William Petty, Duane A. Meeter.

Scope: For teaching initial and remedial reading to diverse populations on a broad scale, and for teaching standard English, the DOVACK System utilizes multisensory devices and the computer to individualize the language experience approach. Pupils dictate, from real and vicarious experiences, instructional materials which are transcribed by an aide for input to the computer. *Method:* The computer 1) processes each pupil's own lessons for his independent study with dictation equipment; 2) periodically obtains for each pupil from his own computer-stored word bank random sample vocabulary recognition tests; and 3) after every test, compiles and prints pupil-criterion-referenced reports including individual reports, class summaries and current norms, and graphs plotted with pupil IDs. To teach standard English for alternative situational use without disparaging the language and culture of the pupils, and to provide individually pertinent materials for making the transition to conventional reading matter, the teacher and others prepare third-person narrative versions of each pupil's dictations. Unique data-collection and compilation procedures, designed for process and product evaluation, can be used for broad-scale research. Workshops are necessary for the teachers.

Type of computer: CDC 6400. *Size of storage:* 65k 60 bit words. *Language:* FORTRAN IV. *Disks or tapes:* 1 CDC 6638 disk. *Processing:* Batch; independent of class time; average CP time on CDC 6400 one minute per pupil per month. Program is available.

Status: ESEA Title 3 funded a project to field test the DOVACK system in Monticello, Florida (OEG-4-8-060040-0069-056; $204,000; 1968-69–1970-71) for adaptability, effectiveness, and economic feasibility. The three-year report is available from ERIC. Consulting services for planning, workshops for teachers, and further information are obtainable from the investigator.

References: DOVACK: Method for Teaching Reading (End of Project Period Report), ERIC ED 071 268 (available from ERIC Documentation Reproduction Service, P.O.Drawer O, Bethesda, MD 20014); "The DOVACK Model," *EDUCOM Bulletin* 4, 5 (October 1969), 5-6; "The 'Language Experience' Approach in Teaching Reading—Computerized," *Computers and Automation* 19, 9 (September 1970), 28-31; *The DOVACK SYSTEM* (black and white 16mm 18-minute film, available on loan from Area Audio-Visual Center, Box 499, Monticello, FL 32344).

See also G64, L317, L326, L412, L458, L515, L531, L542, L551.

General

G1. Bibliographic On-Line Display (BOLD)

Principal investigator: Harold Borko, UCLA, Los Angeles, CA 90024. *Associate:* Howard P. Burnaugh.

Objective: Design of a large-scale, interactive, computer-based document storage and retrieval system. *Method:* Using both teletype and CRT display scopes in an on-line time-shared environment, the user is able to browse and/or search through a store of documents and have the abstracts displayed for his review.

Type of computer: AN/FSQ-32V. *Size of storage:* 64k. *Language:* JOVIAL (80 percent) scamp machine code (20 percent). *Disks or tapes:* Data Products disk, 4 million words; 12 IBM 729 IV tapes. Program is available.

Status: The BOLD system was terminated in 1968 after the principal investigator left SDC and the corporation acquired more modern computer systems. Much useful information was learned regarding user inquiries, displays, file organization and retrieval techniques. This knowledge was incorporated into the design of more advanced information storage and retrieval systems.

References: "Interactive Displays for Document Retrieval," *Information Display* 3, 5 (September-October 1966), 47-90.

G3. Research in Semiotic Systems

Principal investigators: E.D. Pendergraft and T.W. Ziehe, Senior Scientists, Tracor, Inc., 5500 Tracor Lane, Austin, TX 78721. *Associates:* N. Hirst, R. Jonas, C. Perkins, R. Schanks.

Objective: To design and implement semiotic systems that process both linguistic and non-linguistic signs for communication and control.

Type of computer: UNIVAC 1108.

Status: Terminated. [*CHum,* May 1967]

G4. Augmented Human Intellect Program

Principal investigator: D.C. Engelbart, Program Head, Systems Engineering Laboratory, Stanford Research Institute, 333 Ravenswood Avenue, Menlo Park, CA 94025. *Associates:* W.K. English, J.F. Rulifson, R.E. Hay, D.I. Andrews, J.D. Hopper, M.A. Watson, S.R. Levine, C.H. Dawson, B. Pollock, J. Stein.

Objective: To explore possibilities offered by direct, interactive computer aids for improving human intellectual effectiveness. *Method:* A predominantly empirical approach, in which the development of experimental computer aids is limited explicitly to those which can serve a small user community composed of the researchers themselves. Interactive procedures for composing, modifying, and studying specially structured working records on a computer-display console. Modifications of computer aids, information structures, working conventions and procedures, and concepts and skills studies are developed in a whole-system, evolutionary approach.

Type of computer: CDC 3100, shifting to SDS 940 Time-Sharing System. *Size of storage:* CDC 16k; SDS 65k. *Language:* CDC: locally developed machine-language assembler with associated CRT for on-line debugging. SDS: locally developed machine-oriented compiler, with on-line debugging facility. *Disks or tapes:* CDC: 3 high-speed tapes, 1 1311 disk file; SDS: 3 high-speed tapes, Bryant 4061 disk file. *Special features:* New type entry prepared on a paper-tape-punching typewriter and on-line CRT display. Program is available.

Status: Our own developments currently being used for drafting and updating (and publishing if appropriate) all of our on-going specifications, design, and reporting documentation work—even with limited computer availability of a single CRT station shared with other groups. *Future plans:* A time-shared computer system was to be installed in the summer of 1967, to support ten CRT stations, located in research offices, where the entire computer display system is devoted to the support of this exploratory work.

References: Available on request from D.C. Engelbart. [*CHum,* May 1967]

G5. Some Applications of Computers to Literary Stylistics, Textual Criticism, and Bibliographical Analysis

Principal investigators: Thomas Clayton, Professor, Dept. of English, U. of Minnesota, Minneapolis, MN; Ralph Cohen, Professor, Dept. of English, U. of Virginia, Charlottesville, VA.

Objective: The production of concordances and lexico-syntactical aids to literary study, and orthographical and paleographical aids to textual study.

Scope: Thomson's *Seasons* and *Castle of Indolence,* Cowper's *Task,* Suckling's non-dramatic works, Milton's *Paradise Lost* and portion of *The Books of Sir Thomas Moore* in Shakespeare's hand.

Type of computer: IBM 7094 and 1401. Work completed on IBM 360/75.

Status: Completed. Realized objectives included upper/lower-case concordance and related materials drawn from machine-readable tapes of Thomson's *Seasons* (concordance).

References: T. Clayton, "The Preparation of Literary Text for Milton's *Paradise Lost* (concordance), Suckling's non-dramatic works (concordance, etc.), and the "Shakespearean" addition in the MS *Booke of Sir Thomas Moore* (concordance, orthographical indexes, index litterarum, and edited Elizabethan and modern texts). Professor Cohen has printouts of the concordances of *Paradise Lost* and *The Seasons,* and Professor Clayton has printouts of the concordances and other materials drawn from *Paradise Lost,* Suckling's non-dramatic works, and the "Shakespearean" *More* addition. "Multiple Automated Studies: Comprehensive Identification and the Provision of Discriminants," *IBM Literary Data Processing Proceedings,* September 9-11, 1964 (New York: Modern Language Association, 1964), pp.172-199,+ Appendices A-E; "A Computer-Produced Aid to the Study of Shakespeare's Spelling and Handwriting," *Shakespeare Newsletter* 16 (February 1966), 6; "Computers in Humanistic Research: a Panel Discussion," *Computers in Humanistic Research,* ed. Edmund A. Bowles (Englewood Cliffs, NJ: Prentice-Hall, 1967), pp.245-248; *The "Shakespearean" Addition in the Booke of Sir Thomas Moore: Some Aids to Scholarly and Critical Shakespearean Studies,* Monograph No.1, Center for Shakespeare Studies, University of South Carolina (1969); T. Clayton, ed., *The Non-Dramatic works of Sir John Suckling* (Oxford: Clarendon Press, 1971); Ralph Cohen, *The Unfolding of the Seasons* (Los Angeles and Berkeley: University of California Press, and London: Routledge, 1970).

G6. A Computerized Bibliography in American Studies, 1865-1916

Principal investigator: Clarence C. Mondale, Associate Professor, American Civilization, George Washington U., Washington, DC 20006.

Objective: To supply a complete bibliography of published materials treating social themes for the period stated.

Type of computer: IBM 360/40.

Status: Now undertaking systems analysis and programming. *Future plans:* An extended index will supply clues for bibliographical search, when the bibliography is reasonably complete.

References: "A Bibliographical Survey of Economic and Political Writings, 1865-1900," *American Literature* 15 (January 1944), 381-410. [*CHum,* May 1967]

G7. Computer Editing and Publishing of Text [Terminated] [*CHum,* November 1967]

G8. SMART Automatic Text Processing and Retrieval System

Principal investigator: Gerard Salton, Professor, Dept. of Computer Science, Cornell U., Ithaca, NY 14850. *Associates:* Joel Zumoff, Barbara Galaska, Joyce Leslie, Ellen Lundell.

Objective: Design and implementation of a fully automatic information retrieval system, and investigation of techniques leading to an accurate and efficient information retrieval environment; currently using collections of articles in the fields of computer science, documentation, aerodynamics, medicine, and world affairs. Design and implementation of an on-line information retrieval system, and future library systems. *Method:* Fully automatic content analysis of documents, abstracts, and search requests. Automatic search and retrieval using user-controlled search strategies. Feedback studies to improve queries. Document space modification and cluster experiments to improve the data base. Automatic dictionary construction. Use of foreign-language texts. Term discrimination models.

Type of computer: IBM 360/65. *Size of storage:* 200-300k. *Language:* FORTRAN IV, PL/I, BAL. *Disks or tapes:* 1 2314 disk. Program is available.

References: Series of reports on "Information Storage and Retrieval" to the National Science Foundation, Nos. 1-20, 1962-1972, Harvard Computation Laboratory and Computer Science Dept., Cornell University; *Automatic Information Organization and Retrieval* (New York: McGraw-Hill, 1968); *The SMART Retrieval System* (Englewood Cliffs, NJ: Prentice-Hall, 1971); Joel Zumoff, *User's Manual for the SMART Information Retrieval System,* Department of Computer Science Technical Report 71-95, Cornell University, Ithaca, NY 14850.

G9. Journal Citation Reports

Principal investigator: Eugene Garfield, President, Institute for Scientific Information, 325 Chestnut Street, Philadelphia, PA 19106. *Associates:* Calvin Lee, Henry Small.

Scope: To study and evaluate scientific journals by determining the frequency with which journals cite one another. Applications are in the area of journal evaluation, library management, and science policy. *Method:* The *Science Citation Index* tapes are used as the data base. The frequency of citation from one journal to another is determined. Three major listings are produced: 1) a journal ranking list. Journals are ranked by times cited and by average times cited (impact factor); 2) detailed listings: a) cited journals listed for each citing journal and b) citing journals listed for each cited journal.

Type of computer: IBM 370/145. *Language:* BAL.

References: E. Garfield, "Citation Analysis as a Tool in Journal Evaluation," *Science,* 178 (4060) (1972), 471-479.

G10. Automatic Indexing; Development of Computer-Based Indexes, Bibliographies, and Catalogs; Various Research Applications in the Study of Subject Headings and the Structure of Index Entries

Principal investigator: Theodore C. Hines, Associate Professor, School of Library Service, Columbia U., New York, NY 10027. *Associates:* Maurice F. Tauber, Jessica L. Harris.

Objective: Development of groups of unit operations for the production of bibliographies, catalogs, and indexes; and related research into the nature of indexing systems. *Method:* In general, the technique is to combine theoretical and empirical techniques in the development of appropriate unit operations, and to use these both for production purposes and in implementing further research.

Type of computer: IBM 7090, 1620, 360/30, 360/40, 360/75. *Language:* SNOBOL, COBOL, and PL/I. Some of the programs are available.

References: Computer Filing of Index, Bibliographic, and Catalog Entries (Newark, NJ: Bro-Dart Foundation, 1966). [*CHum,* November 1967]

G11. Computerized Catalog of Medieval Manuscripts [Suspended] [*CHum,* November 1967]

G12. Alternate Methods of Machine Indexing and Their Effect on the Effectiveness of Document Retrieval

Principal investigator: Norman D. Peterson, Senior Systems Analyst, Branch of ADP Systems, Bonneville Power Administration, 3660 N.E. 133 Avenue, Portland, OR 97230. *Associates:* Dale L. Holmes, Val S. Lava.

Scope: Study the relative cost and effectiveness of indexing text of abstracts vs. titles only. Single keywording vs. multiple (phrase) keywording. Consideration of exclusion (drop) list vs. inclusion list in keywording. *Method:* Application of IBM software system no. 1401-CR-01X, as improved by Bonneville Power Administration, against a known data base of international power distribution literature in narrative form.

Type of computer: IBM 1401. *Size of storage:* 8k. *Language:* AUTOCODER. *Disks or tapes:* Minimum of 4 tapes (5 desirable). Program is available.

References: E.I. Bromberg, G.A. Dubinski, and N.D. Peterson, "The Bonneville Power Administration Selective Dissemination of Information Program," *Special Libraries* 58, 8 (October 1967); N.D. Peterson, "A System for Information Dissemination and Retrieval," *Computer Group News*, Institute of Electrical and Electronics Engineers, 1, 9 (November 1967); N.D. Peterson, "Notes on Computers and Legal Research," *The Practical Lawyer* (November 1967); E.I. Bromberg, Power Group Conference Paper No. 31 CP 67-513, The Institute of Electrical and Electronics Engineers, July 14, 1967. [*CHum,* November 1967]

G13. GRIPHOS General Retrieval and Information Processor for Humanities-Oriented Studies

Principal investigator: Jack Heller, Professor, Dept. of Computer Science, SUNY, Stony Brook, NY 11790.

Scope: A system within 360/OS which stores textual information in a data bank and retrieves and displays information selectively. Storage for 500,000 textual items of average length 5000 characters. *Method:* Card, typewriter terminal or terminal computer input. Textual information stored on disks and indexed via direct access disk storage.

Type of computer: IBM 370/155. *Size of storage:* 240k. *Language:* PL/1 F. *Disks or tapes:* 1 2314, 1 2321 disks; 2 tapes. Program is available.

Status: This system is being used operationally by the United Nations Library to produce indices of their published documents. The Museum of Modern Art in New York, the Arkansas Archeological Survey, and the Florida State Dept. of Archives and History are using this system to organize their data and produce indices and selected searches. The work is being carried on via the Museum Computer Network. (See V10b)

G14. Pattern Recognition and Formation by Machine

Principal investigator: Kenneth M. Sayre, Professor, Dept. of Philosophy, U. of Notre Dame, Notre Dame, IN 46556. *Associates:* Frederick Crosson, David Burrell, Vaughn McKim.

Scope: Handwritten characters. *Method:* Formal modeling.

Type of computer: UNIVAC 1107. *Language:* FORTRAN. Program is available.

References: K.M. Sayre and F. Crosson, eds., *The Modeling of Mind* (University of Notre Dame Press); K.M. Sayre, *Recognition; A Study in the Philosophy of Artificial Intelligence* (University of Notre Dame Press); F. Crosson and K.M. Sayre, eds., *Philosophy and Cybernetics* (University of Notre Dame Press).

Status: Completed. Final project report in *Pattern Recognition* (1973). [*CHum,* November 1967]

G15. Some Procedures and Programs for Processing Language Data

Principal investigators: Georgette M.T. Silva, S.T.F., Dept. of Linguistics, and Cliff J. Bellamy, Director, Computer Center, Monash U., Wellington Road, Clayton, Victoria, Australia 3168.

Scope: A reference manual for the humanist, containing information required to produce word-frequency counts, word indexes, concordances, etc. *Method:* The manual consists of an introductory chapter surveying the state of computer application to the solution of humanistic problems c. 1968, followed by several others dealing with methodology. The latter include detailed descriptions of procedures, flowcharts, program listings, deck structures, and user's instructions.

Type of computer: CDC 3200. *Size of storage:* 32k. *Language:* FORTRAN. *Disks or tapes:* 2 4 x 10 million-character disks (CDC), 4 tapes (IBM style). *Special equipment:* 30" incremental plotter. Program is available.

References: G.M.T. Silva, *Concordance Generation on the CDC 3200 Computer* (thesis, Monash University, 1966–expanded and revised version of concordance-generating procedures for the CDC 3200 in form of *User's Manual* available from Monash University Computing Center).

G16. Computer-Assisted Interrogation

Principal investigator: Charles T. Meadow, Senior Programmer, Advanced Development Dept., IBM Federal Systems Division, 18100 Frederick Pike, Gaithersburg, MD 20760.* *Associates:* D.W. Waugh, F.E. Miller.

Scope: Develop a language for use in man-machine communication including computer-assisted instruction, documentation, and generation of instruction courses. *Method:* A programming language based on PL/I has been developed which can be used for interrogating a responder to collect from him information not previously available to the computer. It can also be used for writing instructional courses and for assisting authors in the preparation and documentation of their courses.

Type of computer: IBM 360/40 or IBM 360/50. *Size of storage:* 256k. *Language:* PL/I. *Special equipment:* IBM 2260 display terminal.

References: "Computer-Assisted Interrogation," *AFIPS Fall Joint Computer Conference Proceedings* (Washington, DC: Spartan Books, 1966). [*CHum,* May 1968]

G17. Development of On-Line Text Handling Languages [Terminated] [*CHum,* May 1968]

G18. Development of Syntax-Driven Systems for Music and Text Analysis and Question-Answering

Principal investigator: Eors N. Ferentzy, Assistant Professor, Dept. of Computer Science, U. of Toronto, Toronto 5, Canada. *Associates:* T.R. Gabura, G.T. Gottshalks.

Scope: The first phase (a running syntax-driven system) is completed; additions planned during coming two years. *Method:* The Domolki method of syntax-directed compilation is used; the Robinson resolution principle is added for question-answering via automatic theorem-proving.

Type of computer: IBM 360/50. *Size of storage:* 120k bytes. *Language:* OS/360, PL/1. *Disks or tapes:* 2314 disk, 7- and 9-track tapes. Program will be available when completed. [*CHum,* May 1968]

G19. DYSTAL II: A General Purpose Programming Language in FORTRAN with Relocatable Arrays

Principal investigator: James M. Sakoda, Professor, Dept. of Sociology and Anthropology, Brown U., Providence, RI 02912.

Scope: Subroutines cover dynamic storage allocation, relocatable arrays, virtual memory, list-processing, string manipulation and pattern matching, sorting and ranking, statistical and matrix operations. *Method:* Basic FORTRAN IV subroutines and a disk file for secondary storage.

Type of computer: IBM 1130. *Size of storage:* 8000k. *Language:* FORTRAN IV. *Disks or tapes:* 1 disk drive, 2 magnetic tape units. Program is available. [*CHum,* May 1968]

G20. Spindex II (Selective Permutation Indexing)

Principal investigator: John J. Landers, Assistant Executive Director, Office of the Executive Director, National Archives & Records Service, Washington, DC 20408. *Associates:* Edward D. Thomas, Guthrie T. Meade, John Butler.

Objective: To provide a compatible computer system which will index existing archival and manuscript collection-finding aids and create new finding aids from raw data. A grant from the Council on Library Resources, Inc., facilitated development of the system. Programming, testing, and debugging of both a DOS and an OS system have been completed. Published documentation of the system will be available in 1973. *Method:* After examination of a cross section of archival and manuscript collection-finding aids, elements common to all the finding aids (i.e., dates, quantity, titles) were incorporated into the standard format structure of the system. After the standard format was devised, the programs were written, tested, and debugged. Considerable flexibility, both in terms of input method and data elements, and output report formats, was built into the system. The output programs can produce a variety of indexes and other listings specially tailored to the needs of each institution.

Type of computer: 1) IBM 360/30, 2) IBM 360/50, 3) IBM 370/155. *Size of storage:* 1) 32k, 2) 64k, 3) 80k. *Language:* 1) and 2), ALC-DOS, 3) ALC-OS. *Disks or tapes:* 1) 1 2311 disk, 4 9-track tapes; 2) and 3) 2 2311 disks, 4 9-track tapes. Program is available.

References: F.G. Burke, "Automation in Bibliographic Control of Archives and Manuscript Collections," *Bibliography and the Historian,* ed. Dagmar H. Perman (Clio Press, 1968); *Spindex II: Reports and System Documentation,* National Archives and Records Service (Washington, DC: Government Printing Office, 1973).

G21. Handwritten Recognition by Computers

Principal investigator: Raymond Deschenes, Director of Programming, Computing Center, U. of Miami, Box 8011, Coral Gables, FL 33124.

Objective: To have a computer recognize handwritten block letters. *Method:* A sample of 20 persons is asked to draw the alphabet in block letters on special forms, which are then keypunched and the information transferred onto tape. Each letter is considered to be located in a grid of 80 x 100 points. After the letters are normalized, this space is partitioned to form a vector. Various methods of partitioning this space are presently being studied to minimize the overlapping of the vectors describing each letter.

Type of computer: IBM 7040/1401. *Language:* FORTRAN IV and 7040 MAP. *Disks or tapes:* 10 tape drives, 7-track, IBM 729. Program is available. [*CHum,* November 1968]

G22. Recognition and Analysis of Content by Computer

Principal investigator: Howard P. Iker, Professor, Dept. of Psychiatry, U. of Rochester Medical Center, 260 Crittenden Blvd., Rochester, NY 14642.

Scope: A computer approach for the automatic elicitation, recognition, and analysis of content themes from materials in the English language. The prime factor differentiating this method from other content analytic systems is that the content categories and content themes are generated by the data themselves and are not imposed upon the data by use of predetermined encoding methods. This allows an empirical and general approach to the analysis of input. The system has been successfully applied to psychotherapy interview data, to books, and to projective text materials, and seems also applicable to analysis of textual cohesiveness and structure. *Method:* A word-based, continuity association approach capable of application to any form of natural language data. An input document is divided into segments (pages, number of words, minutes of time, etc.) and the frequency with which each word occurs in each segment is noted. Intercorrelations among and between words are then determined and subjected to factor-analysis. Past research demonstrates that the derived factors routinely represent and identify the major content themes in the raw data. All of the word-reduction and statistical analysis algorithms required to follow this approach have been implemented in a functional software package, WORDS.

Type of computer: IBM 360/65. *Size of storage:* 256k. *Language:* PL/I F level, Ver.5. *Disks or tapes:* 1 2314 disk. Program is available.

References: H.P. Iker and N.I. Harway, "A Computer Systems Approach to the Recognition and Analysis of Content," *Analysis of Communication Content,* ed. G. Gerbner et al (New York: Wiley, 1969); N.I. Harway and H.P. Iker, "Content Analysis and Psychotherapy," *Psychotherapy* 6 (1969), 97-104; H.P. Iker, *WORDS System Manual* (Rochester, NY, 1971). [*CHum,* May 1972]

G23. IMIJ/ Preparation of a Language to Construct Imagery Systems for the Production of Music, Graphics, and Sculpture

Principal investigator: John Stehura, Illiac Project, 1095 East Duane Avenue, Sunnyvale, CA 90486. *Associate:* Stephen Schleimer.

Scope: Production efforts are directed toward music making combined with motion picture animation. *Method:* Application of the IMIJ system involves writing statements, then listening to or viewing results. Specification of intermediate inputs may involve interactive control. The IMIJ compiler converts the statements into a secondary language which functions on a software defined machine. The IMIJ machine has a special set of operators for the preparation of one through four dimensional structures. The work originally concentrated on parameter-controlled systems for music and animation design; seven such systems were prepared and tested since 1961. The idea of an artificial intelligence or self-regulation design system was then developed; the film *Cibernetik 5.3* is the preliminary test of such a system. Work then evolved into a non-sequential, system-building language which integrates the previous methods.

Type of computer: IBM 360/75. *Size of storage:* 512k bytes. *Language:* PL/I F and META5. *Special equipment:* IBM 2282 optical scanner and camera recorder, 2250 display console, 7094 with digital to analog for audio output, and 1800 with automatic milling machine. Program is available.

Status: A system notebook regarding the design and integration of the IMIJ system was to be available by the end of 1973. The work is continuing with emphasis on hardware implementation using medium- and large-scale integrated circuit components; and using large-scale computation systems for system debugging.

References: *Cibernetik 5.3* (film), available from The Creative Film Society, 7237 Canby Avenue, Reseda, CA 91335; description of this work appears in "Expanded Cinema," *The Aesthetics of Machine Language* by Gene Youngblood (New York: E.P. Dutton, 1970).

G24. A Programming Language

Principal investigator: Professor Dr. A. Walther, Institut für praktische Mathematik der Technischen Hochschule, 61 Darmstadt, West Germany.[Deceased] [*CHum,* November 1968]

G25. Graphical Input/Output of Nonstandard Characters

Principal investigator: Susumu Kuno, Associate Professor, Dept. of Linguistics, Aiken Computation Laboratory, Harvard U., 33 Oxford St., Cambridge, MA 02138. *Associates:* Sheila Duncan, Hideyuki Hayashi, Tamotsu Mukaii.

Scope: This system was originally developed as an answer to the need for inputting and outputting Chinese characters for automatic processing, and has since been expanded to include other nonstandard orthographies as well. The system operates in two modes: new character entry and text inputting. A text may be prepared in Chinese, Japanese, Tamil, Pahlavic, or Hebrew, and stored on magnetic tape for future processing or printing. In addition to being alterable during its encoding process, the text file may also be edited after it has been stored on tape. The primary emphasis and goal has been on a flexible, rapid system which allows for the manipulation of nonstandard characters in their original orthography, eliminating cumbersome codes.

Type of computer: PDP-1. *Size of storage:* 24k. *Language:* Decal, a compiler. *Disks or tapes:* 1 drum, 2 magnetic tapes. *Special equipment:* Display scopes and RAND tablet, Stromberg-Carlson 4020 as output printing device. Program is available.

References: Susumu Kuno, Hideyuki Hayashi, and Sheila Duncan, "Graphical Input-Output of Non-standard Characters," *Communications of the ACM* 2, 9 (September 1968), 613-618. [*CHum,* May 1969]

G26. An Information Processing Laboratory for Education and Research in Library Science [Terminated] [*CHum,* May 1969]

G27. A Study of the Organization and Search of Bibliographic Holdings Records in On-line Computer Systems [Terminated] [*CHum,* May 1969]

G28. Generation of Data on Offset Masters Using Micromation Technology

Principal investigators: William A. Levy, Director Client Services; James A. Somervell, Director Programming and Systems Engineering, Advanced Research and Planning, Micromation Systems, Inc., 5218 Monroe Place, Hyattsville, MD 20781.

Objective: To develop and evaluate the feasibility of generating offset duplicating masters in various formats and styles using computer-generated microfilm as the principal medium. The end product will be a low-cost method of high-speed reproduction. The format and type style will be under computer control. *Method:* We are combining several major techniques from the disciplines of photo-optics, computer-generated data display, microphotography, and reprographics.

Type of computer: 1) IBM 360/30, 2) Stromberg Carlson 4360, 3) 3-M E.B.R. *Size of storage:* 65k. Program proprietary but information is available. [*CHum,* May 1969]

G29. On-line Computer-Based Systems for Information Management: The N.U.RIMS System (Northwestern University Remote Information Management System)

Principal investigators: G.K. Krulee, Professor, and B. Mittman, Associate Professor, Industrial Engineering/ Management Sciences Dept., Vogelback Computing Center, Northwestern U., 2129 Sheridan Road, Evanston, IL 60201.

Scope and method: This project is concerned with the development of on-line, computer-based information management systems. Primary consideration has been given to designing and implementing effective file management techniques, a user-oriented query language, a data structure which would accommodate both numeric and textual information, and a processor which could integrate all of these into a usable system. N.U.RIMS is designed to handle alphanumeric and numeric data, dates and numbers expressed in any consistent set of units. These data can be defined in a record structure and can then be input and searched. Textual material can be indexed and sorted. The search phase can be entered either on-line or in the batch processing mode.

Type of computer: CDC 6400. *Size of storage:* 65k. *Language:* FORTRAN IV. *Disks or tapes:* 2 854 disks, 1 6603 disk; 4 607 tapes. *Special equipment:* CDC 200 User Remote Terminal. Program is available.

Status: The major accomplishment to date has been the successful implementation of an information management system that meets certain design goals. These goals included: an internal file organization which would emphasize speed of search over speed of file creation and update, a language which could be learned easily by nonprogrammer personnel, a simple record structure, and an on-line search capability.

References: "TRIAL: An Information Retrieval System for Textual Information User's Manual," Vogelback Computing Center, Document No. NUCC 118, Northwestern University, December 1968; "INFOL for the CDC 6400—Reference Manual," Vogelback Computing Center, Document No. NUCC 152, Northwestern University, November 1968; "Development of a Remote Information Management System-RIMS," *Proceedings, American Society for Information Sciences* 6 (1969), 199-206; "Computer-Based Information Systems for University Research and Teaching," *Proceedings, 6th National Information Retrieval Colloquium,* 1969; L. Borman, R.Chalice, and D. Dillaman, "On-line File Interrogation Using the RIMS Remote Information Management System," report for the 1969 ASIS Annual Meeting, American Society for Information Sciences, October 1969; B. Mittman, R. Chalice, and D. Dillaman, "ASIS File

Management System Exercise: Report on RIMS," report for the 1969 ASIS Annual Meeting, American Society for Information Sciences, October 1969. [*CHum,* May 1970]

G30. CROSSPATE (Coordinative Retrieval of Selectively Sorted Permuted Analog-Title Entries)

Principal investigator: David G. Noel, Systems Librarian, University Library, U. of Western Australia, Nedlands, WA 6009, Australia.

Objective: Development of a suite of programs for automatic indexing. The programs are intended to accept uncontrolled English descriptions of the archives or objects being indexed. Every word in these analog titles is sorted, stoplist words are excluded, and all others are entered in a data file for subsequent retrieval by coordinative methods. Analog titles may be permuted en masse or after retrieval of a selected group.

Type of computer: PDP-6. *Size of storage:* 32k. *Language:* FORTRAN, MACRO. *Disks or tapes:* 2 interchangeable 8M character packs; 8 micro, 2 magnetic tapes. [*CHum,* May 1969]

G31. Hand-Printed Books (HPB Project)

Principal investigator: William J. Cameron, Dean, School of Library and Information Science, U. of Western Ontario, London, ON, Canada.

Objective: To establish data banks from which a universal bibliography may be developed, and from which subject bibliographies, chronological and other listings, and information of various kinds may be retrieved. *Method:* To establish bibliographical and computer-readable conventions that will permit a union catalog to be made that can be developed into the universal bibliography.

Type of computer: IBM 370/50. *Size of storage:* 128k. *Language:* COBOL E. *Disks or tapes:* Magnetic tapes. *Special equipment:* IBM 1403 N1 printer/1416 dual case print train; IBM 2740 typewriter terminal.

References: William J. Cameron and Brian J. McMullin, *The HPB Project: Phase I,* School of Library and Information Science, U. of Western Ontario, 1968; Phase II, U. of Western Ontario, 1970; W.J. Cameron, *A Perfectible Milton Bibliography,* 1972.

G32. Inquirer III

Principal investigator: Philip J. Stone, Professor, Dept. of Social Relations, Harvard U., Cambridge, MA 02138. *Associate:* Frances Morse, D.C. Dunphy, A.P. Coxson, A. Chalmers.

Objective: To construct, for as many words of English as possible, in order of decreasing frequency, dictionary entries containing sets of instructions which will enable a computer to guess reliably the meaning of an ambiguous token from features of its context, including inflectional endings, other words, and word-class markers assigned during sentence processing. *Method:* Routines are manually constructed from inspection of a concorded sample of approximately 500,000 words of "typical" behavioral science textual data. The dictionary will be used as a preprocessor for General Inquirer studies; it is processed by the tagging program.

Type of computer: IBM 360 series. *Size of storage:* 400k. *Language:* PL/I. Operational at Columbia, Edinburgh, Chicago, New South Wales, Chicago, and Harvard universities.

Status: Project completed.

References: The General Inquirer (Stone, Dunphy, Smith, Ogilvie); P.J. Stone et al, chapter in *The Analysis of Communication Content,* ed. Gerbner et al. (New York: Wiley, 1969); P.J. Stone, "Social Indicators Based on Communication Content," *FJCC Proceedings,* 1972.

G33. Statistical Analysis of Bibliographic Structures

Principal investigator: Alan Pritchard, Research Assistant, School of Librarianship, North-Western Polytechnic, 207-225 Essex Road, London, N.1, England.

Objective: Analysis of computer literature published in 1966. To help librarians, abstracting services, etc., and to provide basic data on the structure of a literature and the characteristics of abstracting services (overlap, coverage, delay, etc.). *Method:* Approximately 17,000 items coded with 25 parameters—country, language of publication, author, affiliation and publisher details, level, how published (journal, report, etc.), subjects. Analyzed by social survey program—MVC (Multiple Variate Counter) to produce counts and two-dimensional tables of one parameter against another.

Type of computer: Atlas. *Language:* MVC. MVC manual is available from London University, Institute of Computer Science.

References: Two progress reports—one presenting interim results; one giving details of computer analysis and parameters—are under project title by Pritchard. Available from DFSTI AS PB 180 391, or from author. Also "A Guide to Computer Literature" by A. Pritchard (London: Clive Bingley, 1969). [*CHum,* November 1969]

G34. MARC Record Handling System

Principal investigator: R. Bregzis, Associate Librarian (Bibliographic Systems and Technical Services), University of Toronto Library, U. of Toronto, Toronto 5, ON, Canada.

Objective: Local adaptation and integration of LC/MARC and BNB/MARC records into the technical processing procedures of the Library. Operational implementation of the MARC record-handling procedures in the cataloging departments of the University of Toronto Library. *Method:* Locally required MARC records will be identified and printed out from LC and BNB/MARC tapes. The printouts on structured data sheets will be reviewed against the books on hand, required modifications introduced and the resulting adapted records added to the Library's file of bibliographic records in machine-readable form. The adapted records will be used for current printout of catalog cards ready for filing in the various catalog files of the university library system. Listings of current acquisitions and retrospective catalogs to be developed later will also be supported by the adapted records in the Library's machine-readable bibliographic record file.

Type of computer: XDS SIGMA 7. *Size of storage:* 100k. *Language:* Metasymbol. [*CHum,* May 1970]

G36. Generalized I.R. as Applied to a Museum

Principal investigator: S.S. Harris, Assistant Director, Dept. of Sociology, W.H. Over Dakota Museum, U. of South Dakota, Vermillion, SD 57069. *Associates:* Michael Kelly, Lanny Hoffman.

Objective: To computerize the collections catalog of the museum to allow ease of printing of holdings and retrieval of classifications or individual items. *Method:* A three-phase method of total control over classification has been programmed. Logical questions (AND, OR, NOT) may be asked in an English language using classification terms, if desired. Special search algorithms are used.

Type of computer: 1) IBM 1620-I, 2) IBM 360/30. *Size of storage:* 1) 40k, 2) 32k. *Language:* 1) SPS III, 2) BAL. *Disks or tapes:* 1) 1 1311 disk, 2) 3 2311 disks, 2 2415 tapes. Program is available in SPS for the 1620, in BAL for the 360. [*CHum,* May 1970]

G37. General Purpose Information Storage and Retrieval

Principal investigator: William F. MacDonald, Design Specialist, Dept. 595-6, General Dynamics/Convair, P.O.Box 1128, San Diego, CA 92117.

Objective: To procure a mechanized system for information storage and retrieval that will not require complete new-generation reprogramming, one that is data-application independent, relatively computer independent, and easily transferable to new configurations of data processing equipment. The system should also be flexible enough to accommodate the diverse data bases of data management and link or correlate these data bases for both communications between the various data bases and people. *Method:* Continual daily operation of the system since 1963 over a variety of data bases from technical reliability data to biomedical data. The software package is a repertory of programs assembled in modular fashion according to the requirements of a given application. Three major considerations are to accommodate: 1) the data fields of the customer's data base in variable-length capability; 2) the overhead data of the customer in variable-length capability including all terminology and file structure control techniques; 3) the necessary software/hardware overhead to permit efficient manipulation of the data base. These three considerations are separable and individually manipulatable without undue disturbance to each other.

Type of computer: 1) IBM 360/30, 2) IBM 360/35, 3) CDC 6400. *Size of storage:* 1) 64k bytes, 2) 768k bytes, 3) 64k bytes. *Language:* 1) Assembly, 2) Mag. Sort, 3) in research. *Disks or tapes:* 1) 2 2311 disk packs; 1 7-track, 4 9-track tapes; 2) 2314 disk packs, 1 1-2311 disk pack; 33 9-track, 1 7-track tapes; 3) 7.5 x 10 million-word disk, 8 7-track tapes. *Special equipment:* Stromberg Carlson (SC) 4020 high-speed microfilm printer.

References: "Conversion of Large-Scale IS&R Systems for General-Purpose Operation," Convair Division of General Dynamics Report No. ERR-1301, U.S. Clearinghouse No.AD-680-962, December 1968; "Design Criteria for a Technical Information System Computer Software Repertory," Convair Division of General Dynamics Report No.ERR-AN-1131, U.S. Clearinghouse No.AD-834 659L, IDEP No.347.10.00-D5-01, December 1967; "Information Storage and Retrieval (IS&R) Used as an Analysis Tool," Convair Division of General Dynamics Report No.ERR-AN-1137, U.S. Clearinghouse No. AD-854 064L, December 1967. [*CHum,* May 1970]

G38. BALLOTS: *B*ibliographic *A*utomation of *L*arge *L*ibrary *O*perations using a *T*ime-*S*haring System

Principal investigators: David C. Weber, Director; Allen B. Veaner, Assistant Director; University Libraries, Stanford U., Stanford, CA 94305. *Associates:* A.H. Epstein, Donn Martin, L. Stovel, E. Montague, W. Davison, G. Cady, G. Chang, B. Moyer, R. Levitt, J. Schroeder, J. Powell, C. Will, J. Hartzell, R. Guertin, W. Kiefer, T. Martin, N. Roth.

Objective: Apply time-sharing, central computer facility plus CRT terminals in production environment to support library technical processing in a large research library, with hospitality to potential participants in regional network. Beyond technical processing, include public services, such as circulation, and through a sister project, SPIRES, public access to library files. Provide interactive searching of files and on-line diagnostics for new input to minimize errors; provide batch printed outputs. *Method:* Move from development to production operations in eleven modules, the first of which is now fully operational as a production system. Full range of technical processing services available from the first module, including file searching, input of new variables, change of existing values of data elements (deletion, insertion, addition), and all necessary printed outputs. Modules: 1) MARC, 2) in process file, 3) catalog data file, 4) purchase order and original cataloging, 5) inventory file of machine-readable catalog for undergraduate library, 6) non-purchase order, 7) production of book catalog for undergraduate library, 8) standing orders and O-Ps, 9) automatic claiming and canceling, 10) circulation, 11) reserve processing. All modules scheduled for completion by Fall 1974. Train department heads, supervisors, and employees; produce system reference handbooks and training aids.

Type of computer: 1) IBM 360/67, 2) PDP-11, 3) Sanders 804 programmable terminal. *Size of storage:* 1) 1250k, 2) 16k, 3) 4k. *Language:* 1) PL360, 2) Assembler, 3) Assembler. *Disks or tapes:* 1) CDC double density disks; 4 IBM 9-track tapes. *Special equipment:* Multidrop boxes to share communication lines 9600 baud. Program will be available upon full completion of project.

References: A.H. Epstein et al, "Bibliographic Automation of Large Library Operations Using a Time-Sharing System: Phase II, Part 1 (July 1970-June 1971)," Final Report, Project 7-1145, U.S. Office of Education, February 1972; A.H. Epstein and A.B. Veaner, "A User's View of BALLOTS," paper presented at 1972 Clinic on Library Applications of Data Processing (in press); Robert W. Burns, Jr., "Review of System Scope Document," *College and Research Libraries* 32 (May 1971), 236-238; K. Klinger, "Stanford Library Turns to Computer," *San Jose Mercury*, June 10, 1971; A.H. Epstein, "Information Flow Analysis," Canadian Society for Information Science, San Francisco Bay Area Chapter, Palo Alto, CA, November 1970.

G39. Automatic Informative Indexing, Abstracting, and Extracting

Principal investigator: Lois L. Earl, Senior Mathematical Analyst, Information Sciences Dept., Lockheed Missiles and Space Co., Inc., 3251 Hanover Street, Palo Alto, CA 94304. *Associates:* Martin A. Fischler, Oscar Firschein.

Objective: The development of algorithms for producing indexes, extracts, and abstracts, and the development of measures for evaluating these algorithms. *Method:* Investigation began with morphological studies which were preparatory to the current studies in 1) the utilization of syntactic information in automated extracting and indexing, 2) the coding of English word government patterns as an aid to syntactic and semantic analysis, 3) the development of techniques for describing pictorial structures.

Type of computer: IBM 360/40. *Size of storage:* 128k bytes. *Language:* Assembly. *Disks or tapes:* 4 2311 disks, 2 Model 2415 tapes. *Special equipment:* IBM and Sanders display units and a laboratory-developed optical scanner. Program is available.

References: "Structural Definition of Affixes from Multisyllable Words," and "Part-of-Speech Implications of Affixes," *Mechanical Translation* 9, 2 (June 1966); "Automatic Determination of Parts of Speech of English Words," *Mechanical Translation and Computational Linguistics* 10, 3a and 4 (September and December, 1967); M.A. Fischler, "Machine Perception and Description of Pictorial Data," *Proceedings of the International Joint Conference on Artificial Intelligence, May 1969;* "Experiments in Automatic Extracting and Indexing," *Information Storage and Retrieval* 6, 313-334; O. Firschein and M.A. Fischler, "Describing and Abstracting Pictorial Structures," *Pattern Recognition* 3, 4 (November 1971).

G40. DYSTAL II: Dynamic Storage Allocation Language in FORTRAN

Principal investigator: James M. Sakoda, Professor, Dept. of Sociology, Brown U., Providence, RI 02912.

Objective: Expands FORTRAN's capabilities by providing subroutines and functions for dynamic storage allocation, virtual memory, data structuring, list-processing, string-processing, sorting and ranking, statistical and matrix operators. For advanced FORTRAN programmers who desire to use its special features to write programs in such areas as text analysis, simulation, statistical packages. *Method:* DYSTAL consists of some 80-odd subroutines and functions written in Basic FORTRAN IV. A program written in DYSTAL consists of numerous calls to DYSTAL functions and subroutines intermixed with regular FORTRAN instructions. An early version of DYSTAL II was tested on an IBM 360/50 and a more recent version (1970) on an IBM 1130. Work is in progress on a two-byte integer version for smaller computers. Can be implemented on other machines.

Type of computer: 1) IBM 1130, 2) IBM 360/50. *Size of storage:* 1) 32k bytes, 2) 128k bytes. *Language:* 1) Basic FORTRAN IV, 2) E-level FORTRAN IV. *Disks or tapes:* 1) 2315 disk, 2) 2314 disk. Program is available. $15 for minitape, $5 for DYSTAL Manual.

References: DYSTAL II Manual, rev. 1970, Dept. of Sociology, Brown University; *Symbol Manipulation Languages and Techniques,* ed. J.G. Bobrow (North Holland: Amsterdam, 1968), pp. 302-311.

G41. Cost Reduction of Information Storage

Principal investigator: W.G. Hoyle, Senior Research Officer, Radio and Electrical Engineering Div., National Research Council, Montreal Road, Ottawa, ON K1A 0R9, Canada.

Scope: Cost reduction of information storage and retrieval including classification and indexing.

Type of computer: 1) and 2) IBM 360/67. *Size of storage:* 1) and 2) 1.5M. *Language:* 1) PL/I. *Special equipment:* Control units, card reader, punches, EAI 3500 data plotter, paper tape converter. Program is available.

References: "Automatic Classification and Indexing," *Mezhdunorodny Forum po Informatike*" (Moscow: Science Publishers, 1969, and The Hague: F.I.D., 1969); "On the Number of Categories for Classification," *Information Storage and Retrieval* 5 (1969), 1-6; "Automatic Indexing and Generation of Classification Systems by Algorithm," *Information Storage and Retrieval* 9 (1973), 233-242.

G42. APG: Ausgabe-Programm-Generator für nichtnumerische Daten

Principal investigators: F.Schulte-Tigges, Leiter der Abteilung Nichtnumerik; W. Schenkel, Doktor, Leiter der Gruppe Linguistik und Literaturwissenschaften, Deutsches Rechenzentrum, D-61 Darmstadt, Rheinstrasse 75, Bundesrepublik Deutschland. *Associate:* A. Mathe.

Scope: Allgemeines Programm-System für die Ausgabe nichtnumerischer Ergebnisdaten komplexer Struktur auf Schnelldrucker oder im konventionellen lochstreifengesteuerten Satz. *Method:* Problemnahe Definition der Datenstruktur in einer speziellen Sprache. Erzeugen eines geeigneten Programms mit Hilfe eines Programm-Generators.

Type of computer: 1) IBM 7094/I, 2) IBM 1401. *Size of storage:* 1) 32k, 2) 8k. *Language:* 1) FORTRAN IV, MAP. *Disks or tapes:* 1) 1 IBM 1301 disk, 13 IBM 729 IV tapes. Nach abschluss der Programmierarbeiten. [*CHum,* November 1970]

G43. Analyse und Auswertung kategorisierter nichtnumerischer Daten

Principal investigators: F.Schulte-Tigges, Leiter der Abteilung Nichtnumerik; W. Schenkel, Doktor, Leiter der Gruppe Linguistik und Literaturwissenschaften, Deutsches Rechenzentrum, D-61 Darmstadt, Rheinstrasse 75, Bundesrepublik Deutschland. *Associate:* U. Göbel.

Scope: Analyse und Auswertung lexikalischer Sammlungen. Voruntersuchung zu einem allgemeinen Analyseprogrammsystem für nichtnumerische Daten. *Method:* Problemnahe Definition der Datenstruktur in einer speziellen Sprache. Erzeugen eines geeigneten Programms mit Hilfe eines Programm-Generators.

Type of computer: 1) IBM 7094/I, 2) IBM 1401. *Size of storage:* 1) 32k, 2) 8k. *Language:* 1) FORTRAN IV, MAP. *Disks or tapes:* 1) 1 IBM 1301 disk, 13 IBM 729 IV tapes. Program is available. [*CHum,* November 1970]

G44. SORBAS-Sortieren von Bändern mit alphamerischen Strings

Principal investigators: F. Schulte-Tigges, Leiter der Abteilung Nichtnumerik; Th. Rathgeber, Leiter der Gruppe Informationswissenschaften, Deutsches Rechenzentrum, D-61 Darmstadt, Rheinstrasse 75, Bundesrepublik Deutschland.

Scope: Sortieren alphamerischer Strings nach mehreren Gesichtspunkten mit jeweils verschiedenen Ordnungsprinzipien. *Method:* Problemgerechte Definition des durchzuführenden Ordnungsvorgangs mit Hilfe einer speziellen Sprache. Prüfen der Definitionen auf Vollständigkeit und Widerspruchsfreiheit und Durchführung der Sortierung nach einem abgewandelten Mischverfahren.

Type of computer: 1) IBM 7094/I, 2) Telefunken TR440. *Size of storage:* 1) 32k, 2) 128k. *Language:* 1) FORTRAN IV, MAP, 2) FORTRAN IV, TEXAS. *Disks or tapes:* 1) 1 IBM 1301 disk, 8 IBM 729 IV tapes, 2) 6 MDS 252 tapes. Program is available.

Status: Completed 1970.

References: Th. Rathgeber, "Sortieren von Binärbändern mit alphanumerischen Strings," Programmsystem SORBAS, Programm-Information PI-41 des Deutschen Rechenzentrums, Darmstadt, 1970.

G45. A Pilot Project for the Computerized Storage and Retrieval of Theatrical Playbill Information

Principal investigators: Linn S. Kovar, graduate student; Fredrick M. Litto, Professor, U. de Sao Paulo, Escola de Communicasoes e Artes, Cidade Universitaria, Armando de Salles Oliveira, Sao Paulo, Brazil.

Scope: A pilot project for the creation of a system of FORTRAN programs and subprograms which will store, manipulate, and retrieve information from theatrical playbills, and will compile lists, indexes, and other aids for the program user. The system will be designed to handle many thousands of playbills; a trial series of about eighty playbills from Spencer Research Library, Kansas University, is now being used.

Type of computer: GE 635. *Size of storage:* 200k. *Language:* FORTRAN, GMAP, CONVER. *Disks or tapes:* 2 GE 270 disks, 3 7-track tapes.

Status: Project is terminated and available under same title as Masters thesis at the University of Kansas, 1970.

G46. PRECIS (Preserved Context Index System)

Principal investigator: Derek Austin, Project head, British National Bibliography, 7 & 9 Rathbone Street, London, W1P 2AL, England.

Scope: Computer printing of entries for alphabetical subject index. *Method:* Studies of relational analysis controlling citation order of terms in precoordinated input string; coding of terms as entry words, typographical style, etc.; computer rotation of string.

Type of computer: IBM 360/30. *Size of storage:* 64k. *Language:* Assembler. *Disks or tapes:* 4 2311 disks; 4 9-track, 1 7-track tapes.

References: PRECIS: A Manual of Concept Analysis and Subject Indexing (London: British National Bibliography, in press).

G47. Computer Typesetting Research Project

Principal investigators: C.J. Duncan, Director, Dept. of Photography and Visual Aids; L. Molyneux, Senior Lecturer, Dept. of Physics, U. of Newcastle upon Tyne, Newcastle, England. *Associate:* E.S. Page.

Objective: To produce general-purpose programs for processing natural language and symbolic texts through to graphic arts quality printed output (for Ministry of Technology).

Type of computer: 1) English Electric KDF9, 2) Digico MICRO 16. *Size of storage:* 1) 96k, 2) 16k. *Language:* 1) ALGOL, etc., 2) Machine Code. *Disks or tapes:* 1) 4 tapes, 2) 64k drums. *Special equipment:* LINOTRON 505 Reproducer/various converters, etc.

Status: Project is now terminated except for information exchange. The final report, entitled "Computer Composing" by C.J. Duncan, is available from Her Majesty's Stationery Office or the Department of Trade and Industry. Internal Notes Nos. 1-87 are also available at cost in 3 volumes from the Department of Trade and Industry.

G48. United Nations General Assembly Bibliography

Principal investigators: Jack Heller, Professor, Dept. of Computer Science, SUNY, Stony Brook, NY 11790; G.S. Martini, United Nations Chief of Indexing, United Nations Library, New York, NY 10017.

Objective: Construct a direct access data base of the bibliographic information of the documentation of the United Nations General Assembly. Produce permuted indices, selective searches and camera-ready printouts in English, French, Russian, and Spanish. *Method:* The GRIPHOS 06 (*G*eneral *R*etrieval and *I*nformation *P*rocessor for *H*umanities-*O*riented *S*tudies) programs are used to store, edit, and organize the input data (see G13).

Type of computer: IBM 370/155. *Size of storage:* 512k. *Language:* PL/1 F. *Disks or tapes:* 3 2314 disks, 3 60 KC tapes.

G49. Direct Access Data Base of MARC II Data [Deleted] [*CHum*, November 1970]

G50. Braille Translation and Production on a Small Computer

Principal investigator: Hans Karlgren, Linguistic Consultant, Research Group for Quantitative Linguistics, Södermalmstorg 8, S-116 45 Stockholm, Sweden.

Scope: The project consists of three phases: 1) To revise the existing Braille abbreviation scheme for Swedish, making a simple, consistent scheme which is easy to read and learn, and which abbreviates text by 20 to 30 percent; 2) to develop a computer program which can translate Swedish text into the new version of Braille on a small, inexpensive computer; 3) to develop a complete Braille production system using a computer, the program, and a special output

device. *Method:* 1) Statistics on frequencies of words and parts of words are used to select strings which should be abbreviated, considering such linguistic factors as syllable boundaries, ease of detecting affix abbreviations, etc., using a computer to help evaluate tentative versions of the revision; 2) program must be able to detect boundaries between the parts of compound words (very frequent in Swedish), syllable boundaries, etc., and perform translation on word-to-word basis, disregarding context.

Type of computer: 1) TRASK, 2) MINITRASK (P100). *Size of storage:* 1) 2k internal core, 30k external core; 2) 6k 20-bit words. *Language:* 1) ALGOL 60, 2) Assembly. *Special equipment:* Zoltan Braille Embosser for on-line high-volume production (25,000 page/hr).

References: Roger House, "A Tentative Approach to Braille Translation by Machine," *KVAL PM* 481; "A Proposal for the Development of a Computerized System for Swedish Braille," *KVAL PM* 483; "Survey of Braille Production Devices with Emphasis on the Planned Zoltan Braille Embosser," *KVAL PM* 485; "Computerized Braille Production in England, Germany, and Denmark," *KVAL PM* 499; "Description of KVAPT2," *KVAL PM* 513; "Computerized Braille Translation and Production in Sweden," *KVAL PM* 517; Märit Magnusson, "Description of KVA PT3," *KVAL PM* 631.

G51. Catalog of the Museum of the American Indian

Principal investigators: Jack Heller, Professor, Dept. of Computer Science, SUNY, Stony Brook, NY 11790. *Associate:* Frederich Dochsteder.

Objective: To construct a computerized data base of the information from the card catalog of the Museum of the American Indian. Using this data base, produce permuted indices and a new card catalog. *Method:* The GRIPHOS 06 (*G*eneral *R*etrieval and *I*nformation *P*rocessor for *H*umanities-*O*riented *S*tudies) programs are used to store, edit, and organize the input data (see G13).

Type of computer: IBM 370/155. *Size of storage:* 512k. *Language:* PL/I F. *Disks or tapes:* 2314 disk, 3 60 KC tapes.

G52. 1) Computer-Based Indexing (=COIN), 2) Computer-Readable Catalog (=CORC)

Principal investigators: I.K. Ravichandra Rao, Programmer; A. Neelameghan, Professor, Head, Documentation Research and Training Centre, Indian Statistical Institute, 112 Cross Road 11, Malleswaram, Bangalore 560003, India.

Objective: 1) COIN System to enable readers to find documents by giving keywords of titles only; 2) CORC System to arrange entries on the magnetic tape by their class numbers by using mixed notation and special characters.

Type of computer: 1) IBM 1401, 2) ICL 1900 series. *Size of storage:* 8k for both. *Language:* 1) AUTOCODER, low level, 2) PLAN, low level. *Disks or tapes:* 1) 4 7730 tapes, 2) 4 tapes. *Special equipment:* Console display. Program is available.

References: A. Neelameghan and S. Venkataraman, "Catalog on Tape," *Library Science* 5 (1968), 317-330.

G53. Output of Non-Standard Characters Using an SD4020

Principal investigator: Susan M. Hockey, Atlas Computer Laboratory, Science Research Council, Chilton, Didcot, Berkshire, England. *Associate:* Alan Jones.

Objective: To output texts and concordances written in nonstandard characters on microfilm. Character sets now available include Greek, Arabic, Persian, Hebrew, Armenian, and Cyrillic. *Method:* Design characters using PDP-15 and VT15 display. Write programs to edit text or output from the concordance program COCOA into a form suitable for use with the graphic output package GROATS.

Type of computer: 1) ICL 1906A, 2) PDP-15. *Size of storage:* 1) 40k, 2) 32k. *Language:*1) FORTRAN IV, PLAN, ALGOL, 2) FORTRAN IV. *Disks or tapes:* 1) 7-track tape. *Special equipment:* SD4020 microfilm recorder. Definitions of characters are available.

References: R.F. Churchhouse and S. Hockey, "The Use of an SD4020 for Output of a Concordance Program," *The Computer in Literary and Linguistic Research,* ed. R.A. Wisbey (Cambridge: Cambridge University Press, 1972); S. Hockey, "A Concordance to the Poems of Hafiz with Output in Persian Characters," *The Computer and Literary Studies,* ed. A.J. Aitken, R.W. Bailey, and N. Hamilton-Smith (Edinburgh: Edinburgh University Press, 1973).

G54. SPIRES (*S*tanford *P*ublic *I*nformation *RE*trieval *S*ystem)

Principal investigator: Edwin B. Parker, Professor, Institute for Communication Research, Stanford U., Stanford, CA 94306. *Associates:* A.H. Epstein, Douglas Ferguson, Donn Martin, John Schroeder.

Scope: SPIRES seeks to create an on-line reference retrieval system serving the research/teaching functions of a major university. *Method:* SPIRES is based on a behavioral science analysis of the information needs of a target population. This information was the basis for developing the SPIRES I system now serving the Stanford Linear Accelerator Center (SLAC) Library. SPIRES II, a production on-line information system, is scheduled for operational service in the fall of 1972. It will have the capability to serve a variety of social science and humanities disciplines in the university. SPIRES II is designed for high reliability, cost acceptability, fast recovery, and a strong user orientation. It will be operational on a daily basis and will be available throughout the Stanford area wherever a terminal can be connected.

Type of computer: IBM 360/67. *Size of storage:* 1000k. *Language:* PL 360, PL/I. *Disks or tapes:* 2314 disks. *Special equipment:* 2741 typewriter terminals and video terminals.

References: "Behavioral Research in the Development of a Computer-Based Information System," delivered at the Conference on Communication among Scientists and Technologists: The Study of the Production, Dissemination and Use of Information by Scientists and Technologists, 28-30 October 1969, at the Johns Hopkins University, to be published in the *Conference Proceedings;* "Democracy and Information Processing," *EDUCOM* (Bulletin of the Inter-University Communications Council) 5, 4 (1970), 2-6; "Developing a Campus-Based Information Retrieval System," *Proceedings, Stanford Conference on Collaborative Library Systems Development,* Stanford University, 4-5 October, 1968 (Stanford, CA: Stanford University Libraries, 1969), 213-230, ERIC document no. ED 031 281; "Information Utilities and Mass Communication," *Information Utilities and Social Choice,* ed. N. Nie and H. Sackman (Montvale, NJ: AFIPS Press, 1970); "Potential Interrelationships Between Library and Other Mass Media Systems," delivered at the Conference on Interlibrary Communications and Information Networks, 28 September-2 October 1970, at Warrenton, VA; "System Theory Analysis of Feedback Mechanisms for Information Systems," delivered at the FID International Congress of Documentation, 21-24 September 1970, at Buenos Aires, Argentina, to be published in the *Congress Proceedings; SPIRES Annual Report:* 1969-70 Report, ERIC Document no. ED 042 481.

G55. Design for a Long-Range Research Program in Libraries and Information Science (Phase I)

Principal investigator: Miles A. Libbey, Director, Research Center for Library and Information Science, Graduate Library School, Indiana U., Bloomington, IN 47401. *Associates:* Bernard M. Fry, Joseph C. Meredith.

Objective: To identify and characterize current research relevant to service operations of Federal technical libraries, information centers, and information analysis centers to establish a basis for identifying technical information problems and policies relating to Federal libraries and information services; to determine priorities for those issues requiring attention; and to develop research designs for each project recommended. *Method:* Research project summary information from COSATI, DOD, ERIC SIE, and other sources is being gathered and screened to identify those projects which appear to have a present or future potential for Federal technical library and information service operations. Projects which qualify for inclusion are classified according to a faceted scheme built around functional aspects of information acquisition, storage, retrieval, and dissemination (including hardware and software considerations). They are also described in a matrix that merges summary information while preserving identity of source. Both files are to be machine-based with retrieval according to combinations of facets and/or descriptors.

Type of computer: 1) CDC 6600, 2) IBM 360/50, 3) IBM 360/30. *Size of storage:* 1) 96k, 2) 256k + 1000k LCS, 3) 64k. *Language:* 1) COMPASS, FORTRAN, 2) and 3) Assembler, FORTRAN. *Disks or tapes:* 1) 1 6638 disk, 6 7-track, 1 9-track tape; 2) 1 2314 disk, 6 9-track tapes; 3) 5 9-track tapes. *Special equipment:* IBM 2741, TTY 33.

References: Bernard M. Fry and Miles A. Libbey, "Status of Library Research in the Federal Government," *Drexel Library Quarterly* 6, 3 and 4 (July and October 1970), 290-305. [*CHum,* May 1971]

G56. Serials Automation Project

Principal investigator: R.F. Doust, Deputy Principal Librarian, Library of New South Wales, Macquarie Street, Sydney, N.S.W. 2000, Australia.

Objective: To maintain records control of all serials held by the library (over 32,000, of which 17,000 are current titles). The system covers accessioning, claiming, binding, and reference work.

References: Public Library of New South Wales, *Automatic Data Processing Study of Serials Section, 1) Report of Exploratory Study,* 1966, 2) *Report of Feasibility Study,* 3 vols., 1967-68; R. Rothwell, "The Serials Automation Project at the Public Library of New South Wales," *The Australian Library Journal* 18, 4 (May 1969), 136-142.

G57. Information Retrieval and Question-Answering from Classical Texts and Archaeological Data

Principal investigator: Gary G. Phillips, Instructor, Dept. of Computer Science, North Carolina State U., Raleigh, NC 27607. *Associates:* Stylianos Danielopoulos, Yolanda Danielopoulos.

Scope: A general system for information retrieval from classical texts, archaeological data, and research articles to facilitate comparative studies in ancient history and gathering literally all available information on any topic. *Method:* Programs are currently being written to process Greek text for grammatical structure by recognizing inflected forms and to retrieve sections from the text which contain certain keywords or complex logical combinations of keywords (using *and, or,* and *not*). This will provide an early version for question-answering from Greek text. The Loeb edition of Polybius is being prepared for the data base.

Type of computer: 1) IBM 360/75, 2) IBM 1130. *Size of storage:* 1) 300k, 2) 16k. *Language:* 1) PL/I, SNOBOL, 2) FORTRAN IV. *Disks or tapes:* 1) 2 2311 disks, 3 tapes; 2) 1 1130 disk. *Special equipment:* Several terminals and Greek type elements. Program is available.

References: Brief description of research in progress in *Calculi* (March 1971), p.121. [*CHum,* November 1971]

G58. Indexing and Classifying the Scientific Study of Biblical Literature

Principal investigator: Simon J. De Vries, Professor, Methodist Theological School, Delaware, OH 43015.

Objective: A comprehensive verse-by-verse and section-by-section index of diverse and widely scattered bibliographical materials offering various kinds of scientific treatment of the Bible. Three specialists, one in biblical exegesis, one in bibliographical research, and one in data processing, will supervise a team of indexers and processors when the project is fully under way; but this must be preceded by a survey of library resources, a pilot indexing project, and a conference of cooperating persons.

References: Arithmoi 1, 2 (August 1971), 13-15, available from Professor Richard E. Whitaker, Central College, Pella, IA 50219. (See also L458.) [*CHum,* November 1971]

G59. Bibliographies [Deleted] [*CHum,* May 1972]

G60. Index to English Biography, 1000-1485 A.D.

Principal investigator: Jerome V. Reel, Jr., Associate Professor, Dept. of History, Clemson U., Clemson, SC.

Objective: To provide a useful index to biographical studies in doctoral dissertations in the United States, Great Britain, and Canada between 1930 and 1970.

Type of computer: IBM 360/50. *Language:* PL/I.

G61. Computer Photocomposition in Classical Publication

Principal investigator: David W. Packard, Assistant Professor, Dept. of Classics, U. of California, Los Angeles, CA 90024.

Objective: To develop programs for use in computer typesetting of scholarly monographs containing Greek and other nonstandard alphabets. *Method:* A typesetting program has been written and two books have been set in type: a monograph of the American Philological Association and a book to be published by the University of California Press. Both books have footnotes at the bottom of the page, and one has elaborate linguistic formulae.

Type of computer: 1) IBM 360/91, 2) RCA Videocomp 840. *Size of storage:* 1) 150k, 2) 64k. *Language:* Assembler. *Disks or tapes:* 1 2314 disk, 1 9-track tape.

References: The American Philological Association Monograph –31, W.F. Wyatt, *The Greek Prothetic Vowel,* was set in type by this program. [*CHum,* May 1972]

G62. Selective Dissemination of Microfiche Documents in a University Setting

Principal investigator: Joseph C. Meredith, Systems Librarian, Professor, Learning Resources Center, Governors State U., Park Forest South, IL 60466.

Objective: To develop, test, and evaluate an economical system for using the Selective Dissemination of Microfiche (SDM) Service of the National Technical Information Service (NTIS) on a local university-wide basis. *Method:* 1) Arrangement with NTIS covering a composite profile of categories and subcategories of interest to the faculty. 2) Provision of a local copying and redistribution service, originals to be accumulated in a central file. Each participant to be issued a personal microfiche reader. 3) Recast composite profile at the end of three months, based on faculty survey. 4) Evaluate cost effectiveness of service; obtain faculty assessmemt of categorical relevance.

Type of computer: Computer and language are extensions of whatever processing system is used by NTIS in operating SDM. *Special equipment:* Arcata Fiche-to-Fiche Copier and Developer.

References: R.R. Tschudi and J.C. Meredith, "The 'PROBE' Retrieval Program—A Description" (February 10, 1972), in ERIC/RIE. Copies obtainable from Learning Resources Center, Governors State University, Park Forest South, IL 60466.

G63. Integrated Text Processing

Principal investigator: Andrew J. Symonds, Manager, Interactive Applications, IBM Cambridge Scientific Center, 545 Technology Square, Cambridge, MA 02139. *Associates:* William Barrett, Charles Goldfarb, Raymond Lorie, Edward Mosher.

Objective: To develop a conversational environment for text processing in a virtual machine, to formulate model of textual structure for application processing, and to apply existing relational data structure capability.

Type of computer: IBM 360/67 with virtual machines. *Size of storage:* 768k. *Disks or tapes:* 2314 disks.

References: Several reports available from above address, including: "Integrated Text Processing for Publishing and Information Retrieval" (1971), "Use of a Relational Access Method Under APL" (1971), "A Software Associative Memory for Complex Data Structures" (1970), "Programming of Professional Society Meetings with an Integrated Text Processing System." [*CHum,* May 1972]

G64. Essay Grading

Principal investigator: George Petty, Professor, Dept. of English, Montclair State College, Upper Montclair, NJ 07043.

Objective: Preliminary analysis of freshman writing. *Method:* Define and count objective indicators of competence in writing.

Type of computer: IBM 360/67. *Size of storage:* 500k. *Language:* SNOBOL4. A simple program has been written in FORTRAN IV and is available from Dr. Ellis B. Page, Bureau of Educational Research, University of Connecticut, Storrs, CT. The title of his project is Project Essay Grade. Information about advanced versions of the program from John Green, University of Connecticut, Bridgeport, CT.

References: Ellis B. Page, "The Imminence of Grading Essays by Computer," *Phi Delta Kappan* (January 1966), pp.238-243. Other articles in mimeo available from Dr. Page. [*CHum,* May 1972]

G65. A Computer Analysis of Sexual Dominance in English Literature

Principal investigators: George A. Borden, Associate Professor; Susan M. Jenkins, Instructor; Dept. of Speech, Pennsylvania State U., University Park, PA 16802.

Scope: To develop a Masculine/Feminine dominance quotient for English literature. *Method:* 1) A massive literature search to determine what terms have been reported to have masculine or feminine projections. 2) A psychological experiment to determine the respective weightings of these terms. 3) A computer analysis of graded material to determine their M/F dominance quotient, using C.L.A.S. (Computerized Language Analysis System). 4) A psychological experiment to see if this quotient correlates with perceived masculinity-femininity of the material.

Type of computer: IBM 360/67. *Size of storage:* 280k. *Language:* PL/I. Program is available.

References: George A. Borden and James J. Watts, "A Computerized Language Analysis System," *Computers and the Humanities* 5, 3 (January 1971), 129-141. [*CHum,* November 1972]

G66. A Cumulative Bibliography of Computer Applications in Humanities Disciplines

Principal investigator: John B. Smith, Assistant Professor, Dept. of English, Research Consultant, Computation Center, Pennsylvania State U., University Park, PA 16502. *Associate:* Susan Jenkins.

Objective: To collect in machine-readable form a comprehensive, cumulative bibliography of materials relevant to computer applications in humanities disciplines. The collection contains references to actual computational analyses as well as formal studies that might be helpful for future studies. The collection now contains 8000 items. *Method:* Bibliographic references are typed directly into the computer. There they are collected, sorted, and merged with the master file. This file can be accessed with a retrieval program or it can be used to produce camera-ready printout in upper/lower case.

Type of computer: IBM 360/67. *Size of storage:* 1024k. *Language:* Pl/I. *Disks or tapes:* 10 tapes. *Special Equipment:* Datel RJE Terminals. Program is available. [*CHum,* November 1972]

G67. Computerized Card Catalog for the Institute for Advanced Studies of World Religions

Principal investigators: Richard A. Gard, Director of Institute Services, Institute for Advanced Studies of World Religions, 532 Melville Memorial Library; Jack Heller, Professor, Dept. of Computer Science, SUNY, Stony Brook, NY 11790. *Associate:* John Forsythe.

Scope: The Institute's holdings of some 24,000 volumes and 12,000 manuscripts and rare books in microform will be cataloged in bibliographic detail by computer techniques. A data base will be constructed and usable via the GRIPHOS information retrieval system. *Method:* Conventional cataloging will be performed by the cataloging staff of the Institute. This information will be converted into a computerized data base from which card catalogs and specialized and global indexes will be produced.

Type of computer: IBM 370/155. *Size of storage:* 240k. *Language:* PL/I. *Disks or tapes:* 3 disks, 2 tapes. *Special equipment:* Mag-Card communicating typewriter. Program is available. Contact Museum Computer Network, Museum of Modern Art, 21 West 53 Street, New York, NY 10019. See also P18 and L531. [*CHum,* November 1972]

See also L7, L26, L29, L40, L45, L51, L79, L103, L119, L146, L191, L206, L213, L388, L396, L410, L414, L415, L416, L446, L447, L458, L471, L481, L530, M6, M16, M17, M24, M29, M49, M50, M51, M52, M53, M62, M68, P7, P10, P11, P13, P14, P15, S29, S31, S37, S61, S88, S91, S100, S103, S108, S110, S148, S169, S174, V5 V10b, V12, V13, V22, V23, V25.

Language and Literature

L1. Computer Concordances and Analyses of Medieval German Lyrics

Principal investigator: Stanley N. Werbow, Professor, Dept. of Germanic Languages, U. of Texas, Austin, TX 78712. *Associate:* Rudolf Jansen.

Objective: The creation of computer concordances of the scholarly collections of medieval German lyrics, and the analysis of these lyrics from thematic and formal points of view. *Method:* Concordances will be prepared for each collection separately with the dual objective of a computer-housed and a published concordance. The latter can be prepared with or without context. A joint glossary will also be prepared to facilitate the study of lyric diction.

Type of computer: IBM 7040. *Disks or tapes:* 12 tape drives. *Language:* JURNAL–Special program on CDC 160A; concordance program in BAP. *Special features:* IBM 1052 Console typewriter. Program is available.

Status: The text of *Minnesangs Frühling* is on tape, but project is now inactive.

References: R. Jansen, "Computer Concordance and Thematic Analysis of *Minnesangs Frühling*" (Ph.D. dissertation).

L2. Concordances of Middle Iranian Languages

Principal investigator: Richard N. Frye, Professor, Harvard U., Cambridge, MA. *Associates:* John Moyne, Hamid Mohammedi.

Objective: To provide the spade work for dictionaries and glossaries.

Type of computer: IBM 360/50. *Language:* PL/I.

Status: Middle Persian inscriptions processed. *Karnamake Ardashire Papakan* partially processed.

L3. Thomas Mann Index

Principal investigator: Herbert H. Lehnert, Professor, Dept. of German, School of Humanities, U. of California, Irvine, CA 92664. *Associates:* Robert G. Porter, Frederick Ruecking, Jr.

Objective: An index of names, works, self-commentaries and topics in nonfictional writings of Thomas Mann. First section 1889-1914. *Method:* 1) Excerpt documents. 2) Transfer to special worksheets for keypuncher with "tags," a coded system for retrieval. 3) Keypunch. 4) Sort.

Type of computer: IBM 7040. *Size of storage:* 32k. *Language:* COBOL. *Disks or tapes:* 4 729 V tapes and 1 1302 disk storage.

Status: Inactive.

References: CHum 1 (1967), 65-71; Robert G. Porter, *The Preparation of a Computerized Index to the Nonfiction of Thomas Mann,* Rice University Studies 55, 2 (1969).

L4. Phonemic Structure of the Spanish Syllable

Principal investigator: Paul M. Lloyd, Professor, Dept. of Romance Languages, U. of Pennsylvania, Philadelphia, PA 19104. *Associate:* Ronald D. Schnitzer (no longer active).

Objective: To obtain statistical data on various types of syllables in Spanish. *Method:* All syllables from the Spanish Academy Dictionary were recorded on IBM cards and analyzed by the computer.

Type of computer: IBM 7040. *Language:* FORTRAN IV.

Status: Initial investigation completed. Now inactive.

References: "A Statistical Study of the Structure of the Spanish Syllable," *Linguistics* 37 (1967), 58-72.

L5. Analysis of Function-Marking Morphemes and Syntactic Structures in Homeric Greek

Principal investigator: Robert R. Dyer, Professor, Dept. of Classics, U. of Massachusetts, Amherst, MA 01002.

Objective: To investigate certain features inadequately explained in Homeric Greek. *Method:* Concordancing of the *Iliad* in reverse order and indexing on selected internal clusters.

Type of computer: IBM 360. *Language:* Machine language.

References: "The Prospects of Computerized Research on Homer," *Revue de l'Organisation Internationale pour l'Etude des Langues Anciennes par Ordinateur* (1966), No. 4.

L6. Computational Investigation of Swedish Newspaper Prose

Principal investigator: Sture Allén, Docent, Research Group for Modern Swedish, U. of Göteborg, Norra Allégatan 6, 413 01 Göteborg, Sweden. *Associates:* Martin Gellerstam, Sture Berg, Staffan Hellberg, Bo Ralph, Rolf Gavare, Jonas Löfström, Jerker Järborg.

Objective: Qualitative and quantitative analysis of a corpus of 1,000,000 words on a specified set of descriptive levels, including graphic words, homograph components, lemmas, phrases, word elements. *Method:* Processing of punched 6-channel paper tapes from printing works.

Type of computer: DATASAAB D21. *Size of storage:* 24k. *Language:* ALGOL-GENIUS. *Disks or tapes:* 6 tapes. *Special equipment:* Reader for symmetrical 6-channel paper tape. Program is available.

Status: Description on word levels completed. Current level: Word combinations, based on a complete body of recurrent combinations. Level in preparation: word elements (morphemes).

References: Sture Allén, "Vocabulary Data Processing," *The Nordic Languages and Modern Linguistics,* ed.H. Benediktsson, (Reykjavik, 1970); Rolf Gavare, *Graph Description of Linguistic Structures* (Stockholm: Almqvist & Wiksell, 1972); Staffan Hellberg, "Computerized Lemmatization without the Use of a Dictionary: A Case Study from Swedish Lexicology," *CHum* 6 (1972), 209-212.

L7. Development of an On-Line Editing System for the CDC 3600

Principal investigator: Richard L. Venezky, Associate Professor, English and Computer Sciences, Dept. of English, U. of Wisconsin, Madison, WI 53706.

Objective: To develop and implement a text and program editing system for a CRT terminal connected on-line to the CDC 3600. *Method:* Text-editing: File-to-file: insert, delete, replace, with some formatting-right justification, etc. Programs: General editing and execution capability.

Type of computer: CDC 3600. *Size of storage:* 65k. *Language:* COMPASS (assembly language). *Disks or tapes:* 2 tape drives. *Special equipment:* Drum, CDC 212 display unit.

Status: Terminated upon change of computer.

References: "3600 Drum SCOPE Routines for the CDC 212 Terminal," Technical Report No.4, Computer Sciences Dept., U. of Wisconsin, July 30, 1967.

L8. A Computer Analysis of Style and the Influence of Genre upon Style

Principal investigator: Elizabeth Ann Carroll, U.S.Mission, Korea, RDD, APO San Francisco, CA 96301.

Objective: To develop computer programs which will aid in analyzing style, with special attention given to the essays of Dr. Samuel Johnson and his contemporaries. *Method:* To study the periodical essays of Dr. Johnson and his contemporaries; to isolate aspects of style, and to attempt to write a program which will provide for computer analysis of given texts.

Type of computer: CDC 3600. *Size of storage:* 64k. *Language:* FORTRAN IV. *Disks or tapes:* 10 606 tape units; 1 drum.

Status: Exploratory stages. [*CHum,* May 1967]

L9. Comparative Metrics (English, French, German)

Principal investigators: Karl Magnuson, Samuel Rosenberg, Assistant Professors; Frank G. Ryder, Professor, Indiana U., Bloomington, IN 47401.

Objective: To collect distributional evidence to support a statement, or series of statements, about the line-unit in its internal shape and historically varying aspects.

Type of computer: CDC 3400-3600. *Size of storage:* 2-32k. *Disks or tapes:* 6 tapes. *Language:* FORTRAN IV. Program is available.

Status: Several pilot studies have been undertaken. [*CHum,* May 1967]

L10. Analysis of Old French Oral-Formulaic Poetry

Principal investigator: Joseph J. Duggan, Associate Professor, Dept. of French, U. of California, Berkeley, CA 94720.

Objective: To establish a distinction between orally composed epics and works created in writing through the analysis of style. *Method:* Use of concordances with right-hand alphabetic subsorts, word-lists, and frequency counts.

Type of computer: IBM 7094. *Language:* FORTRAN IV. *Size of storage:* 32k. Program is available through Mr. Gio Wiederhold, Stanford University Medical Center, Stanford, CA.

Status: Terminated December 1972.

References: "Formulas in the *Couronnement de Louis,*" *Romania* 87 (1966), 315-344; "The Use of Concordances in Literary and Linguistic Research," *Revue de l'Organisation Internationale pour l'Etude des Langues Anciennes par Ordinateur,* 1966; *A Concordance of the Chanson de Roland* (Columbus, OH: Ohio State University Press, 1970); *The Song of Roland: Formulaic Style and Poetic Craft* (Berkeley: University of California Press, 1972).

L11. Cooke's *Preacher's Assistant*

Principal investigator: J. Gordon Spaulding, Associate Professor, Dept. of English, U. of British Columbia, Vancouver 8, BC, Canada. *Associate:* Werner Dettwiler.

Objective: To create a new edition of John Cooke's two-volume work *The Preacher's Assistant* (London, 1783), a catalog of the sermons published in English from 1660 to 1783, by indexing the records in depth, collating chronologically, and examining them statistically. *Method:* It is hoped that the data base may sometime be integrated with a computerized general catalog of eighteenth-century publications in English.

Type of computer: IBM 360/70. *Size of storage:* 1,000,000 bytes. *Disks or tapes:* 16 ITEL 3101 disk drives (double density); 6 IBM 3420 9-track tape drives to IBM 2301 drum storage devices. *Language:* FORTRAN IV with machine-language subroutines. *Special features:* 2 IBM 1403 printers with PN and TN chains, conversational mode peripheral equipment. Program is available.

Status: Completed.

L12. Stylistic Study of James Thomson's *The Seasons*

Principal investigator: Ralph Cohen, Professor, Dept. of English, U.C.L.A., Los Angeles, CA 90024. *Associate:* Thomas Clayton.

Objective: Use of computers as an aid in literary criticism. *Method:* Concordances and frequency counts.

Status: Near completion. *Future plans:*Stylistic study of William Cowper's *The Task*. [*CHum,* May 1967]

L13. A Concordance to *Don Quijote;* A Concordance to Gustavo Adolfo Bécquer's Poetry

Principal investigator: Enrique Ruiz-Fornells, Professor, Dept. of Romance Languages, P.O.Box 4931, U. of Alabama, University, AL.

Type of computer: UNIVAC 1107. *Size of storage:* 64k. *Disks or tapes:* Two 750,000-word drums. *Language:* FORTRAN with subroutines in SLEUTH. Program is available.

Status: Don Quijote ready for publication in 1973. Gustavo Adolfo Bécquer's Poetry: completed. A concordance to Alonso Fernández de Avellaneda's *Don Quijote* as a basis for future comparative study is partially completed. A concordance to Leopoldo Panero's poetry is partially completed. A concordance to Rubén Darío's *Azul* and a concordance to Carlos Reyles' *El Embrujo de Sevilla* are scheduled for the future.

L14. Computer-Aided Research in Machine Translation

Principal investigator: Gerhard Reitz, Member of Technical Staff, Dept. of Data Sciences, Bunker-Ramo, 8433 Fallbrook, Canoga Park, CA 91304.

Objective: Research in machine translation from Russian to English.

Type of computer: IBM 7094. *Size of storage:* 32k. *Language:* SAP/FAB;STL's System B monitor. *Disks or tapes:* 10 tapes. Program is available.

References: Progress Reports 1-14 and Final Report (3/31/67) under NSF contracts C-233 and C-372.

L15. A Hawthorne Concordance

Principal investigator: Jac Tharpe, Associate Professor, Dept. of English, Texas Technological College, Lubbock, TX 79409.

Objective: Besides concordance, to work on problems involving dates of composition.

Status: Cancelled because of duplication of work by others.

L16. Authorship Attribution in Diderot's *Encyclopédie*

Principal investigator: Richard L. Frautschi, Professor, Dept. of French, Pennsylvania State U., University Park, PA 16802.

Scope: Establishment of a Diderot and a non-Diderot universe; development of discriminants which will "recognize" the style of Diderot; comparison with disputed texts. *Method:* 1) Automatic tabulation of discriminant frequencies. 2) Statistical projections of authorship probabilities.

Type of computer: 1) IBM 360/40, 2) IBM 360/70. *Size of storage:* 512k. *Language:* PL/I.

References: "A Project for Author Discrimination in the *Encyclopédie*" (to be published in the *South Atlantic Bulletin*). [*CHum,* November 1967]

L17. A Thematic History of Canadian Poetry

Principal investigator: Sandra Ann Djwa, Assistant Professor, Dept. of English, Simon Fraser U., Barnaby 2, BC, Canada. *Associate:* D. G. Stephens.

Objective: Keypunching complete works of major Canadian poets from 1830 to 1867. *Method:* 1) Keypunch works of selected poets. 2) Transfer to tapes. 3) Establish general concordance. 4) Edit by thematic categories. 5) Printout of selected index.

Type of computer: IBM 7044. *Size of storage:* 32k. *Language:* FORTRAN IV, machine language. *Disks or tapes:* 4 729V tapes. Program is available.

Status: In progress are concordances and thematic indices to the works of 15 major Canadian poets for use in the preparation of a *Thematic History of Canadian Poetry.* Concordances are being developed to the works of Mair, Sangster, Carman, Crawford, Roberts, Lampman, D. C. Scott, Pratt, Smith, Birney, Klein, Page, Avison, and Layton. Program is available. [*CHum,* November 1970]

L18. English Syntax

Principal investigator: Jan Svartvik, Professor, Dept. of English, U. of Lund, Helgonabacken 14, 22362 Lund, Sweden.

Objective: Investigating American English usage by analysis of *A Standard Sample of Present-Day Edited American English* compiled by Professor W. Nelson Francis, Brown University.

Type of computer: UNIVAC 1108. *Size of storage:* 262k. *Language:* FORTRAN IV. Program is available when completed.

References: On Voice in the English Verb (The Hague: Mouton, *Janua Linguarum,* 1966); H.T. Carvell and J. Svartvik, *Computational Experiments in Grammatical Classification* (The Hague: Mouton, *Janua Linguarum,* 1967); "Plotting Divided Usage with *Dare* and *Need,*" *Studia Neophilologica* 40, 1 (1968), 130-140.

L19. Dictionary of the Old High German Glosses

Principal investigator: John C. Wells, Assistant Professor, Dept. of German, Tufts U., Medford, MA 02155. *Associates:* Taylor Starck, Arthur S. Fink, J.Peter Stein, Jere S. Armen.

Objective: Using computational resources in compiling for publication a dictionary and word index to Steinmeyer-Sievers, *Die althochdeutschen Glossen;* printout of a complete dictionary. *Method:* KWIC index with all variants, keyed to headword, and a lexicon of syntactical combinations; multiple sorting into glossaries, cross-listed compounds, reverse alphabetization for suffixes, and semantic groupings.

Type of computer: 1) IBM 7094, 2) IBM 360/50, 3) SDS 940. *Size of storage: Size of storage:* 1) 32k, 2) 1024k, 3) 64k. *Language:* 1) MAP, 2) and 3) SNOBOL4. *Disks or tapes:* 1) 10 tapes, 16 disks. Program is available when completed.

Status: Completion by 1975.

References: "A Word-Index and Dictionary to the Old High German Glosses," *Literary Data Processing Conference Proceedings* (White Plains, NY: IBM, 1965); "'Die althochdeutschen Glossen' after Fifty Years," *Colloquia Germanica* (Bern: Francke Verlag, 1970), 185-191; *Althochdeutsches Glossenwörterbuch,* ed. Taylor Starck and J.C. Wells, Fascicle I (1972), Fascicle II (1973) (Heidelberg: Carl Winter: Universitätverlag).

L20. Concordance to the Poems of Garcilaso de la Vega

Principal investigator: Edward Sarmiento, Professor, Harpur College, Binghamton, NY 13901.

Type of computer: IBM 1620. *Language:* SPS II. Program is available.

Status: Finished so far as computer is concerned. Editing being done by hand. *Future plans:* Concordance to Juan Boscán on a 360/40.

References: The concordance will be published in Spain (Editorial Castalia) and the U.S.A. (Ohio S.U. Press) as the second volume in the *Obras completas* of Garcilaso de la Vega. (Vol. I, text, edited by E. L. Rivers.) [*CHum,* May 1967]

L21. Dictionary of Russian Homonyms

Principal investigator: William W. Derbyshire, Assistant Professor, Dept. of Russian, SUNY, Binghamton, NY.

Objective: To compile and analyze all homonyms as located in five basic (Soviet) dictionaries of the Russian language. *Method:* Storage of and final analysis of c.2000-3000 homonyms in all categories of speech in Russian.

Type of computer: IBM 360/40. *Size of storage:* 256k. *Disks or tapes:* 2 tapes: 2402/Mod.3; 3 disk drives. *Language:* FORTRAN IV (G level).

Status: Just begun. Plans are to publish findings in dictionary form with lengthy introduction. [*CHum,* May 1967]

L22. A Computer-Based Concordance to *The Anglo-Saxon Poetic Records*

Principal investigator: J.B. Bessinger, Jr., Professor, Dept. of English, New York University, New York, NY 10003. *Associate:* Philip H. Smith, Jr.

Objective: Exhaustive concordance to all extant Old English verse. *Method:* 1) Edit for hyphenation and otherwise standardize six volumes of text. 2) Keypunch. 3) Transfer to tapes. 4) Sort and alphabetize. 5) Re-edit for homographs, etc. 6) Re-sort, re-alphabetize, and print.

Type of computer: IBM 360/75. *Size of storage:* 150k bytes. *Language:* PL/I. *Disks or tapes:* 5 tapes. Program is available in Watcon program package.

Status: Completed in 1972. Publication by Cornell University Press.

References: "Computer Techniques for an Old English Concordance," *American Documentation* 12 (1961), 227-229; "A Concordance Generator," *IBM Systems Journal* 3 (1964), 105-111; "A Computer Program to Generate a Text Concordance," *Literary Data Processing Conference Proceedings* (1964), 112-127; *A Concordance to "Beowulf"* (Cornell University Press, 1969).

L23. A Concordance of the Montaiglon-Raynaud Edition of the Old French Fabliaux

Principal investigator: Ben L. Honeycutt, Assistant Professor, Dept. of French, U. of Missouri, Columbia, MO 65201. *Associate:* James Wagner.

Objective: A complete concordance of the Fabliaux with each word presented in context (one octosyllabic line). Reference to poem number and line number in the Montaiglon-Raynaud Edition. *Method:* Over 40,000 lines and 226,000 words placed on cards transferred to tape, then sorted, and printed out. Frequency lists both alphabetical and in descending order of frequency were produced.

Type of computer: IBM 7094. *Size of storage:* 32k. *Disks or tapes:* 15 729 tape units. *Language:* SCATRAN (Symbolic), FAP (Machine-Language level).

Status: Complete except for correction of certain spelling errors. [*CHum,* May 1967]

L24. Concordance to the *Essais* of Montaigne

Principal investigators: Eliane Jasenas, Associate Professor, Dept. of French; G. Mallary Masters, Project Director, Associate Professor, Dept. of Romance Languages, SUNY, Binghamton, NY. *Associate:* Toby Campbell, Alfred G. Lynn.

Method: Division of words by three categories: frequency, location, and/or context.

Type of computer: IBM 360/40. *Size of storage:* 256k. *Disks or tapes:* 2 tapes and 3 disks. *Language:* FORTRAN G. *Special features:* TN5: Upper/lower-case print chain and special slugs. Program is available.

Status: The program package is informally operational; optimization of transcription fidelity and computing efficiency is in process. Final cleanup of the 400,000-word text is scheduled to be complete this year, as will preliminary documentation of the final system. [*CHum*, November 1970]

L25. Studies in Oral and Written Style

Principal investigator: Jane Blankenship, Assistant Professor, Dept. of Speech, U. of Massachusetts, Amherst, MA. *Associates:* William Price, William Scheff. *Objective:* To determine quantitative differences between oral and written style; to check certain basic assumptions made by previously published quantitative studies in this area.

Type of computer: 1) CDC 3600, 2) IBM 1620. *Size of storage:* 1) 32k, 2) 40k. *Disks or tapes:* 6 tapes and 2 drums. *Language:* FORTRAN IV, COMPASS, ALGO, ALGOL, CODAP. *Special features:* Optional binary programs. Program is available.

Status: First of the 6 studies ("On the Oral Sentence") completed. [*CHum,* May 1967]

L26. The Growing Machine

Principal investigator: John W. Carr III, Professor, Moore School of Electrical Engineering, U. of Pennsylvania, Philadelphia, PA 19104. *Associates:* H.J. Gray, S. Gorn.

Objective: To produce high-level linguistic communication between humans and machines. *Method:* Linguistic growth of a machine system under human tutelage.

Type of computer: IBM 7040, IBM 7090, GE 635, IBM 360, PHILCO S-2000. *Disks or tapes:* Disk system required. *Language:* Self-definitional language. Program is available.

Status: General linguistic structure and file system established. *Future plans:* System to become a common system for a large group of graduate students.

References: T.J. Ostrand, "An Expanding Computer Operating System," Interim Technical Report, Moore School of Electrical Engineering; J.W. Carr III, "The Growing Machine," submitted to the *Journal of the A.C.M.* [*CHum,* May 1967]

L27. Analysis of Dactylic Meter in Middle High German Lyric Poetry

Principal investigator: Rudolf Hirschmann, Assistant Professor of German, U. of Southern California, Los Angeles, CA 90007. *Associate:* Bruce A. Beatie.

Scope: To determine norms and limits of variability of syllable quality in relation to position in the Middle High German dactylic verse, and to develop generalized computer programs capable of simulating a human reader performing metrical analysis. *Method:* Generate concordance. With natural language, accent-positions flagged, serving as input, the computer detects syllable length (five possible values), under-filled and over-filled feet, hiatus and elision for all line positions; it calculates totals and averages and outputs them in tabular and graphic form. This has been done for each poem, each poet, and the corpus as a whole, as well as selected verse positions within the strophe. Accuracy is over 99 percent.

Type of computer: IBM 360/65. *Size of storage:* 275k. *Language:* FORTRAN IV. Program is available. [*CHum,* May 1969]

L29. Information Retrieval through Bibliographic Integration

Principal investigators: Lewis Sawin, Professor, and Charles Nilon, Professor, Dept. of English, U. of Colorado, Boulder, CO 80302. *Associate:* Roger C. Clark.

Objective: Complete and test information retrieval system based on complex merge (integration) of existing subject-area bibliographic sources. Originally concerned with English Studies. System now seen as generally applicable. *Method:* Computer analysis of bibliographical citations; computer merging of ostensible "duplicates." Computer retrieval from merged master file.

Type of computer: CDC 160A. *Language:* CDC 160A machine language. All will be rewritten in FORTRAN. Program is available on request through U.S. Office of Education.

Status: 6 programs written. 20,000 bibliographical citations in machine-readable form. Automated analysis of unedited bibliographical citations achieved. *Future plans:* Complete and test the system.

References: Lewis Sawin, "An Integrated Bibliography for English Studies," *College English* 23 (1961), 141-145; "The Integrated Bibliography for English Studies: Plan and Project," *PLA Bulletin* 19 (February 1964), 7-19; Lewis Sawin and Roger Clark, eds., *Integrated Bibliography Pilot Study Progress Report,* Boulder, CO, 1965; Lewis Sawin, Charles Nilon, and Roger Clark, *The Integration, Storage, and Retrieval of Bibliographic Data in English Studies,* Final Report, Cooperative Research Project No. 2189, U.S. Department of Health, Education, Welfare, Office of Education, Boulder, CO, 1965. [*CHum,* May 1967]

L30. University of Colorado Concordance Project

Principal investigator: Lewis Sawin, Professor, Dept. of English, U. of Colorado, Boulder, CO 80302. *Associate:* Roger Clark.

Objective: To develop a concordance program which researchers can use to produce small concordances (up to 25,000 lines) quickly and cheaply; the output is not suitable for publication because of print-chain limitations.

Type of computer: IBM 7044. *Disks or tapes:* Tape oriented. *Language:* FORTRAN IV, machine language. Program is available.

Status: The program (University of California DISCON) is now running and one concordance has been completed (Samuel Daniel's Sonnet sequence *To Delia*). In the works at present are: The elegies and sonnets of Rilke (concorded separately and together); the corpus of Middle High German Lyric poetry; *The Pearl;* Meredith's *Modern Love;* and Rossetti's *House of Life.* [*CHum,* May 1967]

L32. A Computerized Concordance to John Calvin's *Institutes of the Christian Religion*

Principal investigators: F.L. Battles, Professor, Dept. of Church History, Pittsburgh Theological Seminary, 616 North Highland Avenue, Pittsburgh, PA 15206; Charles Miller, Professor, Dept. of Physics, Trinity College, Hartford, CT 06105.

Objective: To produce an exhaustive KWIC concordance to the Latin text of John Calvin's *Institutio Christianae Religionis* (1559). *Method:* 1) Text to card to input tape to segmented output tapes to condensing routine to merging routine to final merged output tapes of key-word-in-context; 2) text to card to input tape . . . to merging routine . . . to type list with frequency count to computer printed type cards to hand analysis of lemmata to keypunching of lemmata to hand sorting of type cards under lemmata to tape–lemmatic index.

Type of computer: Combination of IBM 1410 and IBM 7094, for various stages as available on five locations. Program is available, UNICON/UNICOUNT.

References: F.L. Battles, "John Calvin and the Computer," *ATLA Proceedings,* 1969; Part I, Introduction, Statistical Tables, Corrigenda and Lemmatic Index (246p) printed; Part II, Concordance Proper in 3 rolls of 16mm microfilm (8300 frames). Published December 1972 as No.8 of *Bibliographia Tripotomopolitana,* Barbour Library, Pittsburgh Theological Seminary, Pittsburgh, PA. If there is sufficient demand, the concordance will be printed, offset, in reduced format, from computer printouts.

L33. PRORA (Programs for research on Romance authors) [Deleted] [*CHum,* May 1967]

L34. *Sententiae et Epistulae Hadriani*

Principal investigator: A. Arthur Schiller, Professor, Law School, Columbia U., New York, NY 10027.

Objective: To establish the credibility of the text; restore the *oratio recta* of Hadrian as far as possible. *Scope and method:* Preparation of a corpus of Latin and Greek texts of the Hadrianic epoch for comparative linguistic and stylistic research.

Type of computer: IBM 7040, 7094 II. *Size of storage:* 7094: 32k. *Language:* SNOBOL3. *Special features:* Direct Data Channel 7094/7040. [*CHum,* May 1967]

L35. Shakespeare Concordance

Principal investigator: Marvin Spevack, Professor, Englisches Seminar, Universität Münster, Johannesstrasse 12-17, 44 Münster, Germany.

Objective: The first complete concordance to Shakespeare, it will consist of three interlocking concordances–for individual dramas, characters, and the complete works–with a host of statistical information.

Type of computer: IBM 7094. *Size of storage:* 32k. *Disks or tapes:* 12 tapes. *Language:* FAP and FORTRAN IV.

References: Vol. 1: *Drama and Character Concordances to the Folio Comedies* (1968); Vol.2: *Drama and Character Concordances to the Folio Histories and the Poems* (1968); Vol. 3: *Drama and Character Concordances to the Folio Tragedies* (1968); Vols. 4-6: *A Concordance to the Complete Works* (1969-1970) (Hildesheim, Germany: Georg Olms Verlagsbuchhandlung).

L36. Latin Medieval Urban Terminology in Northern Europe to c. 1350

Principal investigator: Hans Andersson, Fil. Lic., Dept. of History, U. of Gothenburg, Bengt Lidnersgatan 7, Göteborg S, Sweden.

Objective: To investigate the use of terms which refer to urban settlements of different types, taking into account regional and chronological differences, differences in the language of the chancelleries and so on. Regions investigated: Sweden, Denmark, and Westphalia in Germany. Sources are "Urkundenbücher" from these parts of Northern Europe.

Type of computer: DATASAAB D21. *Size of storage:* 24k x 24 bits. *Disks or tapes:* 5 tape units. *Language:* ALGOL-GENIUS. Program is available.

Status: Completed in 1970.

References: Published in Hans Andersson, *Urbanisierte Ortschaften und lateinische Terminologie.* Studien zur Geschichte des nordeuropäischen Städtewesens vor 1350 (Göteborg 1971); Acta Regiae Societatis Scientiarum et Literarum Gothoburgensis. Humaniora 6.

L37. Machine Translation from English and Spanish into French

Principal investigator: B. Pottier and G. Bourquin, 25 rue de Verdun, 54-Nancy, France. *Associates:* M-Claude Bourquin, Josette Lecomte, Annette Evrard.

Objective: Machine translation of scientific texts.

Type of computer: BULL, Gamma 60. *Language:* ALGOL.

Status: Automatic recognition of the input language (syntax and semantics). *Future plans:* A self-contained analysis of the input language is being developed, to be used as a point of departure for translation into any target languages.

References: See *Current Research and Development in Scientific Documentation* (NSF) No.13, p. 229; No.14, pp. 186ff. [*CHum,* May 1967]

L38. A Semantic Keyword Analysis of Selected Texts of Albert Camus

Principal investigator: Robert F. Roeming, Professor of French and Italian, U. of Wisconsin-Milwaukee, Milwaukee, WI 53201.

Objective: To determine a system by which words of highest frequency can be identified in terms of context according to shades of meaning. *Scope and method:* At present texts by Camus, philosophical and fictional, are being compared on a preliminary basis, through contextual relationships within a sentence, to verify whether a valid system by analysis through a computer can be designed.

Type of computer: IBM 3600. *Size of storage:* 100k. *Disks or tapes:* 2 2430 tapes, 1 2314 disk. *Language:* FORTRAN IV.

Status: Emanating from Camus bibliography on computer tape, published in 1967 by University of Wisconsin Press (out of print). *Future plans:* Continuation of Camus bibliography. Development of the design indicated above in scope and method.

L39. Concordancias de la Obra Poetica de Eugenio Florit

Principal investigator: Alice M. Pollin, Associate Professor, New York U., New York, NY 10003. *Associate:* Henry Weitzer.

Objective: To create the first computerized Spanish-language concordance as a model for the continuing series.

Type of computer: 1) IBM 7094, 2) IBM 1620. *Size of storage:* 1) 32k, 2) 40k. *Disks or tapes:* 6 tapes. *Language:* FORTRAN II and SPS.

Status: Published by N.Y.U.Press in March 1967. *Future plans:* Concordances to other Spanish authors; text has been prepared and keypunched for García Lorca (nondramatic poetry and creative prose). [*CHum,* May 1967]

L40. Annotated Bibliography of Godwin Criticism

Principal investigator: Burton R. Pollin, Professor, Dept. of English, Bronx Community College, CUNY, Bronx, NY 10468. *Associates:* George Logemann, Jonathan Bangs.

Objective: To provide a computerized text to 3379 entries with eleven indices, including persons singled out in texts, titles mentioned, authors of all entries merged. *Method:* The entire text has been punched on over 20,000 cards, in five fields with various denotators for extraction of indices and for printout purposes when published.

Type of computer: IBM 360/30. *Size of storage:* 65k. *Disks or tapes:* 2 2311 disks. *Language:* PL/I, FORTRAN IV.

Status: Completed.

References: Godwin Criticism: A Synoptic Bibliography (Toronto: University of Toronto Press, 1967).

L41. Generation of Old Irish Forms and Paradigms

Principal investigator: Eric P. Hamp, Professor, Dept. of Linguistics, U. of Chicago, Chicago, IL.

Objective: To generate at specified chronological levels attested and nonattested Old Irish words and sets of related words by rule from Proto-Keltic input. *Scope and method:* Updated application of nineteenth-century historical reconstruction.

Status: Current objective is to pool available Breton (esp.Vannetais) dialect forms, on basis of field work on publications by etymon; ultimately to form a nucleus of a British Celtic etymological data bank susceptible of reconstructive manipulation.

L42. Statistical Analysis of Lexical Properties of Various Samples in Russian, Czech, English, and German

Principal investigator: Henry Kucera, Professor, Dept. of Slavic Languages and Dept. of Linguistics, Box E, Brown U., Providence, RI 02912. *Associate:* Raoul N.Smith.

Objective: Investigation of quantitative factors in lexicography. *Scope and method:* Large corpora in four languages. Mathematical methods of various kinds.

Type of computer: IBM 360/50. *Size of storage:* 256k. *Disks or tapes:* 5 9-channel tape drives; 1 7-channel tape drive; 1 2314 disk drive, 8 disks. *Language:* PL/I (OS release 8). *Special features:* IBM 2250 Display Unit.

Status: Completed for Russian,Czech, and English.

References: H. Kucera and W. N. Francis, *Computational Analysis of Present-Day American English* (Providence: Brown University Press, 1967); H. Kucera, "Some Quantitative Lexical Analyses of Russian, Czech, and English," *American Contributions to the Sixth International Slavists,* ed. H. Kucera (The Hague: Mouton, 1968), pp. 155-198.

L43. A Concordance to Baudelaire's *Petits Poèmes en Prose*

Principal investigator: Robert T. Cargo, Professor, Box 1857, U. of Alabama, University, AL 35486.

Objective: A computer-generated concordance with statistical analysis of vocabulary. The concordance will be complete, *every* word will be included, output will be in KWIC format.

Type of computer: UNIVAC. *Language:* FORTRAN with subroutines in SLEUTH.

Status: In progress is a concordance to Baudelaire's *Journaux intimes,* projected for 1973. Future work includes Baudelaire's *Paradis artificiels,* art criticism,and other prose works.

References: Concordance to Baudelaire's "Petits Poèmes en Prose" (University, AL: University of Alabama Press, 1971); *A Concordance to Baudelaire's "Les Fleurs du Mal"* (Chapel Hill: University of North Carolina Press, 1965).

L44. Latin Syntax

Principal investigator: Maurice P. Cunningham, Professor, Dept. of Classics, Lawrence U., Appleton, WI 54911.

Objective: Development of a technique of grammatical analysis for Latin which will permit verifiable syntactical description and analysis on the basis of answers to a series of either/or questions.

Status: Inactive but work to be resumed soon.

References: Brief notice in *Revue* of International Organization for Analysis of Ancient Languages by Computer, 1966, no.1, pp.85-86.

L45. Whole Text Processing

Principal investigator: E.P. Miles, Jr., Professor of Mathematics and Director Computing Center, Florida State U., Tallahassee, FL. *Associates:* John Nall, Darrel Allen, Joe White III, Julia Lawson.

Objective: Development of a general library of programs for processing textual material as a service to librarians, literary scholars, or authors. *Scope and method:* Working with faculty members in English Dept. of library school to develop indexing programs (KWIC, KWOC), concordance programs, word-count programs, search programs, etc., to meet research on instructional program requirements.

Type of computer: 1) CDC 6400, 2) IBM 140. *Size of storage:* 1) 32k 60-bit words, 2) 8k 6-bit characters. *Disks or tapes:* 1) 4 tapes, 1 74-million-character disk, 2) 4 tapes. *Language:* FORTRAN-ASCENT-ASPER/SPS-AUTOCODER. Program is available.

Status: Program developed for IBM 709 being reprogrammed for its replacement, the CDC 6400. [*CHum,* May 1967]

L46. Word-Index to Hakonar Saga, Knytlinga Saga

Principal investigator: Dr. Marina Mundt, Sydnesplassen 9, 5000-Bergen, Norway; U.i Bergen, Nord Inst.

Scope and method: Different listings to determine stylistic peculiarities.

Type of computer: IBM 360/50 H. *Disks or tapes:* 2 IBM 1311 disks. *Language:* Assembly. *Future plans:* Publication of word indices.

L47. Concordance to the Fables and Contes of Jean de la Fontaine

Principal investigator: J. Allen Tyler, graduate student; Dept. of Romance Languages, U. of Virginia, Charlottesville, VA.* *Associates:* Bryant C. Freeman, Alan P. Batson.

Objective: 1) Part of doctoral dissertation on vocabulary and style of La Fontaine, 2) publication in book form.*Method:* Line concordance,with word-indexes for "common words"; upper/lower case, with all punctuation and accents.

Type of computer: Burroughs B5500. *Size of storage:* 192k-character core memory. *Disks or tapes:* 4 66-KC IBM-compatible tape drives; 9,600,000-character disk. *Language:* COBOL and ALGOL. *Special feature:* Input on standard Flexowriter paper tape.

Status: In progress, punching well begun. [*CHum,* May 1967]

L48. Concordance du théâtre et des poésies de Jean Racine

Principal investigator: Bryant C. Freeman, Professor, Dept. of French and Italian, U. of Kansas, Lawrence, KS 66044. *Associate:* Alan P. Batson. *Method:* Line concordance, with word indices for "common words"; upper/lower case, with all punctuation and accents.

Type of computer: Burroughs B5500. *Size of storage:* 192k-character core memory. *Disks or tapes:* 4 66-KC IBM-compatible tape drive; 9,600,000-character disk. *Language:* COBOL and ALGOL. *Special feature:*Input on standard Flexowriter paper tape.

Status: Completed.

References: Concordance du Théâtre et des Poésies de Jean Racine, 2 vols. (Ithaca, NY: Cornell University Press, 1968).

L49. A Concordance to the Kuo yu

Principal investigator: Charles MacSherry, Professor of History and of Art, Wright Hall, Smith College, Northampton, MA 01060.

Objective: To provide a linguistic analysis of a Chinese text from the 4th/3rd century B.C., including peculiarities of style and possible regional variations. *Scope and method:* Chinese characters represented by numbers; page and column locations also by numbers.

Type of computer: IBM 1130. *Language:* FORTRAN. Program will be available.

Status: Incipient; computer not yet installed. [*CHum,* May 1967]

L50. Study of Russian Children's Verses by S. Marshak

Principal investigator: Dina E. Crockett, Assistant Professor, Dept. of Russian, Vanderbilt U., Nashville, TN 37203.

Objective: Analysis of phonological, lexical, and grammatical features. *Scope:* 14,000 words of verses for preschool and school-age children; combination of computer analysis and use of code numbers and card sorter.

Type of computer: IBM 1620. *Language:* SPS and FORTRAN II.

Status: Completed.

References: Ph.D. thesis, University of Pennsylvania, December 1966. [*CHum,* May 1967]

L51. A Bibliography of American Doctoral Dissertations on Subjects of Drama and Theatre

Principal investigator: Fredric M. Litto, Professor, Dept. of Theatre, U. of Sao Paulo, University City 05508, Sao Paulo, Brazil. *Associates:* Stephen Callahan, William Marsh, Horst Claus.

Objective: A reference book to be titled *American Dissertations on Drama and Theatre: A Bibliography,* with references to 4565 dissertations from the 1870s through the end of 1965; sections to include the bibliography proper, author index, KWIC index to titles, and a traditional, but machine-produced, subject index.

Type of computer: 1) IBM 1401, 2) GE 625. *Size of storage:* 1) 8k, 2) 72k. *Language:* Symbolic Programming 2, AUTOCODER/FORTRAN IV, COBOL, GMAP. *Disks or tapes:* 1) 4 7330 tapes, 2) 7 MT-26 tapes. Program is available.

Status: Completed.

References: American Dissertations on the Drama and the Theatre: A Bibliography (Kent, OH: Kent State University Press, 1969).

L52. Development of German-English Machine Translation System

Principal investigators: W.P. Lehmann, Director; Rolf A. Stachowitz, Research Coordinator; Linguistics Research Center, U. of Texas, Austin, TX 78712. *Associates:* Annette Stachowitz, Bary Gold, Deanna Cluck, Helen-Jo Hewitt, Karin Zull, Mimi McKnight, Robert Pendleton.

Objective: To develop an operational system for machine translation of German scientific and technical literature into English. *Method:* Computational analysis.

Type of computer: 1) CDC 6400; 2) CDC 6600. *Size of storage:* 1) 65k words; 2) 131k words. *Language:* CDC FORTRAN 2.1. *Disks or tapes:* 6 607 tapes; 2 6638 and 1 841 disks. *Special equipment:* 500k words extended core storage. Program should be available in 1974.

References: W.P. Lehmann, Rolf A. Stachowitz, *Normalization of Natural Language for Information Retrieval, Final Technical Report,* prepared for U.S. Air Force, Air Force Office of Scientific Research, Report No. AFOSR-TR-72-0480 (University of Texas at Austin); Rolf A. Stachowitz, "Requirements for Machine Translation: Problems, Solutions, Prospects," in *Feasibility Study on Fully Automatic High Quality Translation,* prepared for Rome Air Development Center, Griffiss Air Force Base, Rome, NY. Report No. RADC-TR-71-295, vol. 2 (University of Texas at Austin); W.P. Lehmann and Rolf A. Stachowitz, *Development of German-English Machine Translation System; First Technical Report, Feb. 1970-Jan.1971,* prepared for Rome Air Development Center, Griffiss Air Force Base, Rome, NY. Report No.RADC-TR-72-47 (University of Texas at Austin); W.P. Lehmann and Rolf A. Stachowitz, *Development of German-English Machine Translation System; Second Technical Report, Feb.1971-Jan.1972,* prepared for Rome Air Development Center, Griffiss Air Force Base, Rome, NY. Report No. RADC-TR-72-74 (University of Texas at Austin).

L53. Research in Chinese Lexicography

Principal investigator: W.P. Lehmann, Director, Linguistics Research Center, U. of Texas, Austin, TX 78701. *Associate:* L.W. Tosh.

Objective: The Chinese vocabulary, and that portion of the core vocabulary found in modern scientific and technical books and journals, are being studied. *Method:* Entries will be compiled mechanically in a form permitting their automatic publication or their use in mechanical translation and other programs for machine handling of Chinese.

Type of computer: 1) CDC 1700, 2) CDC 6600. *Size of storage:* 1) 16k, 2) 131k. *Disks or tapes:* 1) 3 tapes, 1 disk; 2) 6 tapes, 4 disks. *Language:* FORTRAN; Assembly language. *Special features:* Satellite link to CDC 6600; multiprogrammed; priority interrupts; memory protect; cathode-ray-tube conversational mode station. Program is available.

Status: Completed.

L54. Research in Russian-English Machine Translation on Syntactic Level

Principal investigator: W.P. Lehmann, Director, Linguistics Research Center, U. of Texas, Austin, TX 78701. *Associate:* L.W. Tosh.

Objective: Development of a Russian-English machine translation system capable of producing translations on the syntactic level. *Scope and method:* Initial experiments were performed via the Linguistic Research System on a test corpus of 100,000 words of technical Russian. Methods were developed to determine Russian-English syntactic transfer rules from lexical equivalents.

Type of computer: 1) CDC 1700, 2) CDC 6600. *Size of storage:* 1) 16K, 2) 131K. *Disks or tapes:*) 3 tapes, 1 disk, 2) 6 tapes, 4 disks. *Language:* FORTRAN; Assembly language. *Special features:* Satellite link to CDC 6600; multiprogrammed; priority interrupts; memory protect; cathode-ray-tube conversational mode station. Program is available.

Status: Completed.

L55. Research in Syntactic Translations

Principal investigator: W.P. Lehmann, Director, Linguistics Research Center, U. of Texas, Austin, TX 78701. *Associate:* L.W. Tosh. *Objective:* Emphasis centered on improving the programming and linguistic techniques for syntactic description with particular application to English. *Scope and method:* Programs used in syntactic analysis improved to allow the linguist to work on-line with his descriptions.

Type of computer: 1) CDC 1700, 2) CDC 6600. *Size of storage:* 1) 16k, 2) 131k. *Disks or tapes:* 1) 1 disk, 3 tapes; 2) 4 disks, 6 tapes. *Language:* FORTRAN, Assembly language. *Special features:* Satellite link to CDC 6600; multiprogrammed; priority interrupts; memory protect; cathode-ray-tube conversational mode station. Program is available.

Status: Completed.

L56. The Syntax of Nouns in Twelfth-Century French

Principal investigator: B. Woledge, Emeritus Professor, Dept. of French, University College, London WC1E, 6BT England. *Associates:* David Cox, S.L. Burch.

Objective: Analysis of the syntax of nouns in twelfth-century French. *Method:* Word-counts and word-order analysis, using our own grammatical code.

Type of computer: IBM 360/65. *Size of storage:* 256k. *Language:* FORTRAN IV. *Disks or tapes:* 3 magnetic 7-track tapes. Program is available.

References: B. Woledge et al., "La Declinaison des Substantifs dans *La Chanson de Roland*," Pt. 1, *Romania* 88 (1967), 145-174; Pt. 2, *Romania* 90 (1969), 174-201; B. Woledge, "Le philologue et l'ordinateur," *Marche romane* 40 (Liège, 1970), 41-46.

L57. A Reverse Dictionary of Old French

Principal investigator: Ralph Paul deGorog, Professor, Dept. of Romance Languages, U. of Georgia, Athens, GA 30601.

Objective: A listing of all the Old French words found in Godefroy and Tobler-Lommatzsch which will be grouped according to suffixes, for the purpose of studying word formation in Old French and with possible implications for those interested in Old French poetry. *Method:* The investigator's program for his "Modern French-Old French Lexicon" is being used to program a dictionary of Old French arranged according to word endings. Material from Tobler-Lommatzsch, *Altfranzösisches Wörterbuch* is being added.

Type of computer: 1) IBM 1401, 2) IBM 7094, 3) IBM 360/65. *Size of storage:* 1) 8k, 16k, 2) 32k, 3) 512k high speed, 1000k low speed. *Language:* 1) AUTOCODER, 2) FORTRAN IV, MAP, SORT, 3) FORTRAN H, PL/I. *Disks or tapes:* 1) 2 729V tapes, 4 7330 tapes, 2) 12 729V tapes, 3) 2 7-track, 4 9-track tapes; 1 2314, 2 2311 disks, 2301 drum, 2321 data cell, 2741 remote console disks. *Special equipment:* High-low-equal compare.

References: "Una Concordancia del Poema de Fernan Gonzalez," *Boletin de la Real Academia Espanola* 49, 187 (1969), 279-316; 50, 188 (1970), 137-172, 315-336, 517-557; *Lexique français moderne-ancien français* (Athens, GA: University of Georgia Press, 1973); "Trends in Spanish Vocabulary–1913-1963," *Hispania* 48 (1965), 645-667; "A Plan for a French-Old French Lexicon," in "Notes on Old French Lexicology and Historical Phonetics," *Linguistic Studies Presented to André Martinet,* vol. 2, also appearing in *Word* 24, 193-206; "The Medieval French Prepositions and the Question of Synonymy," *Philological Quarterly* 51 (1972), 345-364; "Les noms des impôts médiévaux en France: Synonymie et formation," *French Review,* Special Issue No.3 (Medieval & Renaissance Studies, 1971), 59-76; "The Concepts 'difficult' and 'easy' in Old French," *Studies in Honor of Mario A. Pei* (Chapel Hill: University of North Carolina Press, 1972), 115-126; "The Concept 'to destroy' in Old French and the Question of Synonymy," *Linguistics* 93 (1972).

L58. Computer-Aided Investigation of Literary Influence

Principal investigator: Joseph Raben, Professor, Dept. of English, Queens College, Flushing, NY 11367. *Associates:* Seymour Goodman, David V. Lieberman, Philip H. Smith, Jr.

Objective: Location of passages in several literary works which suggest intellectual correspondence between their authors. *Method:* Whole texts (Milton's *Paradise Lost* and Shelley's *Prometheus Unbound*) are scanned sentence by sentence to identify passages with high number of significant words in common.

Type of computer: XDS Sigma 7. *Size of storage:* 64k. *Disks or tapes:* 3 tapes. *Language:* FORTRAN IV. Program is available.

References: "A Computer-Aided Investigation of Literary Influence: Milton to Shelley," *Literary Data Processing Conference Proceedings* (Armonk, NY: IBM, 1965), pp.230-274; Seymour Goodman and Raymond D. Villani, "An Algorithm for Locating Multiple-Word Co-occurrence in Two Sets of Texts," ibid., pp.275-292.

L59. Development of a Program to Phonemicize English Text Phonemically

Principal investigator: George K. Monroe, Assistant Professor, Dept. of Languages, Lafayette College, Easton, PA 18042.

Objective: To devise an algorithm and program to convert graphic English text to phonemic transcription. To obtain a large corpus of phonemically transcribed English for linguistic research, and to have this corpus in a form suitable for computer processing. *Method:* Present-Day American English compiled under the general direction of Professor W.N. Francis at Brown University. The method is to perform the transcription, making use of signals of phonemic value contained in the graphic form of English. Such signals are found in certain grapheme-to-phoneme correspondences as well as in isolable segments of graphic words, such as suffixes.

Type of computer: IBM 360/50. *Size of storage:* 256k. *Language:* PL/I. *Disks or tapes:* 3 2311 disk packs and 4 high-density 7- and 9-channel tapes. Program is available.

Status: The segmentation and transcription of graphic suffixes was carried out originally on the IBM 7070. The segmentation program has been rewritten in PL/I and is being revised and tested. *Future plans:* To segment graphic prefixes, if tests support such a stage, and proceed to the transcription of affixes and stems. [*CHum,* May 1968]

L60. Dictionary of *American Regional English*

Principal investigator: F.G. Cassidy, Professor, Dept. of English, U. of Wisconsin, Madison, WI 53706. *Associates:* James Hartman, Margaret Waterman.

Objective: To produce a full-scale historical dictionary of regional usages throughout the United States, collected primarily by direct interview of native speakers, supplemented from many other sources, written and oral. *Method:* On basis of population and settlement history, 1000 communities in 50 states were chosen to represent the nation proportionally. An 1800-question standardized questionnaire was completed in each chosen community by direct interview. Responses, as well as lexical items from many other written and oral sources (disks, tapes, books, diaries, previous dialect collections, etc.) will be tabulated in a Data Summary (from questionnaires only) and data corpus from which editors will produce edited entries. Computer aids: file processing, printouts, CRT, CalComp plotting, final printing.

Type of computer: 1) CDC 3600, 2) UNIVAC 1108, 3) OCR Scanner. *Size of storage:* 1) 32k, 2) 256k, 3) 4k. *Language:* 1) INFOL, 2) COBOL and FORTRAN, 3) Assembly. *Disks or tapes:* 1) BCD tape, 2) BCD and SDF tapes; 25,000,000-word drum (FASTRAND); 3) 9-channel EBCDIC tapes. Program will be available when dictionary is completed, probably 1976.

References: "American Regionalism and the Harmless Drudge," *PMLA* 82 (June 1967), 12-19; "Dialectology and the Electronic Drudge," *Leeds Studies in English* NS II (1968), 135-143; "Collecting the Lexicon of American Regional English," *The Promise of English* (NCTE), 1970; F.G. Cassidy and R.B. LePage, *Dictionary of Jamaican English* (Cambridge: Cambridge University Press, 1967); R.L. Venezky, "Storage, Retrieval, and Editing of Information for a Dictionary," *American Documentation* 19, 1 (1968), 71-79.

L61. A Graphotactic Analysis of the Old English St. Chad

Principal investigator: Nils Erik Enkvist, Professor, Dept. of English, Abo Akademi, Abo, Finland. *Associate:* Winthrop H. Vermillion.

Objective: An analysis of the letter distribution, illustrating OE graphemics and phonemics as well as methods and problems of graphotactics.

Type of computer: IBM 1401.

Status: Computer listings are being analyzed. [*CHum,* May 1967]

L62. CCACL

Principal investigator: Stephen V.F. Waite, Kiewit Computation Center, Dartmouth College, Hanover, NH 03755.

Objective: Analysis of Latin meter, with the eventual aim of setting up chronological criteria; also lexicographical compilation. *Method:* Work with the texts of Statius' *Achilleid* and selected texts of Vergil. Prepare and correct texts in time-sharing. Carry out automated scansion with programs along the lines of those described in Nathan A. Greenberg, "Scansion purement automatique de l'hexamètre dactylique," R.E.L.O. *Revue* 3 (1967), 1-30.

Type of computer: Honeywell 635. *Size of storage:* 160k (16k partition used for all time-sharing programs). *Language:* BASIC 5. Program is available.

Status: Completed: Subsumed in November 1969 in the American Philological Association's Repository of Greek and Latin Texts in Machine-Readable Form.

References: "Metrical Indices to Statius' *Achilleid,*" *Hephaistos* 1, 1 (Summer 1970), 28-70.

L63. Syntactic and Metric Analysis of Blank Verse Styles

Principal investigator: Georgia S. Dunbar, Professor, Dept. of English, Hofstra U., Hempstead, NY 11550. *Associate:* Samuel Dunbar.

Objective: Determination of frequency of significant syntactic and metric features characterizing various styles in English blank verse.

Scope: Machine counting of frequency of symbols used in hand-coded analysis of 25,000 lines of blank verse.

Type of computer: IBM 1620. *Language:* SPS (New College of Engineering Modification).

Status: Completed 1967.

References: "The Verse Rhythms of *Antony and Cleopatra,*" *Style* 5 (Fall 1971), 231-245.

L64. Concordances to the Works of Hartmann von Aue, Wolfram von Eschenbach, and Gottfried von Strassburg

Principal investigator: Roy A. Boggs, Assistant Professor, Dept. of Germanic Languages, U. of Pittsburgh, Pittsburgh, PA 15213.

Type of computer: CDC 6600. *Size of storage:* 13k. *Language:* FORTRAN, Assembly. *Disks or tapes:* 6 tapes, 4 disks.

Status: Project is completed. A complete concordance to the works of the medieval German poet Hartmann von Aue was scheduled to appear in the summer of 1973 in the series *Indices zur deutschen Literatur* (Athenäum Verlag, 2 vols.). A concordance to Gottfried von Strassburg's *Tristan* will appear in 1974 in the same series. Concordances to the works of Wolfram von Eschenbach have also been compiled but are not scheduled for publication. Except for a limited number of forms such as *der (din, daz)* whose occurrences are indexed by verse number, inflected forms of each word are assigned a common lemma and alphabetically ordered for location and identification. An index of proper names is appended to each word.

References: "Computer Konkordanzen zu den Werken Hartmanns von Aue," *Maschinelle Verarbeitung mittelhochdeutscher Texte,* ed. W. Lenders (Berlin: Erich Schmitt, 1973).

L65. Latin Syntax [*CHum,* May 1967] [Deleted]

L66. Differential Word Usage in Descartes

Principal investigator: John M. Morris, 4 Proctor Avenue, Clinton, NY 13323.

Objective: Determine changes in doctrine over time. *Method:* Word count, concordance, nonparametric statistical studies.

Type of computer: CDC 3600. *Size of storage:* 64k. *Language:* FORTRAN IV, Assembly. *Disks or tapes:* 4. Program is available.

Status: Completed June 15, 1969.

References: John Morris, *Descartes Dictionary* (New York: Philosophical Library, 1971); André Robinet, "Descartes à l'ordinateur," *Les Etudes Philosophiques* (Avril-Juin 1970) 219-223; John Morris, "A Plea for the French Descartes," *Dialogue* 6 (1967), 236-239; "Raison, Connaissance, and Conception in Descartes' Meditations," *Sophia* 36 (Rome, 1968), 265-272; "Pattern Recognition in Descartes' Automata," *Isis* 60 (1969), 451-460; "Descartes' Natural Light," *Journal of the History of Philosophy* (April 1973); "The Essential Incoherence of Descartes," *Australasian Journal of Philosophy* 50 (1972), 20-29.

L67. The Use of Computers in the Analysis of Classical Texts

Principal investigator: Fred C.C. Peng, Assistant Professor, Dept. of Linguistics, International Christian University, 10-2, 3 Chome, Osawa, Mitaka, Tokyo 81, Japan. *Associate:* Noah S. Brannen.

Objective: Automatic recognition and assortment of certain types of verbal expressions, using one text. *Method:* Parsing algorithm.

Type of computer: IBM 1130.

Status: Project terminated 1970.

References: "A Recognition Routine for Recognizing Basic Substantive Phrases (Noun Phrases) in Chinese," *Computation in Linguistics: A Case Book*, ed. Paul L. Garvin (Bloomington, IN: Indiana University Press, 1966).

L68. A Study of Gerard de Nerval's Poetical Works

Principal investigator: Christiane Allais, Associate, Dept. of French, U. of California, 1353 Francisco, Berkeley, CA.

Objective: Psychological analysis. *Method:* Cluster analysis.

Type of computer: IBM 7094. Program is available.

References: Statistical analysis of "Les Liaisons Dangereuses" (in progress). [*CHum,* May 1967]

L69. Computational Stylistics

Principal investigators: John S. Bross, Assistant Professor, and Robert A. Sencer, Associate Professor, Dept. of Language and Literature, Rensselaer Polytechnic Institute, Troy, NY.

Objective: Assembly of computer programs and instructional materials into a manual usable by those interested in pursuing computer-oriented studies in the humanities. *Method:* 1) Examination of the needs. 2) Collection and description of suitable computer programs currently in operation. 3) Writing and testing of new subroutines.

Type of computer: IBM 360/50. *Size of storage:* 131k + 1000k large storage. *Language:* 360 Assembly, LISP, FORTRAN IV.

Status: Preliminary and experimental. [*CHum,* May 1967]

L70. Word Frequency in the Modern German Short Story

Principal investigators: George A. Scherer (deceased), Professor, Jon E. Zimmerman, Professor, Dept. of Foreign Languages, Fullerton, CA 92634.

Objective: To find most frequently occurring words in this modern literary genre.

Scope: Approximately 2 million running words in total body of material; 160,000 randomly selected items were counted with help of computer.

Type of computer: IBM 709. *Tapes:* 3. Program is available through University of Colorado Computing Center, Boulder, CO.

Status: Completed.

References: Ph.D. dissertation in Norlin Library, University of Colorado, Boulder. Final report of project sent to U.S. Office of Education (NDEA TITLE IV).

L71. Greene-Chettle Authorship Project

Principal investigator: Warren B. Austin, Professor, Dept. of English, Stephen F. Austin State U., Nacogdoches, TX 75961.

Objective: To test by computational stylistics the hypothesis that *Greenes Groatsworth of Wit* and *The Repentance of Robert Greene* were posthumous forgeries, and specifically, that the *Groatsworth of Wit* was fabricated by the purported editor, Henry Chettle. *Method:* By the use of computer-generated concordances, to ascertain Greene's and Chettle's distinctive language practices in their unquestioned prose; then to compare these usages with those found in the two works of doubtful authorship.

Type of computer: IBM 7094. *Language:* COBOL. Program is available.

Status: Completed. Computerized study of the authorship of *The Defence of Cony-Catching* and of Chettle's share in collaborative plays is planned.

References: Warren B. Austin, "Greene's Attack on Shakespeare: A Posthumous Hoax?" *The Shakespeare Newsletter,* September 1966, pp.29-30; "The Posthumous Greene Pamphlets: A Computerized Study," *The Shakespeare Newsletter* (November-December 1966), p.45; *A Computer-Aided Technique for Stylistic Discrimination: The Authorship of "Greene's Groatsworth of Wit"* (Washington: U.S. Department of Health, Education, and Welfare, 1969, ERIC Reproduction Service, Document ED 030 322).

L72. Graphotactics of Selected Anglo-Saxon MSS.

Principal investigator: Robert D. Stevick, Professor, Dept. of English, U. of Washington, Seattle, WA 98195.

Objective: Formulation of principles and assessment of accuracy of scribal variations in spacing morphic strings of letters. *Method:* For one major MS. at a time, a string analysis of variations in spacing and the recurrent patterns they exhibit.

Type of computer: CDC 6400. *Language:* SNOBOL4. Program is available.

Status: Analysis of graphotactics in Anglo-Saxon MSS. is well advanced; computer-assisted analysis completed for Beowulf MS. *Future plans:* Complete analysis of B.M. MS. Royal 15B.xxii and Tollemache MS. of Orosius' *History.*

References: "Scribal Notation of Prosodic Features in The Parker Chronicle, Anno. 984," *Journal of English Linguistics I (1967); Suprasegmentals, Meter, and the Manuscript of "Beowulf"* (The Hague: Mouton and Co., 1968).

L73. A Concordance to Malory

Principal investigator: William A. Elwood, Assistant Professor, Dept. of English, U. of Virginia, Charlottesville, VA 22901.

Type of computer: Burroughs B5000.

Status: In progress. Text (forthcoming revised edition ed. by Eugene Vinaver) is on uncorrected Flexowriter tapes, which will be corrected when the edition is published. [*CHum,* May 1967]

L74a. Traitement Automatique des Textes Sémitiques de la Bible Analyses Automatiques des Signifiants

Principal investigator: Professor Gérard Weil, Directeur des Recherches, Département Sémitique du Centre de Recherches et d'Applications Linguistiques, Université de Nancy II, B.P. 3397, F-54 Nancy Porte Désilles, France. [*CHum,* May 1969] For current status report write to Professor Weil.

L74b. Zikkaron—Automatisation des Recherches Documentaires

Principal investigator: Professor Gérard Weil,Directeur des Recherches, Département Sémitique du Centre de Recherches et d'Applications Linguistiques, Université de Nancy II, B.P. 3397, F-54 Nancy Porte Désilles, France. [*CHum,* May 1967] For current status report write to Professor Weil.

L75. Centre de Documentation Automatique de la Bible et des Langues Sémitiques

Principal investigator: Professor Gérard Weil, Directeur des Recherches, Département Sémitique du Centre de Recherches et d'Applications Linguistiques, Université de Nancy, II, B.P. 3397, F-54 Nancy Porte Désilles, France. [*CHum,* May 1967] For current status report write to Professor Weil.

L76. Syntax and Mental Disorder: A Computer Study of the Influence of Personality on Literary Style

Principal investigator: Louis T. Milic, Professor, Dept. of English, Cleveland State U., Cleveland, OH 44115. *Associates:* Harlow D. Dunton, Lilla M. Curley.

Objective: To find whether profound personality disturbance (schizophrenia) is reflected in consistent distortion of syntax, and whether such a concept as a "schizophrenic style" has any empirical reality. *Method:* Examination of schizophrenic writings specifically elicited for the project and compared with normal subjects, with reference to the modifications of standard syntax. The grammar is analyzed by a tagmemic model called "Sector Analysis."

Type of computer: IBM 7040/7094 (DCS). *Size of storage:* 32k. *Language:* SNOBOL. *Disks or tapes:* 1 tape. Program is available.

Status: Terminated in 1969. Continuing as a general autostylistic program.

References: "From Aristotle to IBM," *Proceedings of the Conference on Computer Applications to Problems in the Humanities,* April 5, 1969, State University College at Brockport, NY, ed. F.M. Burelbach (Brockport, 1970), 71-92; "Autocoding in Computational Stylistics," *Current Trends in Stylistics,* ed. H.Stahlke and B.B. Kachru (Edmonton, Canada: Linguistic Research, in press).

L77. Aarhus Linguistic Analysis

Status: Terminated. [*CHum,* May 1967]

L78. A Computer-Based Analysis of Television News Writing Style for Listening Comprehension

Principal investigator: Irving E. Fang, Professor, School of Journalism, U. of Minnesota, Minneapolis, MN 55455.

Objective: To find stylistic measures of clarity, including a "listenability" formula, for television news writing; to test TV news and newspaper copy for clarity.

Scope: 156,000 words analyzed by quantitative content analysis methods, 72 samples of copy taken from 6 television and 6 newspaper sources.

Type of computer: IBM 1401. *Size of storage:* 4k. *Language:* AUTOCODER. *Special features:* Multiply-divide, high-low-equal compare, 3 index registers. Fifteen programs are available.

Status: Completed.

References: "The Easy Listening Formula," *Journal of Broadcasting* (Winter 1966); "By Computer: Flesch's Reading Ease Score and a Syllable Counter," *Behavioral Science* (May 1968).

L79. Actors' Specialties in Comedy on the London Stage, 1660-1730

Principal investigator: Ben R. Schneider, Jr., Associate Professor, Lawrence U., Appleton, WI 59411.

Objective: To ascertain actors' specialties in comedy on the London stage, 1660-1730. *Method:* Analysis of all performances of roles (total: 30,000) in the 83 most popular and/or best-known comedies written and produced in the period 1660-1730, in order to discover the predominant traits of the principal roles in each actor's repertoire.

Type of computer: IBM 1620. *Language:* FORTRAN. Program is available.

Status: Computer work is done. Writeups of results are in progress. *Future plans:* If results are considered significant, a series of articles defining the various specialties and assigning them to actors and actresses.

References: "The Coquette-Prude as an Actress's Line in Restoration Comedy During the Time of Mrs. Oldfield," *Theatre Notebook.* [*CHum,* May 1967]

L80. Linguistic and Cultural Affiliation of Place-names in Bengal and Southern Bihar

Principal investigator: Barrie M. Morrison, Assistant Professor, Dept. of Asian Studies, U. of British Columbia, Vancouver 8, Canada. *Associate:* Essop Mia.

Scope: One-half sample of all names in the region, taken from detailed maps. *Method:* 1) Transliterate originals. 2) Edit for various categories. 3) Keypunch. 4) Transfer to tapes. 5) Sort by categories chosen in editing. 6) Print in tabular form.

Type of computer: IBM 7044. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 1 729 mod 5 tape. Program is available.

References: "Political and Religious Change in Bengal from the Fifth to the Thirteenth Century," Association for Asian Studies. [*CHum,* May 1968]

L81. A Concordance of Lessing's Drama

Principal investigator: Edward Dvoretzky, Professor, Dept. of German, U. of Iowa, Iowa City, IA 52240.

Status: Still in planning stage. [*CHum,* May 1968]

L82. A Study of the Text-History of Cassiodorus' *Psalm Commentary*: Analysis of Variants for a Critical Recension of the Text

Principal investigator: James W. Halporn, Professor, Dept. of Classical Studies, Indiana U., Bloomington, IN 47401.

Objective: To develop a method of dealing with large amounts of data from numerous textual variants in a text transmitted in a large number of medieval MSS. *Method:* Development of testing programs to determine feasibility of such a long-range project, including development of programs to indicate MS. relationships based on methods of J. Froger. *Language:* SNOBOL.

Status: Collection of variants continuing; search for suitable programs compatible with CDC equipment continuing. Analysis of data collected from other texts of Cassiodorus is planned.

References: "The Modern Edition of Cassiodorus' *Psalm Commentary,*" *Festschrift Marcel Richard* (Texte und Untersuchungen zur Geschichte der altchristlichen Literatur, 1972); "A New Fragment of Durham Cathedral Library MS B.II, 30," *Classical Philology,* 1973.

L83. 1. Minnesangs Frühling. 2. Walther von der Vogelweide. 3. Minnesangs des 13.Jahrhunderts. 4. Neidhart von Reuenthal

Principal investigator: Rudolph K. Jansen, Instructor, Trinity U., San Antonio, TX. *Associate:* Stanley N. Werbow.

Objective: Index concordances and glossaries for above scholarly collections of Middle High German poetry.

Type of computer: IBM 7040 and CDC 160A. *Size of storage:* 32k. *Disks or tapes:* 12 tape drives. *Language:* Basic Assembly. *Special features*: Teleprocessors, IBM 1052 on JOURNAL program. Program is available.

Status: 1, about completed. 2,3,4, being readied for computer processing. [*CHum,* May 1967]

L84. A Computer-Generated Catalog and Indices to the Title Pages of English Printed Books Preserved in the British Museum's Bagford Collections

Principal investigator: Melvin H. Wolf, Associate Professor, Humanities Program, Pennsylvania State U., Capitol Campus, Middletown, PA 17057.

Objective: To provide catalog and full indices to the title pages of 3045 works which are listed in the original Wing or Pollard & Redgrave Short Title Catalogues and are represented by STC numbers only; the 804 unrecorded title pages are represented by entries for author, title, ornaments, printer, bookseller, place of publication, format, and date. *Method:* The catalog is a printout of all entries in the order in which the title pages appear in the Bagford volumes. Indices to all but format entries have been generated by a library SORT routine.

Type of computer: IBM 360/67. *Size of storage:* 286k. *Language:* FORTRAN IV, SORT routine. *Disks or tapes:* 2 tapes, 1 disk.

Status: Completed.

L85. Concordance to the Gothic Bible

Principal investigator: F.J. de Tollenaere. Institute for Dutch Lexicologie, section Thesaurus, Stationsweg 39a, Leiden, The Netherlands.

Status: This project has been postponed on account of the projects L463 and L464 which are to be considered as a unit, and as a common project of Frans van Coetsem (Cornell), Randall L. Jones, Philip H. Smith (Waterloo) and F. de Tollenaere.

L86. Preliminary Investigation of Sonnet Style

Principal investigator: William Ingram, Associate Professor, Dept. of English, U. of Michigan, Ann Arbor, MI.

Objective: To determine if there is in fact such a thing as "sonnet style," and to define it structurally if it exists. Study limited to English sonnets of the Elizabethan period.

Type of computer: IBM 360/67. *Language:* FORTRAN IV with some SNOBOL.

Status: Temporarily inactive.

L87. Lexicographical and Stylistic Analysis of Middle High German Texts

Principal investigator: Roy A. Wisbey, Professor, Dept. of German, King's College, U. of London, London, England. *Associates:* M.F. Bott, J.G.B. Heal, M. Porter.

Scope: 1) Concordances, reverse indexes, frequency counts and rhyme indexes for poetic works including *Wiener Genesis, Vorau,* and *Strassburg Alexander,* Eilhart's *Tristrant Rolandslied,* Maurer, *Die religiösen Dichtungen,* 2 vols., Rudolf von Ems, *Barlaam und Josaphat.* 2) Provision of word indexes and reverse indexes for prose works, including *Speculum Ecclesiae, St. Trudperter Hohelied,* Schonbach, *Altdeutsche Predigten.* 3) Lexicographical material for complete works of Wolfram von Eschenbach. 4) Stylistic analysis of above texts.

Type of computer: ICT Atlas II. *Size of storage:* 128k. *Language:* Machine language. *Disks or tapes:* 4 Data Products disk files; 6 Potter tape drives. *Special equipment:* Multiprogramming system, multi-access. Program is available, but probably useless because very machine dependent.

Status: Some 5,000,000 words of literary texts in 8 languages on 7-channel tape, in part on magnetic tape.

References: R.A. Wisbey, "Concordance Making by Electronic Computer: Some Experience with the *Wiener Genesis," Modern Language Review* 7 (1962), 161-172 (an extended version appeared in *Uses of the Computer in Anthropology,* Mouton, 1965); R.A. Wisbey, "Mechanization in Lexicography," *Times Literary Supplement,* March 30, 1962 (*Freeing the Mind,* II; the series was republished separately under the latter title in the summer of 1962); R.A. Wisbey, "The Analysis of Middle High German Texts by Computer, Some Lexicographical Aspects," *Transactions of the Philological Society 1963* (1964), pp. 28-46; R.A. Wisbey, *Vollständige Verskonkordanz zur "Wiener Genesis" mit einem rückläufigen Wörterbuch zum Formenbestand* (Berlin: Erich Schmidt Verlag, 1967); M.F. Bott, "Some Problems of Natural Language Computing" (paper given at data processing conference, Graz, Austria, April 1966); R.A. Wisbey and M.F. Bott, "The Work of the Library and Linguistic Computing Centre, Cambridge," Czech Academy of Sciences, Prague. [*CHum,* May 1968]

L88. Vollständiger Verbalindex und Konkordanz aller mit A beginnenden Wörter zu Herrand und Wildoni

Principal investigators: Drs. Mantey and Mater, Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, Berlin W1, DDR.

Type of computer: SOEMTRON und BULL mit speziellen Umbauten und Programmen.

Status: Abgeschlossen. *Future plans:* Gleicher Index und Konkordanz zu dem Stricker (Schwänke). [*CHum,* May 1967]

L89. Verzeichnis der Lemmata aller alphabetisch geordneten Wörterbucher des 15. bis 18. Jhdts. für die Anfangsbuchstaben Aa bis Az

Principal investigators: Drs. Mantey and Mater, Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, Berlin W1, DDR. *Associate:* Dipl. phil. Schröter.

Type of computer: SOEMTRON and Bull mit speziellen Umbauten und Programmen.

Status: Abgeschlossen. [*CHum,* May 1967]

L90. Englisches Wörterbuch

Principal investigators: Dr. habil Agricola and Dr. Mater, Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, Berlin W1, DDR.

Type of computer: SOEMTRON und BULL mit speziellen Umbauten und Programmen.

Status: Abgeschlossen 1967. *Future plans:* Maschinelles Erzeugen aller Flektierten Formen. [*CHum,* May 1967]

L91. Index und Frequenzliste zum englischen Wortschatz verschiedener Fachgebiete

Principal investigators: Dr. habil. Agricola and Dr. Mater, Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, Berlin W1, DDR.

Type of computer: SOEMTRON und BULL mit speziellen Umbauten und Programmen.

Status: Umfang c. 250.000 Wörter. *Future plans:* Fortführung bis c. 5.000.000 Wörter. [*CHum,* May 1967]

L92. Personen- und Werkregister für *Quellen und Zeugnisse zur Druckgeschichte von Goethes Werken*

Principal investigators: Dr. Hagen and Dr. Mater, Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, Berlin W1, DDR.

Type of computer: SOEMTRON und BULL mit speziellen Umbauten und Programmen.

Status: Abgeschlossen. [*CHum,* May 1967]

L93. Verzeichnis der Lemmata aus den 15 grossten Wörterbuchern der deutschen Gegenwartssprache für die Anfangsbuchstaben H bis Z

Principal investigators: Dr. Klappenback and Dr. Mater, Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, Berlin W1, DDR. *Associate:* Dr. Scharnhorst.

Type of computer: SOEMTRON und BULL mit speziellen Umbauten und Programmen.

Status: Abgeschlossen bis Buchstabe N. *Future plans:* Fortführung der Arbeit bis Z. [*CHum,* May 1967]

L94. Verzeichnis der deutschen Polyseme und ihrer einzelnen Bedeutungen

Principal investigator: Dr. Mater, Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, Berlin W1, DDR.

Type of computer: SOEMTRON und BULL mit speziellen Umbauten und Programmen.

Status: Fertig bearbeitet Anfangsbuchstaben A bis L. *Future plans:* Abschluss der Arbeiten M bis Z. [*CHum,* May 1967]

L95. Deutsche Verben, Band 1-10

Principal investigator: Dr. Mater, Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, Berlin W1, DDR.

Type of computer: SOEMTRON und BULL mit speziellen Umbauten und Programmen.

Status: Band 1-4 abgeschlossen. *Future plans:* Band 5-6 in Arbeit.

References: Deutsche Verben, Band 1: *Alphabetisches Gesamtverzeichnis* (Leipzig: VEB Bibliographisches Institut, 1966), Band 2 (im Druck), erscheint April 1967; Band 3 (im Druck), erscheint 1967; Band 4 (im Druck). [*CHum,* May 1967]

L96. Rückläufiges Wörterbuch zu Manghol- und Niuca Tobca'an

Principal investigators: Dr. Vietze and Dr. Mater, Humboldt-U., Berlin, Ostasiatisches Institute; Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, Berlin W1, DDR. *Associate:* Dipl. Phil. Zeuner.

Type of computer: SOEMTRON und BULL mit speziellen Umbauten und Programmen.

Status: Auf Lochkarten übertragen. *Future plans:* Weitere Auswertung des Materials.

References: Wird gedruckt im Verlag VEB Enzyklopädie, Leipzig. [*CHum,* May 1967]

L97. Rückläufiges Wörterbuch der deutschen Gegenwartssprache

Principal investigator: Dr. Mater, Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, Berlin W1, DDR.

Type of computer: SOEMTRON und BULL mit speziellen Umbauten und Programmen.

Status: Abgeschlossen. *Future plans:* Untersuchung von Graphemdistributionen.

References: Rückläufiges Wörterbuch der deutschen Gegenwartssprache, VEB Bibliographisches Institut, Leipzig, 1965. [*CHum,* May 1967]

L98. KWIC Study of Finnish Language

Principal investigator: Paavo Siro, Professor, U. of Tampere, Kalevantie 4, Tampere, Finland.

Objective: Study of word order in Finnish language. *Scope and method:* 28,000 words of text are indexed by a SHARE library program EI KWIC.

Type of computer: IBM 7090. *Size of storage:* 32k. Program is available from SHARE. [*CHum,* May 1967]

L99. "Breakdown"

Status: Discontinued. [*CHum,* May 1967]

L100. Semitic Language Concordances from Computers

Status: Discontinued. [*CHum,* May 1967]

L101. Computational Stylistics

Principal investigator: Sally Y. Sedelow, Professor, Depts. of Computer Science and Linguistics, U. of Kansas, Lawrence, KS 66044. *Associate:* Walter A. Sedelow, Jr.

Objective: Computer-aided analysis of literary style.

Type of computer: IBM 360/65, IBM 360/7; Honeywell 635. *Size of storage:* 360/7, fast core, 512k/2000k characters slow core; Honeywell 635, 200k total memory of 36-bit words, 128k of microsecond memory, 72k of 2 microsecond memory. *Language:* IBM/360: PL/I; Honeywell: FORTRAN IV. Program VIA is available.

Status: Statistical and other procedures are being tested.

References: S.Y. Sedelow, "Shakespeare Studies and the Computers," *Shakespeare 1971, Proceedings of the World Shakespeare Congress,* ed. Clifford Leech and J.M.R. Magnuson (Toronto: University of Toronto Press, 1972); S.Y. Sedelow, *Automated Language Analysis,* September 1, 1971, to August 31, 1972, University of Kansas; S.Y. Sedelow and Walter A. Sedelow, Jr., "Models, Computing, and Stylistics," *Current Trends in Stylistics,* ed. B.B. Kachru and H.F. Stahlke (Linguistic Research, Inc., 1972); S.Y. Sedelow, *Automated Analysis of Language Style and Structure in Technical and Other Documents,* Technical Report No.1, Defense Documentation Center, Springfield, VA, DDC –735-134; S.Y. Sedelow and Walter A. Sedelow, Jr., "Categories and Procedures for Content Analysis in the Humanities," *The Analysis of Communication Content,* ed. George Gerbner et al. (New York: John Wiley, 1969); S.Y. Sedelow, W. A. Sedelow, Jr., and T.L. Ruggles, "Some Parameters for Computational Stylistics," *Proceedings, Literary Data Processing Conference* (Yorktown Heights, NY: IBM, 1964).

L102. Evaluation of Thai Poetry

Principal investigator: Udom Warotamasikkhadit, Dept. of Humanities and Social Sciences, College of Education at Prasarnmit, Bangkok, Thailand.

Objective: 1) To determine whether poems tested are well-formed poems according to the meters and rules in Thai poetry. 2) To count the occurrence of internal and external rimes. 3) To determine how good the poems sound phonologically. *Method:* I am comparing the poems written by a famous Thai poet and my own poems. At first I took random samples of poems and tested the program.

Type of computer: IBM 7090. *Language:* MAD. Program is available.

Status: In process of writing programs for other types of Thai poetry. *Future plans:* To use computer to determine the names of authors of few Thai literatures which are unknown and to set up machine translation project in Thailand. [*CHum,* May 1967]

L103. An Information Retrieval System for Speech Communication Research

Principal investigator: George A. Borden, Associate Professor, Dept. of Speech, Pennsylvania State U., University Park, PA 16802. *Associates:* Susan M. Jenkins, John D. Stone.

Objective: To build a thesaurus and a computerized retrieval system for speech communication research and our collection of *Vital Speeches. Method:* The manuscripts of all the speeches submitted to *Vital Speeches* since 1934 are deposited in The Pennsylvania State University Library. A data base of articles from nine speech-related journals has been created. The preliminary faceted thesaurus and retrieval system have been developed to give a bibliography of research and speeches pertinent to a user's request.

Type of computer: IBM 360/67. *Size of storage:* 280k. *Language:* PL/I.

References: George A. Borden, Susan M. Jenkins, and John D. Stone, "A Computer-Aided Thesaurus Construction: The Speech Communication Association Information Retrieval System," *Today's Speech* (Spring 1972), pp.11-16.

L104. A Computational View of Style in the Poems of Gerard Manley Hopkins

Principal investigator: Everett G. Powell, Associate Professor, Dept. of English, Del Mar College, Corpus Christi, TX 78404. *Associates:* Wallace Koenning, Jon Wiley.

Objective: To apply computational stylistic analysis to significant factors in the poems; in effect, ranking the poems on the basis of selected criteria. *Method:* 1) Construction of concordance, including variant readings. 2) Compiling lists of occurrence and tables ranking the poems according to three criteria: punctuation for hyperbaton, average line length per poem, and occurrence of compound coinages. 3) Composite table ranking the poems in one list according to the three criteria.

Type of computer: IBM 360/40. *Size of storage:* 64k. *Language:* PL/I and ALC. *Disks or tapes:* 3 2311 disks; 3 2415 tapes.

Status: Completed in May 1969.

References: Unpublished doctoral dissertation, University of Texas at Austin, 1970.

L105. STRIL

Principal investigators: G.M. Gillow, Consultant to the Norwegian Research Council for Science and the Humanities; Kolbjorn Heggstad, Associate Professor, Nordic Institute, U. of Bergen, Postboks 563, N-5001 Bergen, Norway. *Associate:* G. Opheim.

Objective: Construction of a conversational programming system for linguists. The system will, in the first instance, be applied to Norwegian. *Method:* The archive is being built up from existing word-lists in Modern Norwegian and hopefully will be supplemented by semiautomatic excerption from a machine-readable text library.

Type of computer: IBM 360/50. *Size of storage:* 384k. *Language:* PL/I-based language (STRIL). *Disks or tapes:* 1 IBM 2314 disk; 2 IBM 2415 tapes.

Status: A machine-stored data bank in both *bokmaal* and *nynorsk* is being established for linguistic analysis and dictionary construction. At present the material is stored with grammatical classification; eventually semantics will be included.

References: The STRIL Linguistic Programming System: An Intermediate Report, Bergen, 1966 (out of print). [*CHum,* May 1970]

L106. A Concordance to the West Saxon Gospels

Principal investigator: Paul W. Pillsbury, Professor, Dept. of English, Eastern Michigan U., Ypsilanti, MI 48197. *Associate*: Charles W. Peck.

Scope: A complete, syntactically parsed concordance to The West Saxon Gospels. *Method:* Transliterate the West Saxon text onto DIGITEK sheets. Pre-edit a parallel set of DIGITEK sheets containing canonical word forms plus syntactic subsort code keys. DIGITEK cards automatically control computer card printing; list cards and proofread.

Type of computer: IBM 360/67. *Size of storage:* 512k. *Language:* Level F Assembler and FORTRAN G. Program is available from Regina Frey and Laura Gould, *TRICON Concordances* (New Haven: Linguistics Automation Project, Yale University, 1967). The specific format adaptation to this program can be obtained from P.W. Pillsbury. [*CHum,* May 1969]

L107. Short Projects in Grammatical Analysis of the Brown University Standard Corpus of Edited American English

Principal investigators: William Card, Professor Emeritus, and Virginia McDavid, Professor, Dept. of English, Chicago State U., Chicago, IL 60628.

Objective: To establish 1) the standard forms of irregular verbs, 2) the percent of auxiliary uses of *be, have, do,* 3) the structural vocabulary of English, 4) the special order of limiting adjectives in noun phrases. *Method:* Standard corpus only; concordance printout.

Type of computer: IBM 360/50. *Size of storage:* 256k.

L108. The Amorite Language

Principal investigator: I.J. Gelb, Professor of Assyriology, Dept. of Oriental Languages and Civilizations and Dept. of Linguistics, U. of Chicago, Chicago, IL 60637.

Objective: Reconstruction of Amorite as based on personal names. *Method:* Since we lack inscriptions in Amorite, a statistical and linguistic analysis will be made of all known Amorite personal names, each of which has a meaning.

Type of computer: IBM 7090. *Language:* COMIT and FAP. *Special feature:* Will be in "regular" and tabulated form, plus a possible sortkey coding. Program is available.

Status: Project is currently delayed pending implementation of COMIT II on the IBM 360. During this time the data deck is being increased by about 20 percent. Project should be completed early in 1974.

L109. An Index to the *Cahiers du Sud*

Status: Terminated. [*CHum,* May 1967]

L110. Linguistic Analysis of Spanish Colonial Documents

Principal investigator: Peter Boyd-Bowman, Professor, Dept. of Modern Languages, Crosby Hall 229, State U. of New York at Buffalo, Buffalo, NY 14214. *Associates:* Richard Auletta, William Cline, Juan Zamora.

Objective: To find a key to the early differentiation of the major varieties of American Spanish existing today, through the study of precisely dated documents, taken from every region of Spanish America throughout the colonial period. *Method:* Each linguistic feature is assigned a category, place and date, with enough of the surrounding context quoted to illustrate its use and meaning.

Type of computer: IBM 7044. *Size of storage:* 32k. *Language:* FORTRAN IV, MAP. *Disks or tapes:* 12 729 Mod. 5 tape drives. *Special features:* Column binary; buffered I/O; double precision. Program is available.

Status: Completed for the sixteenth century. So far the data bank has furnished raw material for several doctoral dissertations (two of them completed), and for a reference volume entitled *Léxicohispanoamericano del siglo XVI.*Work is expected to begin soon on a similar data bank for the seventeenth and eighteenth centuries.

L111. Rabelais: Concordances and Study of Homogeneity

Principal investigator: J.E.G. Dixon, Professor, Dept. of French, U. of Winnipeg, Winnipeg, MB, Canada. *Associates:* Martin Porter, Monique Johnson.

Objective: To carry out a computer-aided investigation into the authenticity of Rabelais' *Cinquiesme Livre,* largely by comparisons with the *Quart Livre,* and, if positive, to extend it to the *Quart Livre* itself. The main concordance will comprise citations, unite under one lemma all related inflected forms, include variants, and be supplemented by a complete Index Nominum and various verbal indexes. (For want of a complete critical edition of Rabelais, the text used is that edited by Jacques Boulenger, completed and revised by Lucien Scheler [Paris:Gallimard, "Bibliothèque de la Pléiade," 1955].) *Method:* The concordance and indexes will be computer-generated, with editor intervention at various stages. Essential, alike to the projected investigation and to a prose concordance, is pre-editing of the text, entailing insertion of "tags" to delimit contexts, identify proper names, distinguish homographs, mark discourse, identify speaker, etc. The study will bear initially on lexical peculiarities, similarities, and discrepancies; thematic comparison; character analysis, this being initiated by sub-concordances for each of the main personages.

Status: Keypunching of text to be completed in early 1973. Verification in progress. Concordance to be completed 1974. Program is available.

Type of computer: ICT Atlas II. *Size of storage:* 128k. *Language:* Machine language. *Disks or tapes:* 6 Potter tape drives; 4 Data Products diskfiles. *Special Equipment:* Multiprogramming system, multi-access.

L112. Automatic Syntactic Analysis of the Prose of Thomas Carlyle

Principal investigator: Robert L. Oakman, Assistant Professor, Depts. of English and Computer Science, U. of South Carolina, Columbia, SC 29208.

Objective: To see what elements, if any, of Carlyle's style in his syntax are significant, and if syntactic analysis on a large scale is worth while and might be applied to other prose writers. *Scope:*Random samples of 200 paragraphs of Carlyle's prose—approximately 1200 sentences. *Method:* 1) Machine parses random samples (unedited on punch cards) and prints out syntactic elements and patterns in diagrammatic form. 2) Hand analysis of printout. 3) Further runs to test out hypotheses (may include more samples, concordancing, statistical analyses).

Type of computer: IBM 7040. *Language:* COMIT. Program is available.

Status: Completed.

References: "Syntax in the Prose Style of Thomas Carlyle: A Quantitative, Linguistic Analysis," *Dissertation Abstracts* 32 (August 1971), 978A-979A (Indiana).

L113. Statistical Profile of Old Norse

Principal investigator: L. Michael Bell, R.D. 2, Coburn, PA 16832.

Objective: 1) Syntactic: stylistic analysis of Old Norse. 2) Teaching aids: frequency word lists, etc.

Status: Concordance of one 66,000-word text achieved, with possibility of producing an adjective-parsing subroutine to identify inflected forms of adjectives in the text and lump them under their lemmata. See author's remarks, *Computers and Old English Concordances*, ed. A. Cameron (Toronto, 1970), pp.110-111, as well as in his Ph.D thesis.

References: "A Computer Concordance to *Egils saga Skalla-Grímssonar," Studies for Einar Haugen* (The Hague: Mouton, 1972), pp.58-68. Further studies are presented in "Lexical Patterning in *Egils saga Skalla-Grímssonar*—A Computer-Aided Approach" (Unpub. Ph.D. thesis, Harvard University, 1973).

L114. An Index Verborum to the Works of Sallust, Including the Spurious Epistulae

Status: Terminated. [*CHum,* May 1967]

L115. A Comparison of Tilley's *Proverbs* and Kyd's *Spanish Tragedy*

Principal investigator: S.M. Guyton, Major, USAF Academy Preparatory School, Boulder, CO 80860.

Objective: To determine the proverbial content of the play.

Type of computer: Burroughs B5500. *Language:* ALGOL. *Disks or tapes:* 4 tapes. Program is available.

Status: 1) Tilley's proverbs are on cards, 2) index program is completed but not debugged, 3) comparison program begun. [*CHum,* May 1967]

L116. Analyse automatique des textes latins

Principal investigator: L. Delatte, Professor, Laboratoire d'Analyse statistique des Langues anciennes, 110, Boulevard de la Sauvenière, Liège, Belgium.

Scope: 1) Etude des oeuvres de Sénèque (continuation) et d'Horace: analyse des *Lettres à Lucilius* de Sénèque, des tragédies de Sénèque, et des odes d'Horace. Le but est de rassembler sur ces diverse oeuvres toutes les informations possibles relativement au vocabulaire, à la morphologie, à la syntaxe et au style. Cette masse d'information est exploitée dans des études sur la langue, l'art et la personnalité de l'auteur. 2) Enregistrement et exploitation automatiques de papyrus grecs documentaires. Le but de ce travail est d'enregistrer le texte de papyrus et les conjectures et corrections qui s'y rapportent de manière à pouvoir obtenir des transcriptions diplomatiques, des éditions intégrant les corrections et conjectures retenues, des indices et listes des fréquences, et des relevés et concordances relatifs à toute espèce d'éléments (groupes de lettres, mots ou formes, groupes de mots ou de formes, etc.). *Method:* Analyse morphologique et lemmatisation automatiques des textes à étudier. Etablissement automatique des listes, relevés et dénombrements aux divers points de vue cités ci-dessus. Impression automatique de la documentation. Application automatique des textes

statistiques aux données numériques. Exploitation des résultats obtenus; cette exploitation est évidemment le fair des machines.

Type of computer: IBM 360/20 sub.4. *Size of storage:* 16k. *Language:* PL/I(DPS). All programs are available. *Disks or tapes:* 2 2311 disks.

Status: Seneca, *De Beneficiis. Future plans:* General index of Seneca.

References: Revue of the International Organization for Ancient Languages Analysis by Computer (quarterly)

L117. An Investigation of the Structure of an English Lexicon

Principal investigators: John Olney and Carter Revard (Associate Professor of English), System Development Corporation, 2500 Colorado Avenue, Santa Monica, CA; Paul Ziff, Professor, Dept. of Philosophy, U. of Wisconsin, Madison, WI.

Objective: To provide computational aids for obtaining a formal semantic description of English. *Method:* 1) Transcribe the lexical portions of *Webster's Seventh New Collegiate Dictionary* (pp.1-1041) and *The New Merriam-Webster Pocket Dictionary* (pp.1-592) onto magnetic tape; 2) program interactive processor to enable users at remote consoles to work with the transcripts and, in particular, to trace semantic fields for word senses of interest to them.

Type of computer: 1) IBM Special Processor original transcending; 2) IBM 360/65 Special Processor. *Size of storage:* 1) 65k 48 bit words 2*mus* cycle time; 2) 256k. *Language:* 1) LISP, with metacompiler features; 2) Assembler F, PL/I F.

Status: The transcription of *Webster's Seventh* has been completed. An analytic study of word-sense development is under way, and portions of the processor are being programmed.

References: "A Research Plan for Investigating the Structure of an English Lexicon," TM-(L)-3331, System Development Corporation (in progress); to appear in *The Finite String:* "An Interactive Processor for Machine-Usable Version of *Webster's Seventh New Collegiate Dictionary* (W7) and *The New Webster-Merriam Pocket Dictionary* (MPD)." [*CHum,* May 1967]

L118. Applied Linguistics

Principal investigator: John A. Moyne, Associate Professor, Dept. of Computer Science, Queens College, CUNY, Flushing, NY 11367. *Associate:* Morris Halle.

Objective: Use of natural languages for communication with computers. Concordances and indexes of rare texts and inscriptions of ancient and medieval languages (e.g., Khotanese Saka and Pahlavi). Development of proto-RELADES and CUE systems for natural language processing and information retrieval. *Method:* Computer programs and special computer languages.

Type of computer: IBM 360/50. *Size of storage:* 256k plus 4000k extended memory. *Language:* PL/I and special languages developed for the project. *Disks or tapes:* 9 disks, 4 tapes. *Special equipment:* IBM 2260 display screens. Program is available by arrangement with IBM Corporation.

Status: Completed 1971.

References: "Restrictive Language Defining System," *IBM Technical Report,* ACU-011 BAP 23 (SP), December 31, 1963; "A Khotanese Saka Concordance–On the Use of Computers in Linguistic Studies," paper read at the Joint Harvard University and Rockefeller Foundation Symposium, Como, Italy; "Introduction to an Operational RELADES," *IBM Technical Report,* TROO.1462, April 15, 1966; "Towards the Understanding of Natural Languages by Machines," *Proceedings of the 10th International Congress of Linguists,* Bucharest, 1970; "Proto-RELADES: A Restrictive Natural Language Processing," *IBM Technical Report,* EPC 3, October 3, 1967; "An Introduction to Transformational Grammars," *International Journal of Computer Mathematics,* vol.2, 1968, 169-181; "Information Retrieval and Natural Language," *Proceedings of the American Society for Information Science,* vol.6, 1969; "A Survey of Grammars and Recognizers for Formal and Natural Languages," *IBM Technical Report,* FSC 71-6006, May 1971, Washington, DC; "FORTRAN versus English," *Share Fortran Project Newsletter* 2 (December 1971), New York.

L119. A Knowledge Availability System for Mechanical Retrieval of Ephemeral Theatre Materials (USOE Project 6-10-247)

Principal investigator: Ned A. Bowman, Associate Professor, Dept. of Speech and Theatre Arts, U. of Pittsburgh, Pittsburgh, PA. [*CHum,* May 1967]

L120. Identifying Authors of Unsigned Newspaper Editorials

Principal investigator: Wayne Allen Danielson, Professor, School of Journalism, U. of North Carolina, Chapel Hill, NC 27514. *Associate:* Harold L. Jackson.

Objective: To improve our ability to assign editorials to particular writers. *Method:* A number of stylistic "markers" will be studied first in known editorials by W.J. Cash in the Charlotte *News,* and then in the disputed editorials. Comparisons will be made with editorials by other writers.

Type of computer: IBM 360/75. *Language:* Assembly. Program will be available when finished. [*CHum,* May 1967]

L121. Index Verborum to the *Noctes Atticae* of Aulus Gellius

Principal investigator: Charles W. Dunmore, Associate Professor, Dept. of Classics, New York U., New York, NY 10003. *Associate:* Jay Woods.

Objective: Index of words used by Gellius, both Latin and Greek; listing of Latin and Greek writers quoted, with quotations in full; listing of all legal material, and an alphabetical listing of every historical name found in the work.

Type of computer: 1) IBM 360/50, 2) UNIVAC 1108. *Size of storage:* 1) 256k, 2) 132k. *Language:* 1) PL/I F, 2) field maintenance routines. *Disks or tapes:* 1 1316 disk pack, 2) 1 Fastrand drum.

Status: To be completed 1974.

L122. Word Indexes of the *Rolandslied* Fragments

Principal investigator: Clifton D. Hall, Assistant Professor, Dept. of German, U. of Iowa, Iowa City, IA 52240. *Associate:* Roy Wisbey, who is furnishing a concordance of the principal *Rolandslied* manuscript (Wesle edition).

Scope: All of the fragments in the Wesle edition, plus Schilter's edition of MS. A (1727). *Method:* 1) Keypunch one verse per card. 2) Sort according to alphabetical and numerical requirements (reverse indexes will be included). 3) Print in traditional word-index format. 4) Add diacritical marks and superscripts manually.

Type of computer: IBM 360/65. *Size of storage:* 512k. *Language:* COBOL F. *Disks or tapes:* 4 2311 disks; 6 7-track tapes, 2 9-track tapes. [*CHum,* November 1967]

L123. Stylometry in the Works of Xenophon

Principal investigator: Norman Duncanson Thompson, St. Andrews U., 7 Havers Place, Hopeman, Elgin, Morayshire, Scotland.*

Scope: The chronological order of certain of Xenophon's works, and light on disputed works. *Method:* Analysis of sentence length, word length, differential sentence length, differential word length, use of relative frequencies of thirty-seven common words including definite article, plus two stylistic features—hiatus and assonance. Above scheme completed and recorded in B.Phil. thesis at St. Andrews University, Scotland (1965). A statistical investigation on the authorship of the Marprelate Pamphlets (1589-90) is now in hand.

Type of computer: Feranti Mercury, KDF 9. [*CHum,* May 1967]

L124. Metrical Analysis of Latin Hexameter

Principal investigator: Wilhelm Ott, Zentrum für Datenverarbeitung, U. of Tübingen, D-74 Tübingen, Köllestrasse 1, West Germany.

Objective: To analyze all Latin dactylic poetry, and those in other meters for metrical and stylistic studies. *Method:* 1) Double transcription of the text from a printed edition on punched cards or paper tape and automatic collation (instead of proofreading). 2) Automatic scansion (about 96 percent accuracy). 3) Correction of the scansion. 4) Automatic analysis which outputs a series of lists of metrical characteristics on magnetic tape and line printer, and a magnetic tape containing the single words of the analyzed work, including complete metrical characteristics and reference. 5) Cumulation of those word-tapes of all the analyzed works; they are the machine-readable basis for further inquiries (also complex ones) by relatively simple and fast programs. 6) Automatic publication of a part of the material by photocomposition.

Type of computer: 1) CDC 3300, 2) DIGISET 50T1 (photocomp). *Size of storage:* 1) 32k words, 2) 32k bytes. *Language:* FORTRAN IV, COMPASS. *Disks or tapes:* 1) 3 7-track, 1 9-track tapes, 2) 1 9-track tape; 1) 1 854 disk. *Special equipment:* Paper tape reader, plotter. Program is available.

References: "Metrical Analysis of Latin Hexameter By Computer," *Revue* (LASLA, Liège) 4 (1966), 7-20; 1 (1967), 39-64; *Metrische Analysen zur Ars Poetica des Horaz,* Göppinger Akademische Beiträge Nr. 6, Göppingen 1970; publication of the material is continued in the Series "Materialien zu Metrik und Stilistik" (Tübingen: Niemeyer-Verlag, 1972ff). The first 3 vols. contain *Aeneid* VI, I, XII; 4 to 6 vols. a year are planned.

L125. An Index to the *Stationers' Register*

Principal investigators: Thomas Y. Remington, Assistant Professor, Chico State College, Chico, CA 95926; William P. Williams, Instructor, Kansas State U., Manhattan, KS 66566.

Objective: To index all proper names and titles in the Eyre and Rivington volumes of the *Stationers' Register,* and, hopefully, to then form a similar index for the Arber volumes.

Type of computer: IBM 1620. *Size of storage:* 20k. *Language:* Assembly. *Disks or tapes:* 1 1311 disk. *Special equipment:* TNS, TNF, NF, automatic Divide. [*CHum,* November 1967]

L126. Computer-Based Research on Linguistic Universals

Principal investigator: Paul L. Garvin, Dept. of Linguistics, SUNY, Buffalo, NY. *Associates:* R.M. Worthy, Gerhard Reitz.

Objective: 1) Basic research in linguistic methodology. 2) Computer-experimentation, using an on-line system, in the simulation of informant work in linguistics. 3) Interpretation of the results with particular emphasis on insights gained by the use of different basic assumptions in experimentation.

Type of computer: Bunker-Ramo 133 and B-R 90 display console. *Disks or tapes:* 16k. *Language:* Machine language. *Disks or tapes:* 1 disk, 2 tapes. [*CHum,* November 1967]

L127. German Expressionist Drama

Principal investigator: R.W. Last, Lecturer, Dept. of German, U. of Hull, Yorkshire, England.

Scope: Complete word indexes to selected dramas of George Kaiser. *Method:* Punch text onto paper tape, sort alphabetically, post-edit, print in tabular form.

Type of computer: ICL 1905E. *Size of storage:* 64k. *Language:* ALGOL/PLAN. *Disks or tapes:* 4 tapes, 4 EDS disks.

Status: The index to selected Kaiser dramas is nearing completion. Concordances are available to the poetry of August Stramm and Hans Arp; they may be made available in line-printer output form on enquiry.

L128. Procedure e programmi di utilità generale per analisi computazionale di testi

Principal investigator: A. Zampolli, direttore della sezione linguistica del CNUCE (Centro Nazionale Universitario di Calcolo Elettronico), Università di Pisa, 56100 Pisa, Italia.

Scope: Procedures for elaborating projects pertaining to input of several tens of millions of words and for writing relative utility programs. *Method:* Immediate purpose is carrying out numerous projects of the Institutes that utilize the computers of Centro Nazionale Universitario di Calcolo Elettronico di Pisa (*Index thomisticus; Vocabolario storico giuridico italiano del CNR: Lessico Rosminiano; Lessico intellettuale europeo; Studio statistico dell'italiano moderno; Trascrizione automatica in alfabetico fonetico e statistiche fonologiche; Concordanze di Seneca; Analisi dei canti popolari italiani,* ecc.). We have written and tested programs for concordances, for lemmatizing with computer dictionary, and for transcribing from historical to phonologic alphabet.

Type of computer: 1) IBM 7090, 2) IBM 360/30, 3) IBM 1401. *Size of storage:* 2) 72k, 3) 16k. *Language:* 1) FORTRAN IV, MAP, 2) Assembler, COBOL, 3) AUTOCODER, IOCS. *Disks or tapes:* 1) 16 tapes, 2) 4 tapes, 2 disks, 3) 4 tapes, 2 disks. *Special equipment:* Catene speciali di stampa.

Status: E stata compilata una serie di programmi di carrattere generale per lo poglio fonematico, lessicale (indici, concordanze, indici inversi, ecc.), sintattico e per la elaborazione statistica di testi in lingua naturale e per la stampa dei dati in fotocomposizione. I programmi sono stati applicati in testi in 18 lingue per circa 50 milioni di parole. Sono state studiate catene speciali di stampa per alfabeto fonetico, latino, greco, ecc.

References: Studi di statistica linguistica acica sulla composizione fonologica e lessicale dell'italiano, Padua, 1960; *Due conversazioni sulla linguistica computazionale,* 1969. [*CHum,* November 1970]

L129. Lessico Intellettuale Europeo

Principal investigator: Tullio Gregory, Professor, Istituto di Filosofia dell'Università di Roma, Rome, Italy.

Scope: Lessici di autore nell'ambito della storia della filosofia (conschedatura elettronica) e un "Lessico filosofico del XVII e XVIII secolo" (conschedatura elettronica e manuale). *Method:* Oltre ai vari lessici particolari in programma si vuole costituire uno archivio di schede contesto e concordanze a documentazione della storia di parole particolarmente importanti nella cultura occidentale. In particolare verranno spogliati elettronicamente testi filosofici dei secoli XVII e XVIII. La direzione del Centro ha deciso di seguire nella elaborazione dei testi i metodi dell'Accademia della Crusca.

Dopo un primo esperimento su un testo campione si è passati alla preparazione del materiale per i vari Lessici. Sono per ora disponibili le concordanze delle opere italiane di Bruno, delle opere di Galilei, dell'*Aesthetica* di Baumgarten; si è invece alla fase della perforazione per quanto riguarda le opere di Vico.

Size of storage: Da stabilire. [*CHum,* November 1967]

L130. Statistiche Linguistiche dell'Italiano

Principal investigators: Carlo Tagliavini, Lucio Croatto, Professors, Istituto di Glottologia e Fonetica dell'Università di Padova, Padua, Italy.

Scope: Testi letterari; dizionari bilingui; periodici e stampa quotidiana; conversazioni registrate; teste psicologici e audiometrici; ecc. *Method:* Si vogliono stabilire le frequenze d'uso delle principali unita (lettere, fonemi, sillabe, parole, costrutti grammaticali, ecc.) costituenti la lingua italiana ai vari livelli ştrutturali (grafemico, fonologico, lessicale, morfemico, ecc.). Particolare riguardo sara riservato alla statistica dei componenti fonologici, per continuare la collaborazione tra gli istituti di glottologia, di foniatria e logopedia, e di elettronica dell'Università di Padova (iniziata con la tesi del Dr. Antonio Zampolli, *Statistica fonologica e lessicale dell'italiano,* Padova, 1960). E in corso di studio l'organizzazione scientifica.

Size of storage: Da stabilire; per ora si iniziera con 600.000 parole. [*CHum,* November 1967]

L131. Tradizioni Popolari Italiane

Principal investigator: Vittorio Santoli, A.M. Cirese, Professors, Università di Cagliari, Cagliari, Italy.

Scope: Raccolte di canti popolari italiani. *Method:* Per ora—Esiste una fondamentale raccolta di canti popolari (raccolta Barbi), allo stato di manoscritti di difficile lettura. Si vogliono ottenere il rimario, l'incipitario, l'indice, liste de frequenze, concordanze e schemi metrici. Probabilmente il materiale, una volta schedato, sara utilizzabile anche per stampare una edizione più agevolmente diffondibile della raccolta stessa. *Per il futuro:* Si prospetta l'estensione del metodo alle altre raccolte di canti già edite, con applicazioni di tipo documentalistico all'*Enciclopedia delle Tradizioni Popolari* in corso di realizzazione. E in corso l'analisi delle procedure e l'organizzazione del lavoro e in esperimento repliminare. *Size of storage:* Per ora, previste 200.000 righe. Per il futuro, diversi milioni di righe. [*CHum,* November 1967]

L132. Antichi Testi Siciliani

Principal investigator: Francesco Ugolini, Professor, U. di Perugia, Perugia, Italy.

Scope: Documenti siciliani delle origini in via di edizione. *Method:* Si vogliono stampare indici di frequenza e concordanze. E stato caricato su nastro e contrallato l'infero testo. *Size of storage:* 20.000 righe. [*CHum,* November 1967]

L133. *Journal* di André Gide

Principal investigator: Carlo Bo, Professor, U. di Urbino, Urbino, Italy.

Scope: Journal di A. Gide. *Method:* Sulla base delle frequenze e delle concordanze, si vuole studiare l'evoluzione del pensiero religioso di Gide, attraverso l'evoluzione della sua terminologia. Sono state eseguite la preedizione, la perforazione, il carico a nastro, e il controllo a lettura del testo. *Size of storage:* 500.000 parole. [*CHum,* November 1967]

L134. Lexical Analysis by Computer of Works and Documents of the Italian Language

Principal investigator: Aldo Duro, Professor, Staff Manager of the *Historical Dictionary of the Italian Language,* Accademia della Crusca, Piazza dei Giudici 1, Florence, Italy. *Associate:* Antonio Zampolli.

Scope: A general historical reference file of Italian language, consisting of about 100 million samples printed and/or punched on cards. To provide material for editing a *Treasury of the Italian Language from the Origins,* as well as editing the *Historical Dictionary of the Italian Language. Method:* Punched cards: one word per occurrence, abstract sentences limited to 10 lines: main entry word, run-on entry, inflected form, reference and dating in print; lemma, text code number and dating punched. Besides providing context cards, concordances, word-frequency lists, and magnetic tape, records will be available to scholars for further research. We are planning to punch and analyze by computer no less than 1000 texts, mainly of the early part of Italian language (i.e., thirteenth and fourteenth centuries).

Type of computer: 1) IBM 1401, 2) IBM 7090. *Size of storage:* 1) 16k, 2) 32k. *Language:* AUTOCODER IOCS, FORTRAN IV, IBSYS v.13. *Disks or tapes:* 15 729 IV, V, VI tapes. *Special equipment:* Special character chain. In the future we shall operate on a special 120-character chain including diacritic marks, etc. *Future plan:* Private hiring of an IBM 360/30 computer. Program is available.

References: A. Duro, "Les Nouvelles methods du dictionnaire historique de la langue italienne," *Cahiers de lexicologie* 8 (1966), 1; "La rinnovata attività lessicografica dell'Accademia della Crusca," *Studi di Filologia italiana* 24 (1966). [*CHum,* November 1967]

L135. Concordances of the Complete Works of Nathaniel Hawthorne, Based on the Centenary Hawthorne

Principal investigators: James J. Owen, Assistant Professor; John R. Byers, Jr., Associate Professor; Dept. of English, Virginia Polytechnic Institute, Blacksburg, VA 24061. *Associates:* Whitney L. Johnson, Dorothy H. Hobbis.

Objective: To publish Hawthorne concordances from the reproduction of each of the volumes of the Centenary Hawthorne, listing the complete line in which each word is found and preserving on tape all stylistic devices such as punctuation and capitalization. *Method:* 1) Keypunch. 2) Transfer to tapes. 3) Sort by selective categories. 4) Print in tabular form. 5) Additional refinements.

Type of computer: IBM 7040/1401. *Size of storage:* 32k. *Language:* FORTRAN IV, MAP. *Disks or tapes:* 1 1301 disk file and 3 729 tapes. [*CHum,* November 1967]

L136. Analysis of Samuel Beckett's *How It Is*

Principal investigator: Howard M. Harper, Jr., Assistant Professor, Dept. of English, U. of North Carolina, Chapel Hill, NC 27514. *Associates:* Hal Fredrickson, Peggy Goldstein, Barbara Snyder, William Hanson.

Scope: Analysis of the text of *How It Is. Method:* 1) Keypunch the text. 2) Derive and print KWIC word index (concordance).

Type of computer: IBM 360/40. *Size of storage:* 256k. *Language:* FORTRAN E and Assembler. *Disks or tapes:* 3 2311 disks, 4 tapes. [*CHum,* November 1967]

L137. Concordances of the Complete Works of Stephen Crane, based on the forthcoming Crane volumes of the American Authors Series

Principal investigators: John R. Byers, Jr., Associate Professor, and James J. Owen, Assistant Professor, Dept. of English, Virginia Polytechnic Institute, Blacksburg, VA 24601. *Associates:* Whitney L. Johnson, Dorothy H. Hobbis.

Objective: To publish Crane concordances from the reproduction of each of the forthcoming Crane volumes of the American Authors Series, listing the complete line in which each word is found and preserving on tape all stylistic devices such as punctuation and capitalization. *Method:* 1) Keypunch. 2) Transfer to tapes. 3) Sort by selective categories. 4) Print in tabular form. 5) Additional refinements.

Type of computer: IBM 7040/1401. *Size of storage:* 32k. *Language:* FORTRAN IV and MAP. *Disks or tapes:* 1 1301 disk and 3 729 tapes. [*CHum,* November 1967]

L138. Literary Computing

Principal investigator: George Whalley, Professor, Dept. of English, Queen's U., Kingston, ON, Canada.

Scope: A comprehensive system for the scholarly editing of literary texts, with a view to extending the heuristic possibilities of the computer in literary scholarship, criticism, and composition.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* PL/I and Assembler. *Disks or tapes:* 2311 and 1 314 disks; 3 9-track tapes and 1 7-track tape. One disk drive and on-line printer required.

Status: The program is in an elementary state on the 1620 but there are plans to rewrite for an IBM 360. The redesign of the files has begun, but as the program is written for the 360, new sections will be added to incorporate processes which could not be accommodated by 1620 facilities. The initial stages of work will eventually become part of a complex of programs to process and display literary data.

References: Computer Society of Canada Quarterly Bulletin, June 1967; Proceedings of *SHARE* 28; *New Cambridge Bibliography of English Literature,* entry for S.T. Coleridge. [*CHum,* November 1967]

L139. A Linguistic Analysis of Selected John Donne Sermons

Principal investigator: Paul W. Mockovak, Dept. of English, State U. Agricultural and Technical Institute, Morrisville, NY 13408.

Scope: A syntactic and morphological study of six Donne sermons spread uniformly throughout Donne's sixteen-year preaching career, a sample of approximately 45,000 words.

Type of computer: IBM 1401. *Size of storage:* 16k. *Language:* FORTRAN and AUTOCODER.

Status: Completed.

L140. Linguistic and Stylistic Analysis of the Fragments of Lucilius

Principal investigator: Adriano Pennacini, Assistant in Latin Literature, Facoltà di Lettere, U. di Torino, via Carlo Alberto 10, Torino, Italy.

Scope: All fragments. *Method:* 1) Analysis by grammatical and syntactic categories. 2) Analysis by lexical components of the vocabulary. 3) Statistical elaboration of results. Automatic analysis by Laboratoire d'Analyse statistique des Langues anciennes, Liège, Belgium.

Type of computer: IBM 1620. *Size of storage:* 20k. *Language:* SPS. *Disks or tapes:* 2 1311 disks. Program is available.

References: "*Docti* e *crassi nella poetica di Lucilio,*" *Atti della Accademia delle Scienze di Torino* 100 (1965-66), 293-360; "Funzioni della rapresentazione del reele nella satira di Lucilio. In appendice il lessico luciano del reele," *Atti della Accademia delle Scienze di Torino* 102 (1968), 311-345.

L141. Textual Criticism

Principal investigator: Jacques Froger (élève diplômé de l'Ecole des Hautes Etudes, Section historique and philologique, Sorbonne), Abbaye Saint-Pierre de Solesmes, 72-Sablé, France.

Scope: Any literary (or even musical) texts. *Method:* 1) Collection of the MSS. 2) Research of their genealogical relations. 3) In normal cases draw the genealogical tree. 4) In abnormal circumstances find their relations of propinquity. 5) Restore the original text.

References: "Emploi de la machine électronique dans les études médiévales," *Bulletin de la société internationale pour l'étude de la philosophie médiévale* 3 (1961), 177-188 (Louvain); "The Electronic Machine at the Service of Humanistic Studies," *Diogenes* 52 (1965), 104-142; "La collation des manuscrits à la machine électronique," *Institut de Recherche et d'Histoire des Textes,* Bulletin No. 13 (1964-65), 135-171; *La critique des textes et son automatisation* (Initiation aux nouveautés de la science, Paris: Dunod, 1967). [*CHum,* November 1967]

L142. Indexed Catalogs of Strophic Forms in Medieval Lyric Poetry

Principal investigator: Bruce A. Beatie, Professor, Dept. of Modern Languages, Cleveland State U., Cleveland, OH 44115. *Associate:* Rudolf Hirschmann.

Objective: To provide a reliable tool for the study of formal traditions within medieval lyric, in the form of thoroughly indexed catalogs of strophic forms of each substantial body of such lyric poetry. *Scope:* Initially, the forms in the *Carmina Burana* and in *Deutsche Liederdichter des 13. Jahrhunderts* (ed. Carl von Kraus); later, other bodies of lyric poetry. *Method:* Strophic schemata will be generated using programs developed for an earlier project (L27); they will be ordered and printed out as a catalog, with supplementary indices to rhyme-schemes, verse-types.

Type of computer: CDC 6400. *Size of storage:* 32k. *Language:* FORTRAN IV.

References: Bruce A. Beatie, "Strophic Form in Medieval Lyric: A Formal-Comparative Study of the German Strophes of the *Carmina Burana* (Unpub. dissertation, Harvard University, 1967). [*CHum,* November 1967]

L143. Stylistic Analysis (author identification) of "Letter from a Virginian to the Members of the Congress"

Principal investigator: Anne Y. Zimmer, Assistant Professor, Dept. of History, Wayne State U., Detroit, MI 48202.

Scope: Letter has been attributed to two authors but is actually an unsigned pamphlet written in 1774. The study was designed to analyze the style of writing along the lines of non-content words used by Mosteller and Wallace in their study of the *Federalist Papers.* It is credited by the Library of Congress to Jonathan Boucher, but has in the past been credited to Jonathan Odell. *Method:* A table of 127 words designed in part by Friedman, but supplemented with some words selected by Mosteller and Wallace, was used to analyze the anonymous document, a document written by Jonathan Boucher, a document written by Jonathen Odell, and a document written by Samuel Seabury.

Type of computer: IBM 1460, IBM 7074. *Size of storage:* 8k. *Language:* FORTRAN.

Status: Completed 1970.

L144. Documentation Littéraire Médiévale

Principal investigator: Edith Brayer, Adjointe au Directeur de l'I.R.H.T. (CNRS Paris), Centre de Recherches et d'Applications Linguistiques, Faculté des Lettres et des Sciences Humaines de l'Université de Nancy, 54-Nancy, France. *Associates:* Française Ferrand, Madame Steunou.

Scope: 1) Répertoire général sur cartes perforées des manuscrits antérieurs au XVIe s. en latin, arabe, hébreu, grec, français, provençal, et espagnol d'après le documentation de l'IRHT à Paris. 2) Inventaires des poésies palinodiques françaises et des chansons castillanes. [*CHum,* November 1967]

L145. Stylistic Analysis of the Poetry of e. e. cummings

Principal investigator: Irene R. Fairley, Assistant Professor, Dept. of English, Northeastern U., Boston, MA 02155. *Associates:* Susumo Kuno, Carol Smith.

Method: Three-line input (text, syntactic analysis, deviance indicator), and concordance on the three lines three-dimensionally. 1) Map out charts for the three-line input. 2) Keypunch two sets of cards for a partial and whole analysis. 3) Concord partials and wholes for lines 2 and 3, and three-dimensional subsort for key items. Trial runs successful; projected sample run with 20 poems.

Type of computer: IBM 7094. *Language:* FORTRAN IV. Original TRICON program is available in S.M. Lamb and L. Gould, *Concordances and Computers* (Berkeley: University of California, 1964).

Status: Discontinued. [*CHum,* November 1967]

L146. Exploitation sur Ordinateur des Chartes Antérieures à 1100

Principal investigator: Lucie Fossier, Archiviste-paléographe, Ingénieur au C.N.R.S., Centre de Recherches et d'Applications Linguistiques, Faculté des Lettres et Sciences Humaines de l'Université de Nancy, 54-Nancy, France. *Associate:* Geneviève Contamine.

Scope: Etude exhaustive de la langue des diplômes, de la diplomatique des actes, de l'onomastique; étude paléographique. *Method:* Enregistrement des chartes sur ordinateur. Constitution d'index verborum. Relevé des formules diplomatiques (préambules, clauses, etc.); constitution d'index onomastique.

Type of computer: IBM 360/75. *Size of storage:* 1024k. *Language:* COBOL. *Disks or tapes:* 2 2314 disks, 8 2400 tapes. Program is available. [*CHum,* May 1971]

L147. Generelle Untersuchungen zur Ventris'schen Entzifferung von Linear B und Analysen zu speziellen Problemen der in diesen Texten ausserhalb von Zeichengruppen stehenden Zeichen

Principal investigators: Dr. habil. Geiss and Dr. Mater, Deutsche Akademie der Wissenschaften zu Berlin, Leipziger Strasse 3-4, 108 Berlin, DDR.*

Scope: Linear-B-Inschriften aus Knossos, Pylos and Mykenai.

Type of computer: SOEMTRON und BULL mit speziellen Umbauten und Programmen. [*CHum,* November 1967]

L148. Construction of an Algorithm for Stem Recognition of the Hebrew Language

Principal investigator: Rabbi Grainom Lazewnik, Associate Research Scientist, Institute of Hebraic Studies, New York U., New York, NY 10003. *Associate:* Philip H. Smith, Jr.

Objective: Reclassification of Hebrew verb roots on basis of morphology and preparation of a noun reference dictionary. *Method:* Programs in PL/I are being used to program the data to enable the computer to abstract the verb root or noun from the given raw test word.

Type of computer: IBM 360/40. *Size of storage:* 250k. *Language:* FORTRAN Version 4. *Disks or tapes:* 2311 disks, 2415 tapes. *Special equipment:* Hebrew Selectric Typewriter. Program is available. [*CHum,* November 1969]

L149. A Concordance to Pindar [*CHum,* November 1967] [Deleted]

L150. Semiautomatic Grammatical Analysis of Dialectal Spoken English

Status: Cancelled. [*CHum,* November 1967]

L151. Homer

Principal investigator: Rev. Andrew Q. Morton, Production Manager, The Abbey Manse, Culross, Fife, Scotland. *Associates:* John Chadwick, W.B. Stanford, A.J. Beattie, W.C. Wake, M. Levison, R. Edwards.

Scope: To provide basis information about Greek verse comparable to what is known about prose, to teach computer techniques to students of Homer, to make a contribution to two problems connected with the text. *Method:* A series of studies: vocabulary, common words, ranking of words, forward and reverse indices, metrical analysis, vowel sequences, participle sequences, article sequences, sentence structure, etc.

Type of computer: Atlas. *Size of storage:* 20k words. *Language:* Machine language and ALGOL. Paper tape input of text, magnetic tape for use. Program is available. Apply Dr. M. Levison, Dept. of Computer Science, Queen's University, Kingston, ON, Canada. [*CHum,* November 1967]

L152. Chinese-English Machine Translation

Principal investigator: David T. Chai, Technical Staff, Dept. of Data Sciences, Bunker-Ramo Corporation, 8433 Fallbrook Avenue, Canoga Park, CA 91304. *Associates:* Constance Hwang, Milton Worthy.

Scope: Research on linguistic analysis of the Chinese and English languages for machine translation. *Method:* Select corpus of Chinese sentences, analyze these for their syntactic structure, construct dictionary containing all lexical entries, construct grammar rules for analyzing Chinese sentences, construct rules for transferring syntactic structure, construct English generation rules, and write computer programs for testing translations.

Type of computer: IBM 7094. *Size of storage:* 32k. *Language:* SNOBOL3. Program is available.

References: Current Research and Development in Scientific Documentation, No.14, National Science Foundation, pp.164-165. [*CHum,* November 1967]

L153. Traitement Automatique des Textes d'Ancien-Français

Principal investigator: Hélène Naïs, Professeur, Centre de Recherches et d'Applications Linguistiques, Faculté des Lettres et des Sciences Humaines de l'Université de Nancy, 54-Nancy, France. *Associates:* Roberte Tomassone, Simone Monsonego, Madeleine Hénin, Odile Derniame.

Scope: Analyse linguistique de textes historiques et poétiques en français du XIIIe au XVIe siècle. *Method:* 1) Transcription diplomatique des manuscrits. 2) Codage d'éléments paléographiques et grammaticaux. 3) Perforation sur cartes. 4)Enregistrement sur bandes. 5) Traitement linguistique. 6) Impression d'index et de résultats sur listings.

Type of computer: IBM 1401 et 7044. *Size of storage:* 16k. *Language:* COBOL. [*CHum,* November 1967]

L154. Formalized Techniques in Machine Translation

Principal investigator: Gary R. Martins, Dept. of Language Analysis and Translation, Bunker-Ramo Corporation, 7259 Jordan Avenue, Canoga Park, CA 91303. *Associate:* Steven B. Smith.

Objective: Design and implementation of a prototype machine translation system (Russian to English) based exclusively upon formalized linguistic techniques. *Method:* The BEAST system (*B*asic *E*xperimental *A*utomatic *S*yntactic *T*ranslator) parses Russian sentences using a vector-symbol phrase grammar; the resulting analyses are mapped, many-to-many rules, into analyses of equivalent English sentences; these latter analyses are then "unparsed" by an output sentence generator employing a quinary CF grammar of English.

Type of computer: BR-133 (AN/UYK-3). *Size of storage:* 16k. *Size of storage:* Interpretive FORTRAN II supplemented by machine language. *Disks or tapes:* Up to 3 IBM 2311 disks, up to 4 BR-170 tapes. Program is available.

References: "A Parsing Procedure for a Vector-Symbol Phrase Grammar of Russian," NSF-C372, Bunker-Ramo Corporation, December 1965; four Contract Status Reports on Contract AF30(602)-3770, Change A, issued September 1, 1966; December 12, 1966; January 10, 1967; and February 1, 1967, by Bunker-Ramo Corporation; "BEAST: The Basic Experimental Automatic Syntactic Translator," AF30(602)-3770, Bunker-Ramo Corporation, July 1967. [*CHum,* November 1967]

L155. Sumerian Computer Concordance

Principal investigator: Miguel Civil, Associate Professor, Dept. of Near Eastern Languages and Civilizations, Oriental Institute, U. of Chicago, 1155 East 58 Street, Chicago, IL 60637.

Objective: To build a basic file (25,000-30,000 lines, with 15-20 morphemes per line) of Sumerian literary texts to be used for a future "Sumerian Dictionary." *Method:* The monolingual texts are processed using the concordance program UNICON, and the bilingual texts (Sumerian Akkadian) by the program DISCON, both created by the Mechano Linguistics Project of the University of California, Berkeley.

Type of computer: IBM 7094/7040. *Size of storage:* 32k. *Language:* FAP, Assembler, FORTRAN II I/O. *Disks or tapes:* 12 729 VI tapes. Program is available. [*CHum,* November 1967]

L156 . Vocabolario storico della lingua italiana

Principal investigator: Aldo Duro, Professor, Accademia della Crusca, Piazza dei Giudici, 1, Firenze, Italy.

Scope and method: Testi e documenti di lingua dalle origini ai giorni nostri; italiano e (entro i limiti dei secoli 13 e 14) dialetti. Scope immediato del progetto è costituzione di un grande archivio della lingua italiana (su nastro magnetico e in forma di schedario a schede contesto), il quale dovrà fornire gli esempi di uso e la documentazione della storia delle parole ai redattori del Vocabolario Storico della Lingua Italiana che il CNR ha affidato alla Accademia della Crusca.

Size of storage: 100 milioni di parole.

Status: La prima fase del ciclo di elaborazioni previste (perforazione e verifica, carico a nastro, controllo a lettura e correzioni, lista delle parole con frequenza, stampa delle concordanze, lemmatizzazione) è in corso e si svolge regolarmente ormai per 150 opere. Al momento attuale è allo studio l'impianto del Dizionario de macchina per la lemmatizzazione semiautomatica. [*CHum,* November 1967]

L157. *Zauberberg* Theme Analysis

Principal investigator: Francis Bulhof, Assistant Professor, Dept. of Germanic Languages, U. of Texas, Austin, TX 78712.

Scope: Word index of Thomas Mann's novel *Der Zauberberg.*

Type of computer: 1) CDC 6600, 2) CDC 1700. *Size of storage:* 1) 131,000, 2) 16,000. *Language:* 1) FORTRAN IV, 2) Assembly. *Disks or tapes:* 1) 2 6608 disks (376 million characters), 6 607 tapes, 2) 1 1738 disk (6 million characters), 3 601 tapes. Program is available.

References: Transpersonalismus und Synchronizität, Wiederholung als Strukturelement in Thomas Mann's "Zauberberg" (Groningen, 1966). (Copies may be obtained directly from author.) [*CHum,* November 1967]

L158. Authorship Attribution to Rabelais' *Fifth Book.* A Computer-Aided Study

Principal investigator: Gerard J. Brault, Professor, Dept. of French, Pennsylvania State U., University Park, PA 16802. *Associate:* Julie Burgoyne.

Scope: Statistical and stylistic comparisons of random samples from the *Fourth Book, Fifth Book,* and other sixteenth-century prose samples. *Method:* 1) Word-frequency counts. 2) Syntactical measurements.

Type of computer: IBM 360/50. *Language:* PL/I.

Status: Still in progress.

L159. Spogli Elettronici dell'Italiano dell Origini e del Duecento

Principal investigator: Mario L. Alinei, Professor, Italian Institute, U. of Utrecht, L. Bolwerk 16 Utrecht, The Netherlands. *Associates:* S. Scalise, R. Crespo.

Scope: A lexical, morphological, and syntactical inventory of all Old Italian texts, from the Placito di Capua (c.960) to Dante's works. *Method:* More than 6 million words, analyzed in part with mark sensing methods, in part with automatic programs.

Type of computer: IBM 1401. Program will be available at end of project.

References: Dizionario Inverso Italiano (The Hague, 1962); *Lessicografia italiana con l'ausilio di macchine all'Un. di Utrecht: Prime esperienze e prospettive* (IV Cong. Ass. Intern. Studi di Lingua e Lett. Italiana, Magonza, 1962); *Grammatical and Statistical Indexing of Italian Texts with the Help of Punched Card Machines at the U. of Utrecht* (Levende Talen. 219); *La lista di frequenza della "Divina Commedia"* (Miscellanea Dantesca, Utrecht 1965); *Spogli Elettronici dell'It. delle Or. e del Duecento,* Vol.1: Introduzione, *Prose fiorentine* (The Hague: Mouton, 1967); Vol.2: *Il Novellino* (Bologna: Il Mulino, 1971); Vol.3: Brunetto Latini, *Rettorica* (Bologna: Il Mulino, 1971); Vol.4: Iacopone, *Laudi* (Bologna: Il Mulino, 1971); Vol.5: Dante Alighieri, *La Commedia* (Bologna: Il Mulino, 1971); Vol.6: *Il Libro dei Sette Savi* (Bologna: Il Mulino, 1971); Vol.7: Chiaro Davanzati, *Rime* (Bologna: Il Mulino, 1971); Vol.8: Dante Alighieri, *La Vita Nuova* (Bologna: Il Mulino, 1971); Vol.9: *Il Bestiario toscano* (Bologna: Il Mulino, 1971); Vol.10: Bono Giamboni, *Della miseria dell'uomo* (Bologna: Il Mulino, 1972); Vol.11: Dante Alighieri, *Rime* (Bologna: Il Mulino, 1972); Vol.12: Bono Giamboni, *Il Libro dei Vizi* (Bologna: Il Mulino, 1972); Vol.13: Dante Alighieri, *Il Convivio* (Bologna: Il Mulino, 1972); Vol.14: Andrea da Grosseto, *Volgarizzamento di Albertano* (Bologna: Il Mulino, 1972); Vol.15: *Prose Sangimignanesi* (Bologna: Il Mulino, 1972).

L160. Dictionary of Names and Titles in Poe's Collected Works

Principal investigator: Burton R. Pollin, Professor, Dept. of English, Bronx Community College, CUNY, Bronx, NY 10468. *Associate:* Gary Berlind.

Scope: Processing over 10,000 entries derived from the 15 volumes of Harrison edition, producing 250 pages of edited entries, in four listings. *Method:* 1) Keypunching indexed material. 2) Extracting two of the four indices through delimiters. 3) Formatting through program control page proofs in external storage. 4) On-line editing of page proofs. 5) Setting of prose introduction into justified page proofs through Gerald Bernes's FORMAT program.

Type of computer: IBM 360/30. *Size of storage:* 64k. *Language:* Basic Assembler Language. *Disks or tapes:* 3 2311 disk drives and 2321 data cell. Program is available.

Status: Completed.

References: Dictionary of Names and Titles in Poe's Collected Works (New York: DaCapo Press, 1968).

L161. *Dictionarium Tetraglotton sev Voces Latinae Omnes, et Graecae eis respondentes, cum Gallica & Teutonica (quam Passim Flandricam vocant) earum interpretatione,* Antwerp, 1562

Principal investigator: F. de Tollenaere, Institute for Dutch Lexicology, section Thesaurus, Stationsweg 39a, Leiden, Beatrixlaan 7, Warmond, Holland. *Associates:* Marineke Mulder-Frijters, J.B. Veerbeek, F. Claes.

Objective: To reproduce the old book and to make an alphabetical list of the Dutch words of the dictionary. *Method:* 1) Keypunch. 2) Transfer to tapes. 3) Alphabetical sorting. 4) Print in tabular form.

Type of computer: IBM 360/50. *Language:* PL/I.

Status: Completed May 1972.

References: F. Claes, F. de Tollenaere and J. B. Verbeek, *Het Tetraglotton van 1562* (in *Monumenta Lexicographica Hollandica* Series II), Parts I and II ('s-Gravenhage: Mouton, 1972).

L162. Computer-Assisted Linguistic Analysis of the Prose Style of John Dryden

Principal investigator: Harry W. Peter, Dept. of English, Georgetown U., Washington, DC 20007.

Scope: Corpus limited to Dryden's essays; exact size will be determined by statistical necessity after initial samples are processed.

Method: 1) De-transformational history manually made for each sentence. 2) Presence-absence of various syntactic structures noted; labels indicating lexical register and function in paragraph structure assigned. 3) Material processed to determine frequency, proportion, patterns.

Type of computer: IBM 360/50. *Language:* PL/I, FORTRAN. [*CHum,* May 1968]

L163. Cato the Elder

Principal investigator: Stephen V.F. Waite, Kiewit Computation Center, Dartmouth College, Hanover, NH 03755.

Scope: Entire corpus of Cato, including fragments (about 20,000 words). *Method:* 1) Enter and correct originals on time-sharing. 2) Produce preliminary word-lists and sentence-length charts for 1200 word sections (approximately). 3) Produce card output. 4) Get complete word-list in batch-processing. 5) Get statistics on word length, sentence length, repetitiveness of vocabulary and entropy, and clausulae. 6) Compare with results obtained from different works, with additional comparisons to sample texts of Sallust and Livy.

Type of computer: Honeywell 635 (formerly GE 635). *Size of storage:* 160k (16k partition used for all time-sharing programs). *Language:* BASIC 5 (time-sharing); GECOS FORTRAN (batch). Key programs in the series are available in "as-is" condition.

Status: Completed June 1969.

References: For input format and texts used, see *Computers and the Humanities* 2, 3 (January 1968), 135; "A Computer-Assisted Study of the Style of Cato the Elder with Reference to Sallust and Livy," *Harvard Studies in Classical Philology* 74 (1970), 348-349; "Approaches to the Analysis of Latin Prose, Applied to Cato, Sallust, and Livy," R.E.L.O. *Revue* 2 (1970), 91-120.

L164. Stylostatistical Analysis of Anglo-Saxon Battle Poetry to Determine Effect of Theme on Syntax

Principal investigator: T.J. Ray, Professor, Dept. of English, U. of Mississippi, University, MS 38677.

Scope: Random samples (48,000 words) from Anglo-Saxon poetry containing both battle and non-battle theme units. *Method:* 1) Edit for syntactic units and patterns. 2) Keypunch. 3) Transfer to tapes. 4) Sort by units and patterns. 5) Print concordance and separate listings by particular patterns and units.

Type of computer: IBM 360/30. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 4 2311 disk drives. Program is available.

L165. A Computer Investigation of *Pericles*

Principal investigator: Tommy Ruth Waldo, Associate Professor, Dept. of English, U. of Florida, Gainesville, FL 32601.

Scope: Stylistic comparisons of *Pericles,* parts of which are attributed to Shakespeare, and *Antony and Cleopatra,* a work near in date to *Pericles* and known to be by Shakespeare. The purpose is to study the degree of Shakespeare's participation in the composition of *Pericles. Method:* Keypunch texts. Alphabetize words and transfer them to tape. Add coded information about frequency, syntax, and meaning. Sort words in various orders to facilitate comparisons of several elements. Other procedures will be added later.

Type of computer: IBM 360/65. *Language:* PL/I.

References: "*Pericles* and Thomas Wright's *Passions of the Mind,*" Southeastern Renaissance Conference, University of Kentucky, April 6-7, 1973; "*Pericles* and the Question of Structural Unity," International Conference on Computers in the Humanities, University of Minnesota, July 20-22, 1973.

L166. La Rochefoucauld Concordance

Principal investigator: Bryant C. Freeman, Professor, Chairman of Dept. of French and Italian, U. of Kansas, Lawrence, KS 66044. *Associate:* Lynne Huxsaw.

Scope: Maximes and *Réflexions diverses,* with all variants. *Method:* 1) Paper tape. 2) Transfer to tapes (magnetic). 3) Photon printout (variable spacing).

Type of computer: Burroughs B5500. *Size of storage:* 244k. *Language:* Burroughs Extended ALGOL 60.

References: Concordance du théâtre et des poésies de Jean Racine, 2 vols. (Ithaca, NY: Cornell University Press, 1968).

L167. Analysis of Syntax and Grammar of Fiction Prose Style

Principal investigator: Karl Kroeber, Professor, Dept. of English, Columbia U., New York, NY 10027. *Associate:* Gary Goshgarian.

Scope: Samples (2000-11,000 words) from approximately 50 English novels from Fielding to Graham Greene. *Method:* 1) Mark and edit all words in each sample. 2) Keypunch. 3) Transfer to tapes. 4) Analyze according to approximately 25 tests. 5) Print out results in tabular form.

Type of computer: CDC 3600. *Size of storage:* 64k. *Language:* FORTRAN IV. *Disks or tapes:* 10 606 tape units.

References: A Computer Analysis of Fictional Prose Style (Washington, DC: U.S. Office of Education, 1966); "Computers and Research in Literary Analysis," *Computers in Humanistic Research,* ed. Edmund A. Bowles (Englewood Cliffs, NJ: Prentice-Hall, 1967); "The Perils of Quantification," *Statistics and Style,* ed. R.W. Bailey and L.Dolozel (Ann Arbor: University of Michigan Press, 1968). [*CHum,* May 1968]

L168. A Statistical Investigation of the Prose Writings of Henry Fielding and His Contemporaries

Principal investigator: Michael G. Farringdon, Dept. of Computer Science, University College of Swansea, Singleton Park, Swansea, Glamorganshire, U.K.

Scope: To try to find stylistic discriminators in the works of Fielding and some of his contemporaries and to use these on some disputed works. *Method:* 1) Keypunch. 2) Transfer to disks. 3) Derive and print word counts and concordances. 4) Additional refinements.

Type of computer: ICT 1904 and 1906A. *Language:* POP-2 and ALGOL 68.

L169. *Index Verborum* to Xenophon's *Anabasis*

Principal investigator: Cordelia M. Birch, Professor, Dept. of Greek and Latin, 1318 Fourth Ave., Beaver Falls, PA 15010. *Associate:* W. Paul Arnold.

Scope: Complete with references to each section where found. *Method:* Listing of every word used with a line of context for all of Xenophon. All five volumes of Oxford Classical Text keypunched with full symbols. *Anabasis* is completely keypunched without diacritics, subscripts and capitals; it will be redone with all symbols after the other volumes are completed.

Type of computer: IBM 1130. *Size of storage:* 8k. *Language:* Basic FORTRAN IV. *Disks or tapes:* 1 removable disk.

Status: Vols.1 and 3 sent to Dartmouth; vols.2 and 5 proofreading completed; vol.4 still being proofread.

L170. Stylography of Eighteenth-Century English Novels

Principal investigator: Edward M. Jennings, Assistant Professor, Dept. of English, SUNY, Albany, NY 12222.

Scope: With initial sections of *Pilgrim's Progress, Pamela, Tom Jones,* and *Tristram Shandy,* to seek, through analysis of simplified syntactic and semantic patterns, macro-stylistic discriminators of kinds of perceptions of experience. *Method:* 1) Numeric coding of semantic classes in sequence. 2) Machine search for repeated patterns. 3) Evaluation.

Type of computer: CDC 3200. *Language:* FORTRAN IV. Program is available.

Status: Completed. Evaluation in progress.

L171. Automatic Syntactic Analysis and Synthesis of English Sentences

Principal investigator: Herbert Pilch, Professor, Englisches Seminar, Albert-Ludwigs U., D 78 Freiberg, Rottekring 4, Germany. *Associates:* Klaus Detering, Lilo Moessner.

Objective: Machine translation. *Scope:* Automatic translation of English sentences into a target language, preferably Russian. We work on automatic syntactic analysis and synthesis. These we consider to be the prerequisites for automatic translation. Programs for automatic sentence generation and automatic analysis of nominal groups within English sentences have been worked out. We plan to concentrate on automatic analysis of verbal groups as well as on an experimental translation program of nominal groups from English to Russian. *Method:* Splitting up of the translation process into three parts: analysis of complex structures in the object language by reducing them to elementary structures; translation of the elementary structures of the object language into those of the target languages; building up of complex structures in the target language out of its elementary structures. Utterances are understood as sequences of utterance units—programs are projected which analyze the different types of utterance unit. An overall program is to govern the sequence of application of these minor analysis routines.

Type of computer: IBM 7094. *Language:* FORTRAN II. Program is available.

References: K. Detering, Automatische Erzeugung englischer Satze (erscheint); L. Moessner, Automatische Analyse englischer nominater Gruppen (erscheint); H. Pilch, "Der Stand der Freiburger Arbeiten zur maschinellen Ubersetzung," *Syntax u. Datenverarbeitung* 2 (1966), 82-117; U. Oomen, "Automatische syntaktische Analyse," *Janua Linguarum,* Series minor 76, den Haag 1968; L. Moessner, "Strukturtypen englischer nominaler Gruppen," *Beiträge zur Linguistik und Informationsverarbeitung,* Heft 18 (197?), 41-65.

L172. Documents pour l'étude des textes littéraires français: Index et concordances

Principal investigator: Bernard Quémada, Directeur, Centre d'Etude du Vocabulaire française, Faculté des Lettres et Sciences Humaines,18, rue Chifflet, 25-Besançon, France.

Scope: Textes de référence du XIVe siècle à nos jours. *Method:* 1) Bandes perforées. 2) Cartes mécanographiques. 3) Listages + analyse et études critiques. 4) Impression mécanographique.

Type of computer: BULL.

Status: Programme continu. [*CHum,* May 1968]

L173. A Concordance to the Poetry of Jonathan Swift

Principal investigator: James M. Pye, Research Assistant, Data Processing Center, Texas A & M U., College Station, TX 77843.

Scope: All poetry from Williams' 1958 Oxford Edition, including those verses which are of doubtful authorship and those verses by other poets which are in the collection. *Method:* 1) Manually transfer to textual data coding forms. 2) Symbolically designate typefont conventions. 3) Keypunch. 4) Transfer to tape. 5) Produce concordance and associated statistics.

Type of computer: IBM 360/65. *Size of storage:* 512k. *Language:* COBOL. *Disks or tapes:* 1 2314 D.A.S.F., 4 2401/2 tapes, 1 2402/2 tape, 1 2402/3 tape. *Special equipment:* 2260 display terminal. [*CHum,* May 1968]

L174. Computer-Aided English Translation of Petrarch's Latin Letters

Principal investigator: Aldo S. Bernardo, Professor, Dept. of Romance Languages, SUNY, Binghamton, NY 13901.

Objective: To translate primarily Petrarch's *Familiares* and *Seniles* with the aid of word-lists generated by a computer and controlled in such a manner that a Latin-English dictionary of Petrarch's Latin, in addition to an in-context concordance, would result. This has been accomplished for the *Familiares,* and translating from a printout is now in process, listing each sentence in column form on the left side of the sheet, and from one to five English meanings for each word extending toward the right side of the sheet. The present project is to proceed in the same manner as the *Seniles* but the computer will extract the English meaning for each Latin word.

Type of computer: IBM 360/40. *Size of storage:* 256k. *Language:* COBOL DS. *Disks or tapes:* 5 2311 disk drives, 1 2402 tape. *Special equipment:* 1050/2260 terminals. [*CHum,* May 1970]

L175. Concordances to *Osman* by Ivan Gundulic

Principal investigator: Zeljko Bujas, Odsjek sa anglistiku, Filozofski fakultet, Zagreb, Yugoslavia.

Scope: The entire epic (13,500 lines; 65,000 words) in seventeenth-century Croatian. *Method:* 1) Keypunch (transliterating 5 letters with diacritics and 3 accents). 2) Transfer to tapes. 3) Produce KWIC ordinary and reverse concordances. 4) Produce glossary with frequencies and vocabulary rank list. 5) Compile a rhyming dictionary.

Type of computer: 1) CDC 1700, 2) CDC 6600. *Size of storage:* 1) 16k, 2) 131k. *Language:* FORTRAN, Assembly. *Disks or tapes:* 1) 1 disk, 3 tapes, 2) 4 disks, 6 tapes. Program is available for steps 3 and 4. [*CHum,* May 1968]

L176. Concordances to *Povratak Filipa Latinovicza* by Miroslav Krleza

Principal investigator: Zeljko Bujas, Odsjek za anglistiku, Filozofski fakultet, Zagreb, Yugoslavia.

Scope: The entire original in Croatian and its translation into English (2800 + 2800 sentences). *Method:* 1) Keypunch (transliterating 5 letters with diacritics in Croatian text). 2 Transfer to tapes. 3) Produce KWIC ordinary and reverse concordances in both languages. 4) Produce glossaries with frequencies and vocabulary rank lists for both texts. 5) Produce ordinary and reverse contrastive (i.e., two-line two-language) concordances, Croatian-item and English-item ordered.

Type of computer: 1) CDC 1700, 2) CDC 6600. *Size of storage:* 1) 16k, 2) 131k. *Language:* FORTRAN, Assembly. *Disks or tapes:* 1) 1 disk, 3 tapes, 2) 4 disks, 6 tapes. Program is available for steps 3 and 4. [*CHum,* May 1968]

L177. Research in German-English Machine Translation on Syntactic Level

Principal investigator: Winfred P. Lehmann, Professor and Director, Linguistics Research Center, U. of Texas, Austin, TX 78701. *Associate:* L.W. Tosh.

Scope: The study is directed toward the development of a German-English mechanical translation system capable of producing translations of German scientific and technical texts on the syntactic level. *Method:* Two major objectives guide the research: 1) German-English transfer grammar linking descriptions of the two languages and 2) a complete description, including programming systems, which will generate mechanical translation of texts.

Type of computer: 1) CDC 1700, 2) CDC 6600. *Size of storage:* 1) 16k, 2) 13k. *Language:* FORTRAN, Assembly. *Disks or tapes:* 1) 1 disk, 3 tapes, 2) 4 disks, 6 tapes. *Special equipment:* Cathode-ray-tube conversational mode station. Program is available.

Status: Completed.

L178. Study of Russian Verb Auxiliaries

Principal investigator: Raoul N. Smith, Instructor, Dept. of Linguistics, Northwestern U., Evanston, IL 60201.

Scope: Modal auxiliaries appearing in a 200,000-word corpus of contemporary standard Russian. *Method:* Concordance.

Type of computer: CDC 6400. *Language:* FORTRAN. Program is available, as well as copies of concordance of whole corpus.

L179. An Investigation into the Vocabulary of English Dictionaries and Newspapers

Principal investigator: Thomas Finkenstaedt, Professor, Anglistisches Institut, Universität des Saarlandes, 66 Saarbrücken 11, Germany. *Associate:* H. Joachim Neuhaus.

Scope: A statistical investigation into the range of the English language, its etymological composition, the parts of speech and dates of the first appearances of the words in the language. *Method:* The investigation is based on the *Shorter Oxford English Dictionary,* the *Advanced Learner's Dictionary,* and Michael West's *General Service List of English Words.* The body of about 80,000 words is being compared with English newspaper texts.

Type of computer: Electrologica X1. *Size of storage:* 32k. *Language:* Machine code. *Disks or tapes:* Sandwich tape, Scotch 789. Program is available.

References: D. Wolff, "An Investigation into the Vocabulary of English Dictionaries and Newspapers: A Preliminary Report," *Revue* (LASLA) II, 1967, 1-13; Dieter Wolff, *Statistische Untersuchungen zum Wortschatz Englischer Zeitungen,* Saarbrücken, 1969; Joachim Neuhaus, "Studies in Rare, Archaic, and Obsolete Words in *The Shorter Oxford English Dictionary*" (Master's thesis, U. des Saarlandes, Saarbrücken, 1969); Thomas Finkenstaedt, "Uber einige Erstbelege in der englischen Literatur. Ein Beitrag zur Statistik des Sprachschöpferischen," *Literature und Datenverarbeitung,* ed. H. Schanze (Tübingen: Niemeyer, 1972); Thomas Finkenstaedt and Dieter Wolff, *Ordered Profusion—Studies in Dictionaries and the English Lexicon,* with contributions by H.J. Neuhaus and W. Herget (Heidelberg, Winter 1973).

L180. Glossarial Index to Middle English *Morte Arthure*

Principal investigators: Bruce Steele, Senior Lecturer, and Jennifer Strauss, Lecturer, Dept. of English, Monash U., Clayton, Victoria, Australia 3168. *Associate:* Georgette Silva.

Scope: Complete alphabetical word list. *Method:* 1) Keypunch text onto 8-channel paper tape. 2) Transfer to magnetic tape. 3) Produce word index and word-frequency count.

Type of computer: CDC 3200. *Size of storage:* 32k. *Language:* FORTRAN. *Disks or tapes:* Two 4 x 10 million character disks (CDC); 4 tapes (IBM style). *Special equipment:* 30" incremental plotter. Program is available.

References: G.M.T. Silva, *Concordance Generation on the CDC 320 Computer* (Unpub. thesis, Monash University, 1966—expanded and revised version of concordance-generating procedures for the CDC 3200 in form of *User's Manual* available from Monash University Computing Center).[*CHum,* May 1968]

L181. Edition of the Plays of Thomas Southerene (1660-1745)

Principal investigator: Harold H.R. Love, Senior Lecturer, Dept. of English, Monash U., Clayton, Victoria, Australia 3168. *Associate:* Peter Hohaus.

Scope: Four texts of *The Loyal Brother* and six of *The Wives' Excuse* (others to follow). *Method:* 1) Punch on paper tape. 2) Transfer to magnetic tape. 3) Identify corresponding lines, additions, omissions, etc. 4) Print variants of derived texts.

Type of computer: CDC 3200. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* Two 4 x 10 million-character disks (CDC); 4 tapes (IBM style). *Special equipment:* 30" incremental plotter. Program is available.

References: See L180.

L182. Critical Index of Columella

Principal investigator: George G. Betts, Dept. of Classical Studies, Monash U., Clayton, Victoria, Australia 3168. *Associates:* P.M.Waterson, W.D. Ashworth, Georgette Silva.

Scope: Systematic grammatical analysis (including syntax) combined with manuscript attribution of doubtful readings. *Method:* 1) Precode text. 2) Paper tape (Flexowriter). 3) Transfer to magnetic tape. 4) Sort for concordance.

Type of computer: CDC 3200. *Size of storage:* 32k. *Language:* FORTRAN. *Disks or tapes:* Two 4 x 10 million-character disks (CDC); 4 tapes (IBM style). *Special equipment:* 30" incremental plotter. Program is available.

Status: Completed 1971 and published as *Index to the Uppsala Edition of Columella* by G.G. Betts and W.D. Ashworth (Uppsala: Studia Latina Upsaliensia 6, 1971).

References: See L180.

L183. Procédés Stylistiques de Nathalie Sarraute

Principal investigator: Roland Husson, Lecturer, Dept. of Modern Languages, French section, Monash U., Wellington Road, Clayton, Victoria, Australia 3168. *Associate:* Georgette Silva.

Scope: Tenter de définir certains aspects du vocabulaire et du style de Nathalie Sarraute en interprétant notamment les renseignements quantitatifs que peuvent fournir des ordinateurs. *Method:* 1) Keypunch text onto 8-channel paper tape for transfer to magnetic tape. 2) Make word index, word-frequency count and concordances.

Type of computer: CDC 3200. *Size of storage:* 32k. *Language:* FORTRAN. *Disks or tapes:* Two 4 x 10 million-character disks (CDC); 4 tapes (IBM style). *Special equipment:* 30" incremental plotter. Program is available.

References: See L180.

L184. Indonesian (Newspaper) Word Count

Principal investigator: R. Rabinranat Hardjadibrata, Lecturer, Dept. of Modern Languages, Indonesian and Malay Section, Monash U., Wellington Road, Clayton, Victoria, Australia 3168. *Associate:* G. Silva.

Scope: 1) To establish a frequency list of the 2000 most frequent words in "newspaper" Indonesian. 2) To establish the morphemic structure of Indonesian (in paradigmatic fashion). 3) To establish the structure of the Indonesian phrase. *Method:* 1) Keypunch current nationally read newspapers onto paper tape for storage on magnetic tape. 2) Correction of errors. 3) Production of word index. 4) Production of word-frequency count. 5) Study of usage, frequency, and dispersion.

Type of computer: CDC 3200. *Size of storage:* 32k. *Language:* FORTRAN. *Disks or tapes:* 2 disks, 6 tapes. *Special equipment:* 30" incremental plotter. Program is available.

References: See L180. [*CHum,* May 1968]

L185. A Concordance of the Early Works of Friedrich Schiller

Principal investigator: Robert L. Kahn, Professor, Dept. of Germanics, Rice U., Houston, TX 77001. *Associates:* Frederick Ruecking, Jr., Karl-Heinz Boewe, Klaus Neuendorf, Gudrun Heinsen, James McLeod.

Scope: Complete word list. *Method:* 1) Keypunch text. 2) Load and keyword index. 3) Sort. 4) Print temporary word list. 5) Update and correct. 6) Sort. 7) Print final word list.

Type of computer: IBM 7040. *Size of storage:* 32k. *Language:* COBOL. *Disks or tapes:* 4 729 V tape drives. [*CHum,* May 1968]

L186. Russian Linguistic Dictionary in Morpho-Phonemic Form

Principal investigators: Walter D. Lazear, Project Head, and George W. Smalley, Chairman, Dept. of Slavic Languages, Lawrence U., Appleton, WI 54911.

Scope: Eventually, 80,000-word vocabulary in morpho-phonemic form. *Method:* 1) Create morpho-phonemic forms. 2) Keypunch. 3) Determine linguistic requirements. 4) Edit data with program. 5) Print in list form. 6) Evaluate data on list.

Type of computer: IBM 1620 (Model 1). *Size of storage:* 40k. *Language:* AFIT SPS. Program is available. [*CHum,* May 1968]

L187. Structural Analysis of Hebrew Roots

Principal investigator: David E. Y. Sarna, IBM (Israel) Ltd., P.O. Box 3058, Haifa, Israel. *Associate:* Lawrence H. Schiffman.

Objective: To investigate intolerable letter combinations, in Biblical Hebrew, and formulate thereby rules for determining admissibility of questionable or putative nominal roots. We also hope to examine later Hebrew for deviation from the formulated rules, in the hope of observing the language development. *Hapax Legomona* will also be examined in the light of the rules formulated. Presently limited to Biblical Hebrew. To include Biblical Aramaic, possibly Ugaritic, and possibly Modern Hebrew. *Method:* 1) Compile an accurate root list. (Sources: Lexica of Brown, Driver and Briggs, Baumgartner and Koehler, Even Shoshan, and Ben Yehuda, with reference to Mandelkern's *Concordance.*) 2) Keypunch. 3) Sort for single- and multiple-letter presence. 4) Sort for presence by location. 5) Sort for presence and frequency of two-letter combinations. 6) Print out in tabular form both presence and absence lists. 7) Formulate presence-absence rules. 8) List letter combinations presumed intolerable in Biblical Hebrew. 9) Examine low-frequency combinations for possible foreign etymology, which would be suggested by deviation from the rules formulated in 7 above. 10) Examine presumed putative nominal roots for possible deviation from above rules. 11) Compare with associated Semitic languages to observe possible similarity in structure. 12) Compare with Modern and Mishnaic Hebrew to observe language development. 13) Examine *sin-shin-tav* for possible clues as to origin.

Type of computer: IBM 1620. *Size of storage:* 40k core (char.) 2 disks. *Language:* FORTRAN 2D, SPS II D, SNOBOL3. *Special Equipment:* Hebrew Selectric element for listing. Output is available for cost of reproduction.

Status: Completed 1968.

L188. Computer-Aided Manuscript Study

Principal investigators: David E. Y. Sarna, IBM (Israel) Ltd., P.O. Box 3058, Haifa, Israel; Lawrence H. Schiffman, Dept. of Near Eastern and Judaic Studies, New York U., 50 Washington Square South, New York, NY 10003.

Objective: To maximize the automatic processing of Rabbinic texts to aid in manuscript comparison; to prepare critical editions giving textual variants, photo-typesetting, concordancing, statistical analysis.

Type of computer: 1) IBM 360/65, 2) IBM 1130. *Size of storage:* 1) 365k partition, 2) 16k. *Language:* 1) SNOBOL4, 2) FORTRAN IV, ASM. *Special equipment:* High-speed Hebrew print chain, low-speed Hebrew, Greek, and Cyrillic print balls.

References: The Toseftah Sotah: A Computer-Aided Critical Edition (Jerusalem: Makor Press, 1973).

L189. Computerized Collation of Medieval Prose Manuscripts

Principal investigator: Anne-Lise Cohen, Dept. of History, U. of California, 405 Hilgard Avenue, Los Angeles, CA 90024. *Associate:* Peggy Cabannis.

Scope: Word for word, sentence for sentence, paragraph for paragraph (100 words) comparison of two texts. Prints out original and variants. *Method:* Card input, one line of text per card, stored on disk.

Type of computer: IBM 360/75. *Language:* 2 versions of program: COBOL and PL/I. [*CHum,* May 1968]

L190. A Method of Structural Analysis with an Application to *Les Liaisons dangereuses*

Principal investigator: Christiane B. Allais, Instructor, Dept. of French, U. of California, Berkeley, CA. *Associate:* Claude Allais.

Method: Mathematical method for determining the structure of literary works. Graphical representation of the structure.

Type of computer: IBM 7094. [*CHum,* May 1968]

L191. A Study of the Farce-Afterpiece on the London Stage, 1700-1725

Principal investigator: Raja P. Ramdass, graduate student, Dept. of Drama, U. of Southern California, 1010 W. Jefferson Blvd., Los Angeles, CA 90007. *Associate:* James H. Butler.

Scope: To analyze the daily calendar of performances on the London stage (1700-1725) and establish the relative popularity of the different kinds of afterpieces. *Method:* Different afterpieces presented nightly are coded as to their title, type, and authorship from *The London Stage,* and a frequency distribution is presented for each season, type, and play. Analysis of variance is performed on the data.

Type of computer: Honeywell 800. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 8 tape drives. High-speed printer. Program is available.

References: ICRH Newsletter II (February-March, 1967), 2. [*CHum,* May 1968]

L192. A Concordance to the Poems of Sir Philip Sidney

Principal investigator: Herbert S. Donow, Associate Professor, Dept. of English, Southern Illinois U., Carbondale, IL 62901.

Objective: To produce a concordance to the poems of Sir Philip Sidney based on the Oxford edition, edited by William A. Ringler, Jr., with indices in modern American spelling and text in original spellings.

Type of computer: Some programs were designed for use on an IBM 7044; others for the IBM 360/65. Programs are either obsolete or too specialized to be of much use.

Status: Text of poems and the concordance are on tape and finished concordance is awaiting publication by Cornell University Press.

References: "A Concordance and Stylistic Analysis of Elizabethan Sonnet Sequences," *Computers and the Humanities* 3, 4 (March 1969), 205-208.

L193. Syntaktische Analyse frühneuhochdeutscher Prosa

Principal investigator: Ludwig Erich Schmitt, Ordentlicher Professor, Forschungsinstitut für Deutsche Sprache, Deutscher Sprachatlas, U. Marburg, 355 Marburg/Lahn, Krummbogen 28A, Germany. *Associates:* Monika Hager, Susanne Mumm, Bernd S. Müller.

Scope: Johann Eberlin von günzburgs 15 Bundesgenossen (c. 66.000 Wörter). *Method:* 1) Übertragung des Originaltextes auf 8-Kanal-Lochstreifen. 2) Syntaktische Klassifizierung in mehreren Zeilen. 3) Alphabetisches Wortregister mit gleichzeitiger tabellarischer Zuordnung der syntaktischen Kategorien und Stellenangaben. 4) Sortierung nach den syntaktischen Kategorien mit Zuordnung der Textwörter.

Type of computer: TR4. *Size of storage:* 32k. *Language:* Maschinencode TEXAS. *Disks or tapes:* Magnetbandeinheiten. Program is available.

References: Publikation des Index mit den Klassifikationen, ende 1968. [*CHum,* May 1968]

L194. Wortindex zu Gottfried Kellers Novellenzyklus *Die Leute von Seldwyla*

Principal investigator: Ludwig Erich Schmitt, Ordentlicher Professor, Forschungsinstitut für Deutsche Sprache, Deutscher Sprachatlas, U. Marburg, 355 Marburg/Lahn, Krummbogen 28A, Germany. *Associates:* Monika Hager, Niels Sörensen. In Zusammenarbeit mit dem Deutschen Rechenzentrum, Darmstadt, Rheinstrasse 75.

Scope: Zwei Bände von Gottfried Kellers Werken (c. 100.000 Wörter). *Method:* 1) Übertragung des Originaltextes auf 8-Kanal-Lochstreifen under Kennzeichnung der Stellen mit Varianten. 2) Vollständiges alphabetisches Wortregister mit absoluten Häufigkeiten und Häufigkeiten pro Seite. 3) Restriktive Stellenkonkordanz. 4) Rückläufiges Wortregister. 5) Häufigkeitsregister. 6) Namenregister. 7) Kontextregister nach Bedarf.

Type of computer: IBM 7090. *Size of storage:* 32k. *Language:* FORTRAN, FAP. Erfordert Rücksprache mit dem Deutschen Rechenzentrum.

References: Publikation des Index geplant, ende 1968.[*CHum,* May 1968]

L195. Wortindex zu Georg Büchners Werken

Principal investigator: Ludwig Erich Schmitt, Ordentlicher Professor, Forschungsinstitut für Deutsche Sprache, Deutscher Sprachatlas, U. Marburg, 355 Marburg/Lahn, Krummbogen 28A, Germany. *Associates:* Monika Hager, Niels Sörensen. In Zusammenarbeit mit dem Deutschen Rechenzentrum, Darmstadt, Rheinstrasse 75.

Scope: Die Dichtungen Georg Büchners. *Method:* 1) Übertragung des Originaltextes auf 8-Kanal-Lochstreifen unter Einbeziehung der verschiedenen Fassungen. 2) Vollständiges alphabetisches Wortregister mit absoluten Häufigkeiten und Häufigkeiten pro Seite. 3) Restriktive Stellenkonkordanz. 4) Rückläufiges Wortregister. 5) Häufigkeitsregister. 6) Namenregister. 7) Kontextregister nach Bedarf.

Type of computer: IBM 7090. *Size of storage:* 32k. *Language:* FORTRAN, FAP. Erfordert Rücksprache mit dem Deutschen Rechenzentrum.

References: Publikation des Index, Anfang 1969. [*CHum,* May 1968]

L196. The Application of Automated Content Analytic Techniques to the Study of *Paradise Lost*

Principal investigator: Michael J. Zieky, Educational Testing Service, Princeton U., Princeton, NJ 08540. *Associates:* G. Fisher, J. Hiller.

Scope: Complete text of *Paradise Lost. Method:* 1) Map text into categories. 2) Collect scores across character. 3) Run multiple discriminant and classification analyses. 4) Interpret.

Type of computer: IBM 360/65. *Size of storage:* 256k. *Language:* PL/I F. *Disks or tapes:* 4 disks, 1 drum. Program was available September 1968. [*CHum,* May 1968]

L197. Mechanical Recognition of Linguistic Structures

Principal investigator: Hans Karlgren, Research Group for Quantitative Linguistics, Fack, Stockholm 40, Sweden. *Associates:* Benny Brodda, Göran Engström, Bengt Svensson.

Method: Combinatorial analysis, tested on Swedish sentences and lexicological material.

Type of computer: TRASK and CDC 3600. *Size of storage:* 24k (TRASK). *Language:* ALGOL plus string processing procedures.

References: Reports at Grenoble and Tokyo ILA; Hans Karlgren, "Slant Grammar Calculus," *KVAL PM* 337; Benny Brodda, "The Entropy of Recursive Markov Processes," *KVAL PM* 339; "Citation Index and Measures of Association in Mechanized Document Retrieval," *KVAL PM* 295. [*CHum,* May 1968]

L198. A Concordance to the *Second Shepherds' Play*

Principal investigator: Everett G. Powell, Associate Professor, Dept. of English, Del Mar College, Corpus Christi, TX 78411. *Associate:* Jon Wiley.

Scope: Concordance, frequency-of-appearance word list, and alphabetized list of all words, reading each word from right to left, grouping words by endings.

Type of computer: IBM 1401. *Size of storage:* 12k. *Language:* AUTOCODER. *Disks or tapes:* 2 IBM 1311 disks.

L199. Wortbildung der deutschen Zeitungssprache des 19. and 20. Jahrhunderts

Principal investigator: Ludwig Erich Schmitt, Ordentlicher Professor, Forschungsinstitut für Deutsche Sprache, U. Marburg, 355 Marburg/Lahn, Krummbogen 28A, Germany. *Associate:* Bernd S. Müller.

Scope: Frankfurter Tageszeitungen (1800-1960): c.170.000 Wörter. *Method:* 1) Ubertragung auf 8-Kanal-Lochstreifen. 2) Morphologische und semantische Klassifizierung. 3) Sortierungennach den Textwörtern, den morphologischen und semantischen Codes. 4) Ausdruck von zeitlich gestuften Teilsortierungen.

Type of computer: Telefunken TR4. *Size of storage:* 32k. *Language:* Maschinencode TEXAS. *Disks or tapes:* 4 Magnetbandeinheiten. Program is available. [*CHum,* May 1968]

L200. Syntax Discovery Procedure for Natural Languages [Deleted]

Status: Discontinued. [*CHum,* May 1968]

L201. Accusative Object as Verbal Complement in German-search Project

Principal investigator: Edwin A. Hopkins, Professor, Fremdspracheninstitut, Ruhr-Universität Bochum, 463 Bochum,Postfach 2148, West Germany. *Associate:* Kenneth Ekman.

Method: 1) Search tape for accusative object-verb combinations. 2) Print combination with 4 lines of context. 3) Tabulate results.

Type of computer: IBM 7094 model I. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 7 IBM model VI tape drive; 6 model V (we use only 3 at a time). *Special equipment:* CRT display, recorder, adapter available.

Status: Completed.

L202. Glossary of the Late Middle English *Chronicle* by John Capgrave

Principal investigator: Peter J. Lucas, Lecturer, Dept. of English, University College, Belfield, Dublin 4, Ireland. *Associate:* Michael Marr.

Scope: Complete *index verborum* to Capgrave's *Chronicle* (85,000 words) edited from MS. Gg.4.12 in Cambridge University Library. The two main purposes are: 1) to compile a select glossary suitable for publication in a new edition of the text, 2) to arrange all the data for a study of the vocabulary.

Type of computer: IBM 360/50. *Size of storage:* 256k. *Language:* PL/I. *Disks or tapes:* 1 9-track tape.

References: "Consistency and Correctness in the Orthographic Usage of John Capgrave's *Chronicle*."

L203. Application of Computing Technique to a Text in the Welsh Language: the Poetic Debate between Edmond Prys and Wiliam Cynwal (sixteenth century)

Principal investigator: G. Aled Williams, Assistant Lecturer, Dept. of Welsh, University College, Earlsfort Terrace, Dublin 2, Ireland.

Scope: Word index to text (54 poems) edited from Llanstephan MS. 43 in National Library of Wales, Aberystwyth. *Method:* 1) Prepare text by distinguishing homographs and partially de-mutating. 2) Keypunch text onto cards, two lines per card. 3) Sort alphabetically. 4) Print out in tabular form. 5) Post-edit.

Type of computer: IBM 1620/1. *Size of storage:* 40k. *Language:* SPS.

References: Article to appear in future issue of *Llen Cymru* (Cardiff: University of Wales Press). [*CHum*, May 1968]

L204. Concordance to *The Collected Poems of Dylan Thomas*

Principal investigator: Robert C. Williams, Brynarlais, Llanstephan, Carmarthenshire, Wales. *Associate:* W. B. Guindon, S.J.

Objective: To list alphabetically the use and occurrence of principal words used in Dylan Thomas' *Collected Poems* and prepare text for publication. *Method:* 1) Modified IBM KWIC sorting of card-punched lines of poetry to produce a card for each word in line with linear context and coding for identification. 2) Hand sorting through sorter to list alphabetically. 3) Manual correction and editing for printout.

Type of computer: IBM 1620. *Language:* FORTRAN. *Special equipment:*2000-card-per-minute manual sorter.

Status: Completed and published (University of Nebraska Press, 1967). [*CHum*, May 1968]

L205. Projet de Traduction Automatique

Principal investigators: Richard Kittredge, Director; Jules Dansereau, Gilles Stewart, Group de recherches pour la traduction automatique (T.A.U.M.), U. de Montréal, C.P. 6128, Montréal 101, PQ, Canada. *Associates:* Alain Ambrose, Réal Ayotte, Guy Poulin.

Scope: Machine translation from English to French with supporting linguistic research.

*Method:*Transformational analysis, transfer, synthesis implemented using Colmerauer's Q-systems.

Type of computer: CDC Cyber 74. *Size of storage:* 65k (60-bit words). *Language:* COMPASS. Program is available subject to agreement with National Research Council of Canada.

References: Mechanical Translation Project Quarterly Report, Montréal National Research Council and University of Montreal, Faculty of Letters, Applied Linguistics Section (Jan. 31, 1966, and on); A.Dugas, M. Gopnik, B. Harris, and J.P. Paillet, "Le Projet de Traduction Automatique à l'Université de Montréal," International Conference on Computational Linguistics, Preprint No.55 (Stockholm: KVAL, 1969); A. Colmerauer, G. De Chastellier, "W-Grammar," paper presented at the 1969 ACM National Conference and Exposition, San Francisco, 1969; A. Colmerauer, J. Dansereau, B. Harris, R. Kittredge, M. Van Caneghem, *TAUM 71,* Groupe de recherches pour la traduction automatique, Université de Montréal, 1971; M. Van Caneghem, *Morphographie générative du Français,* Département d'Informatique/Groupe de recherches pour la traduction automatique, Université de Montréal, 1971; A. Colmerauer, J. Dansereau, B. Harris, R. Kittredge, M. Van Caneghem, "An English-French MT Prototype," *Proceedings of International Conference on Computational Linguistics,* Debrecen, 1971; I. Bellert, C.Connors, T.R.Hofmann, *Working Papers in the Linguistics of Machine Translation,* Groupe de recherches pour la traduction automatique, Université de Montréal, 1971; R. Kittredge, "Use of Q-Systems for Deriving Normalized Sentence Trees," paper presented at NYU Symposium on

Natural Language Processing, New York, 1971; J. Dansereau, *Modèle génératif de la syntaxe du Français,* Département de linguistique et philologie/Groupe de recherches pour la traduction automatique, Université de Montreal, 1972.

L206. TRIAL Information Processing System

Principal investigator: Donald G. Dillaman, Research Assistant, Vogelback Computing Center, Northwestern U., 2129 Sheridan Rd., Evanston, IL 60201. *Associate:* Lorraine Borman.

Scope: TRIAL is a user-oriented general purpose text-processing system. A master file is created which can 1) be interrogated by commands using natural language (English), 2) retrieve and print records which will satisfy the search query, 3) generate indexes using either the basic KWIC or KWOC output formats, and 4) produce frequency counts of word lists either across all records on the file, or selectively, within specified levels of information. Any combination of these features can be achieved, through one computer run. *Method:* The existing system is fully operational in a batch processing environment. Design specifications are being developed to include remote terminal capabilities, "browsing" through the file, and possible experimentation on the use of thesauri and concept dictionaries.

Type of computer: CDC 6400. *Size of storage:* 65k. *Language:* CDC FORTRAN. *Disks or tapes:* 1 6603 disk file, 2 tapes. Program is available upon written request by the director of the local computing center.

References: Lorraine Borman and Donald G. Dillaman, *TRIAL Information Processing System* (Evanston: Northwestern University,1968).

L207. A Concordance to the English Prose of John Milton

Principal investigator: Laurence Sterne, 544 West 113 St., New York, NY 10025.

Type of computer: 1) IBM 360/75, 2) IBM 360/91. *Language:* FORTRAN IV G. *Disks or tapes:* 2 7- or 9-track tapes. *Special equipment:* Upper/lower-case print chain, bulk storage needed for sorting. Program is available. Can be used on IBM 360/50 or larger.

Status: Publication expected in Spring 1973 by Southern Illinois University Press. Compilation of concordance to Volume 2 begun summer of 1971 under the editorship of J. de Bruyn, U. of British Columbia; Volume 3 is under the editorship of Charles Larson, U. of Missouri, St. Louis; Volume 5 is under the editorship of Gladys Hudson, Baylor University.

Status: Continuing.

References: "A Concordance to the English Prose of John Milton,"*Dissertation Abstracts* 29, Sec. 2 (1968), 579-A.

L208. English Dialectology

Principal investigator: W. Nelson Francis, Professor, Dept. of Linguistics, Brown U., Providence, RI 02912. *Associate:* Gerald M. Rubin.

Type of computer: IBM 360/67. *Size of storage:* 512k. *Language:* PL/I, FORTRAN. *Disks or tapes:* 2 2400 tapes; 1 2314 disk. *Special equipment:* CalComp 563 digital plotter.

Status: The initial plan is to put the whole of *The Survey of English Dialects* (ed. H. Orton, et al., Leeds) on tape. Further investigations will use this corpus to check current theories of structural dialectology and to attempt to resolve some problems in English dialectology. *Method:* Devise coding system for IPA transcriptions; keypunch the corpus and transfer it to tape; isolate specific problems and study them by means of computer sorting of the data and ad hoc maps prepared on the CalComp plotter.

References: "Modal *Daren't* and *Durstn't* in Dialectal English," forthcoming in *Leeds Studies in English*; W.N. Francis, J. Svartvik, G.M. Rubin, "Computer-Produced Representation of Dialectal Variation: Initial Fricatives in Southern British English," COLING Preprint No. 52 (Stockholm, 1969); Gerald M. Rubin, "Computer-Produced Mapping of Dialectal Variation," *Computers and the Humanities* 4, 4 (March 1970), 241-246. [*CHum,* May 1970]

L209. General Stylometry of Contemporary English

Principal investigator: Preston Davis (deceased), Medfield Foundation, Harding, MA 02042.

Scope: The purpose of the investigation was to develop a set of machine-measurable variates which would yield practical description and classification of twentieth-century English prose styles. Numerous features, at letter, word, and sentence levels, of samples from the Brown University corpus and elsewhere were analyzed and reduced, by factorial and other methods, to a limited set of variables.

Type of computer: IBM 360/50. *Size of storage:* 512k. *Language:* Assembler, PL/I. *Disks or tapes:* 1 2314 disk; 2 2400 8-channel tapes. [*CHum,* May 1969]

L210. Text Concordance and Thematic Analysis of *Der grüne Heinrich* by Gottfried Keller

Principal investigator: Edda Schrader Gentry, Associate Professor, Dept. of Germanic and Slavic Languages and Literature, State U. of New York, Albany, NY 12203.

Method: The text of *Der grüne Heinrich* was keypunched on cards and transferred to magnetic tapes. Word frequencies, concordances, a word index, and a reverse index were compiled. This provides basic material for stylistic analysis of the novel as well as nineteenth-century German literary style.

Type of computer: CDC 3100.*Size of storage:* 32k/24k. *Language:* FORTRAN 3200. *Disks or tapes:* 2 854 disks; 5 604 tapes.Program available from Computing Center, SUNY, Albany, NY 12203.

Status: Completed 1971.

References: Wortindex zu Gottfried Keller "Der grüne Heinrich" Erste Fassung. Edited by Edda Schrader, vol. 3, 1-2 *Deutsche Wortindices* (Berlin and New York, 1971).

L211. Dissertations in English Literature Accepted by German, British, and American Universities for the Years 1964-1967

Principal investigator: Lawrence F. McNamee, Professor, Dept. of English, East Texas State U., Commerce, TX 75428. *Associates:* Ed Harris, David Ayres, Lowell Ballew, John Anderson, John McNamee.

Scope: To arrange these dissertations in a bibliography so that the doctoral student can tell in a matter of minutes whether or not his potential dissertation topic has previously been worked on. This is a follow-up to a larger volume completed in 1969 that spanned the years 1865-1963. *Method:* 1) Each dissertation was listed on index cards, coded, and keypunched. 2) Individual listings were sent to the respective universities for emendation and approval. 3) Cards were then resorted by topic (e.g., Shakespeare, Milton, American Colonial Literature), listed on 1402, and book printed from the tape.

Type of computer: IBM 360/40. *Size of storage:* 256k. *Language:* A & S COBOL. *Disks or tapes:* 6 2319 disks; 2 2415 tapes. Program is available.

Status: The first book was 1124 pages long and contained 14,521 dissertations. The supplement came out in 1969, was 450 pages long, and contained 4382 dissertations. It also contained the work of Canada and Australia. The years covered in Supplement One were 1964-1968 inclusive. The current project, Supplement Two, will include dissertations from the years 1969-1973 inclusive. Number of pages in each dissertation will be included.

References: Dissertations in English and American Literature (New York, 1968).

L212. A Computer-Assisted Graphemic Analysis of *Beowulf*

Principal investigator: Edward A. Kline, Assistant Professor, Dept. of English, U. of Notre Dame, Notre Dame, IN 46556. *Scope and method:* To describe the graphemic system underlying the text of the Beowulf by using standard graphemic analysis techniques.

Type of computer: IBM 1620. *Language:* SPS. Program is available.

References: Computers and the Humanities 2, 5 (1968), 211-213. [*CHum,* May 1969]

L213. An Index to the *Stationers' Register*

Principal investigator: William P. Williams, Associate Professor, Dept. of English, Northern Illinois U., Dekalb, IL 60115.

Objective: An index of all persons and titles found in the Eyre transcription of the *Register. Method:* 1) Pre-edit the text of the *Register.* 2) Put this data in machine-readable form using online terminals and WYLBUR text editing program. 3) Edit output using online terminals and WYLBUR. 4) After proofreading, to produce finished printer's copy from the highspeed printer using the WYLBUR data sets.

Type of computer: IBM 360/70. *Language:* WYLBUR, PL/I. *Disks or tapes:* 3 2311 disks; 4 2401 V tapes.

L214. Concordance to the Poems of Jonathan Swift

Principal investigator: Michael Shinagel, Professor, Dept. of English, Union College, Schenectady, NY 12308.

Objective: To compile a computer concordance to the canon of Jonathan Swift's poetry based on the "definitive" three-volume edition of *The Poems of Jonathan Swift,* 2nd edition, ed. Harold Williams (Oxford: Clarendon Press, 1958). This concordance is part of the Cornell Concordance Series under the general editorship of Professor Stephen M. Parrish and follows the program and format of this series.

Status: Published December 1972 by Cornell University Press.

L215. Frequency Study of Large English Corpus [Deleted]

Status: Suspended. [*CHum,* November 1968]

L216. Study of a Large English Corpus for a Semantic Parameter [Deleted]

Status: Suspended. [*CHum,* May 1969]

L217. Design of a Grammar Tester and Synthesizer

Principal investigator: Antonio A.M. Querido, Assistant Professor, Linguistique et Langues Modernes—Section de Linguistique Appliquée, U. de Montréal, B.P. 6128, Montréal 26, PQ, Canada. *Associate:* Gilles Gauthier. *Scope and method:* To obtain a system of programs useful for linguistic research by computer. The first research carried on with this system will be the development of a basic grammar for French. The linguistic model will be generative-transformational grammar according to the latest developments. The system of programs should accept grammars written in the usual format and should permit changes in the rules of each component of the grammar. It should also permit tracing of the generative processes and give different kinds of outputs showing the "state of affairs" at some points in the generative process.

Type of computer: IBM 360/30. *Language:* LISP and PL/I. *Special equipment:* Cathode-ray tube display and on-line access (mainly for the tracing of generative processes). [*CHum,* May 1969]

L218. The Use of the Computer in Reconstructing Proto Ekoid Bantu

Principal investigator: David W. Crabb, Associate Professor, Dept. of Anthropology, Princeton U., 100 Green Annex, Princeton, NJ 08540. *Associate:* Andrea B. Smith. *Scope and method:* To reconstruct the sounds and word structure of an ancestral language by means of operations made by the computer on the descendant languages. We collected the equivalents of 530 English glosses in contemporary Ekoid Bantu languages and coded, keypunched, and stored this as our basic data. We then developed programs that break words into their structural components and isolate and categorize roots according to initial vowels, tone marks, etc. Comparisons have been made between theoretical (expected) and actual patterns of variation between the 14 languages.

Type of computer: IBM 360/67. *Size of storage:* 300,000 bytes. *Language:* PL/I F. *Disks or tapes:* 2 disks; 2+ tapes (varies). This was not a general purpose program, but could be adapted for some other uses. [*CHum,* May 1969]

L219. A Lexicon to Diodorus Siculus

Principal investigator: J. Iain McDougall, Lecturer, Dept. of Classics, U. of Winnipeg, 515 Portage Avenue, Winnipeg 2, MB, Canada.

Scope: The work is to combine the advantages of lexicon and index with the alphabetization of Diodorus being accomplished by computer and the lexicographical section being completed manually.

Type of computer: IBM 360/65. *Size of storage:* 512k. *Language:* Assembler F (OS).

Status: Computer part concluded. Manual analysis now in progress.

L220. Concordance du Théâtre et des Poésies de Molière

Principal investigator: Bryant C. Freeman, Professor and Chairman, Dept. of French and Italian, U. of Kansas, Lawrence, KS 66044. *Associate:* Lynne Huxsaw.

Scope: Line concordance to the complete theater and poetry, with word indices only for common words. Frequency list in descending order in telescoped forms (e.g., *chien* and *chiens* will appear under the singular). Analytical tables for each role in each play. Special appendix of ultracommon words (*à, ce, de, et, le, un*).

Type of computer: Burroughs B5500. *Size of storage:* 244k. *Language:* Burroughs Extended ALGOL (for concording); COBOL (for alphabetizing).

Status: Published by Cornell University Press. [*CHum,* May 1969]

L221. Syntactic Analysis of English

Principal investigator: Adélard Faubert, Research Assistant, Dept. of Applied Linguistics, U. of Montreal, P.O. Box 6128, Montreal, PQ, Canada. *Associates:* Juliette Daigneault, Serge Vallée. *Scope and method:* Syntactic analysis of a 4000-word technical text. Binary + N-ary constituent groupings mechanically analyzed and synthesized by an application of Cocke's and Sakaï's (Kuno-Woode's) algorithms.

Type of computer: CDC 3400. *Size of storage:* 32,768. *Disks or tapes:* 10 tapes. [*CHum,* May 1969]

L222. Reconnaissance automatique des structures du français écrit

Principal investigator: André Dugas, Assistant de Recherches, Linguistique appliquée, U. de Montréal, B.P. 6128, Montréal 26, PQ, Canada.

Method: Instructions programmées à partir de l'algorithme de Cocke.

Type of computer: CDC 3400. *Size of storage:* 32,000 mots. *Language:* FORTRAN IV, ALGOL, COBOL, SLIP. *Disks or tapes:* 12 604 (60 KC) tapes.

References: Etudes de linguistique appliquée sur les structures syntaxiques du français oral de Montréal (Canada), C.E.T.A. (Grenoble, 1966). [*CHum,* May 1969]

L223. Quantitative Studies in Prosody: Blank Verse from Shakespeare to Frost

Principal investigator: John D. Allen, Emeritus Professor (deceased), Dept. of English, East Tennessee State U., Johnson City, TN 37601.

Scope: To determine, in the verse of ten major English and American poets who have used blank verse extensively, the frequencies of occurrence of nine features of prosody varyingly important in creating verbal music. *Method:* Sampling, counting, and preparing for certain features suitable work sheets for keypunch operation, and compiling results for some thirty tables. The computer was chiefly useful in connection with data for alliteration, assonance, and consonance.

Type of computer: IBM 1401. *Size of storage:* 8k. *Language:* AUTOCODER. *Disks or tapes:* 2 1311 disks; 2 7330 tapes.

Status: Volume I of the book is ready for printing. Copy for Volume II is in the hands of the University of East Tennessee Press. [*CHum,* November 1968]

L224. Numerical Classification of Indo-European Languages

Principal investigator: Edward R. Gammon, Associate Professor, Dept. of Linguistics, Fresno State College, Fresno, CA 93726.

Scope: Analysis of Kroeber-Chrétien measure of similarity. *Method:* Apply Bell Telephone programs for 1) clustering and 2) proximities.

Language: FORTRAN IV. Program is available.

References: "Quantitative Linguistic Typologies," *Proceedings of the 10th International Congress of Linguists.* [*CHum,* November 1968]

L225. Herstellung von Sprachkarten

Principal investigator: Ludwig Erich Schmitt, Ordentlicher Professor, Forschungsinstitut für deutsche Sprache, Deutscher Sprachatlas, U. Marburg, 355 Marburg, Kaffweg 3, Deutschland. *Associate:* W. Putschke.

Scope: Anwendbar für sprachliche Elementkarten (Laut-, Wort- und Formenkarten) mit direkt oder indirekt aufgenommenen Material. *Method:* Ausdrucken einer Belegliste; Herstellen einer Häufigkeitsliste der einzelnen Beleg; Kartenausdruck in Einzelzeichenform; Herstellen der Legende und einer Liste der Ausnahmebelege.

Type of computer: Telefunken TR4. *Language:* Maschinencode TEXAS. Program is available.

References: "Uber ein Computerprogram zur Herstellung von Sprachkarten," *Zeitschrift für Mundartforschung, Beiheft.* [*CHum,* November 1968]

L227. Automatic Phonemtransformation Sanskrit-Indonesian

Principal investigator: Harry Spitzbardt, Professor of Theoretical and Applied Linguistics, Sektion Sprachwissenschaft, Friedrich-Schiller U., 69 Jena, West Germany.

Scope: 1) Establishing an inventory of phonemic transformation rules Sanskrit-Indonesian. 2) Morphological typology by statistical methods. *Method:* Word-matching; computation of group (syllable) and class (phoneme) intervals.

Type of computer: CELLATRON C 8206. *Size of storage:* 4096 words at 33 bits each (drum). *Language:* Machine code.

References: "Mathematische Methoden in der Modernen Linguistik," *Wissenschaftliche Zeitschrift der Friedrich-Schiller-Universität Jena, Gesellschafts- und Sprachwissenschaftliche Reihe* 3 (1965), 499ff.; "Zur Entwicklung der Sprachstatistik in der Sowietunion," ibid., 4 (1967), 471ff.; "Einige Bemerkungen zu G. Altmanns statistischer Analyse indonesischer Morphemstrukturen," *Mitteilungen des Instituts für Orientforschung* 1 (1967), 137ff.; "Kritischer Bemerkungen zur Glottochronologie," *Ethnographisch-Archäologische Zeitschrift* 1 (1969), 41ff.; "Automatic Phoneme Transformation: Sanskrit-Indonesian," *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 4 (1972).

L228. Indices zu Abkurzungen aus Knossos-Texten

Principal investigator: Heinz Geiss, Institut fur griechisch-romische Altertumskunde, 108 Berlin, Otto-Nuschke-Str. 23, West Germany. *Associate:* Erich Mater.

Scope: About 3000 abbreviations. *Method:* 1) Transliterate originals. 2) Keypunch. 3) Sort by abbreviations and positions. 4) Count. 5) Print in tabular form.

Type of computer: BULL, General Electric, Model B.S.; Soemtron, Model 401.

References: "Untersuchungen zur Ventrisschen Entzifferung," *Klio* 48 (1967). [*CHum,* November 1968]

L229. Description syntaxique du français et vérification à l'aide d'un ordinateur

Principal investigator: André Dugas, Professeur, Dept. of Linguistics, U. du Québec, B.P. 8888, Montréal, PQ, Canada. *Associates:* Thomas R. Hofmann, Serge Boisvert, Denise Bélanger, Lorne Bouchard, Jean-Pierre Paillet.

Scope: Construction d'une grammaire française à la lumière des théories transformationnelles. *Method:* Theoretical research followed by computation of the data and grammar testing.

Type of computer: CDC 6400. *Size of storage:* 65k. *Language:* FORTRAN IV 60. *Disks or tapes:* 1 6603 (6638) disk; 8 604 tapes.

References: Grammaire française, Université de Québec à Montréal; L.Drapeau et J. Gérard (à paraître), "Là où les implications se compliquent," *Cahiers de linguistique* 3, Montréal, Université du Québec; S. Bouvert et A. Dugas, "OBLING, a Tester for Transformational Grammars," ICCL, Pise, 1973; "Les compléments d'instrument," *Problèmes de sémantique,* ed. A.Dugas et J.McA'Nulty (Montréal: Presses de l'Université du Québec, 1973).

L230. A Word Frequency Compilation of the German Middle Class Tragedy

Principal investigator: Robert B. Youngblood, Assistant Professor, Dept. of German, Washington and Lee U., Lexington, VA 24450. *Associate:* John H. Wise.

Method: 1) Keypunch texts. 2) Proofread printout with text-critical edition of text. 3) Print word counts in order of occurrence, alphabetically, and in order of frequency.

Type of computer: IBM 1130. *Size of storage:* 8k. *Language:* 1130 FORTRAN. *Disks or tapes:* 2 2315 disks. [*CHum,* November 1968]

L232. Concordances to Swedish Poets

Principal investigator: Jan Thavenius, Assistant Professor, Dept. of Literature, U. of Lund, 223 62 Lund, Sweden.

Scope: 1) Development of a concordance program. 2) Production of concordances, in the first place to the poets Esaias Tegnér (1782-1846) and Hjalmar Gullberg (1898-1961). 3) Research work.

Type of computer: SAAB D21. *Language:* Dac, ALGOL-GENIUS. Program is available.

Status: Almost completed.

References: Ordindex till Hjalmar Gullbergs lyrik I-II (Lund 1969-1971) (Word Index to the Poems of Hjalmar Gullberg); *Konkordans till Hjalmar Gullbergs lyrik* (Stockholm, 1971) (A Concordance to the Poems of Hjalmar Gullberg); *Ordindex till Esaias Tegnérs lyrik I-II* (Lund, 1970-1973) (Word Index to the Poems of Esaias Tegnér); *Konkordans till Esaias Tegnérs lyrik* (in prep.) (A Concordance to the Poems of Esaias Tegnér); *Stil och vokabulär* (Lund, 1972) (Style and Vocabulary).

L233. Cyril Tourneur: A Problem of Authorship

Principal investigator: Michael G. Farringdon, Lecturer, Dept. of Computer Science, University College of Swansea, Singleton Park, Swansea, Glam., U.K.

Scope: Study of works attributed to Tourneur with a view of finding if *Revenger's Tragedy* and *Atheist's Tragedy* are by same hand.

Type of computer: ICL 1904S and 1906A. *Language:* POP-2 and ALGOL 68.

L234. Spoglio lessicale dei testi etei cuneiformi

Principal investigator: P. Meriggi, Professor, U. di Pavia.

Scope: Si vuol compilare un dizionario della lingua etea, la lingua indeuropea più anticamente attestata, che conserva nella grammatica preziosi relitti arcaici. Il calcolatore del CNUCE di Pisa fornisce la documentazione lessicografica di base, con le concordanze e le schede contesto. Sono pronte le schede dei testi storici (circa un sesto del tutto). E in corso la preparazione d'un'altra porzione di testi per la perforazione.

L235. Lingue Orientali di Napoli

Principal investigator: N. Radovich, Professor.

Scope and method: Le Biblioteche italiane custodiscono un notevole numero di manoscritti slavi medioevali tuttora inediti, datati e localizati solo approssimativamente. Si é constatato che la codificazione dei manoscritti presenta caratteristiche (inventario, frequenza e possibilità di combinazioni dei segni grafici) che variano a seconda della età e della provenienza dei manoscritti stessi. Sembra pertanto che una statistica delle caratteristiche sopra elencate, effettuata su testi datati e localizzati, possa fornire elementi esprimibili numericamente e facilmente comparabili, validi come punto di riferimento per la datazione e la localizzazione di altri testi di età e provenienza non note. [*CHum,* November 1968]

L236. Filologia Bizantina

Principal investigator: Giuseppe Schiro, Professor, Storico Filologico, Studi Bizantini e Neoellenici, Appia Nuova 96, Italy.

Scope and method: Si tratta di fare elaborazioni lessicali e metriche di testi bizantini. Siamo ai primi contatti. La maggiore difficoltà è rappresentata dalla mancanza di una catena greca. [*CHum,* May 1969]

L237. Opere minori di Dante

Principal investigator: Professor Bosco, direttore Enciclopedia Dantesca, Prof. di Letteratura italiana nell'U. di Roma, Facoltà di Magistero.

Scope and method: Si son prodotte concordanze delle opere minori di Dante, comprese le opere a lui attribuite, *Fiore* e *Detto d'amore,* da affiancare alle Concordanze Pisane della *Divina Commedia;* sono e saranno utilizzate dai redattori per la composizione degli articoli della Enciclopedia Dantesca che è in corso di pubblicazione a cura della Enciclopedia Treccani.

L238. Semantematics of Greek Choral Lyric (texts of Pindar and Bacchylides)

Principal investigator: C.O. Pavese, Professor of Greek, Facoltà di Lettere e Filosofia, U. di Venezia, Venice, Italy. *Associate:* Bruno Gentili.

Method: The texts of Pindar (Epinician odes) and of Bacchylides (all except fragments), recorded with symbols to indicate the themes and motifs of Greek choral lyric, permitted various collections of data in romanized transliteration for a book on the Greek choral lyric, which is hopefully going to appear during 1974.

Status: Completed 1973.

References: "Semantematica della poesia corale greca," *Belfagor* 23 (1968), 389-431.

L239. Concordantiae Senecanae

Principal investigators: R. Busa, S.J., A. Zampolli, auspicio auctoritate Societatis linguae Latinae historice investigandae ab Italico Consilio studiis provehendis (CNR) constitutae. Istituto di Filologia Latina, U. de Padova, "Liviano"—Piazza Capitaniato, 7, 35100 Padova, Italy.

Scope and method: Ci si propone di pubblicare concordanze non lemmatizzate e inoltre indice di frequenza e indice inverso di tutte le opere di Seneca, compresi i frammenti di sicura attributionel. Particolare riguardo è riservato alla elaborazione automatica delle varianti dell'apparato critico. Le *Concordantiae* sono in corso di stampa presso Olms, Hildesheim.

L240. Phonology and Morphology in *Reisen und Gefangenschaft Hans Ulrich Kraffts*

Principal investigator: Professor Dr.P. Hartmann, Sprachwissenschaftliches Seminar der Universität, Domplatz 23, 44 Münster, Germany. *Associates:* K. Brochhaus, Ilpo Piirainen.

Scope and method: Es werden bearbeitet: Laut- und Formenlehre in *Reisen und Gefangenschaft Hans Ulrich Kraffts* (Ulm, 1616). Von diesem Text liegen eine Stellenkonkordanz, je ein Namen-, Wort- und Häufigkeitsregister vor.

Type of computer: IBM 7094. *Language:* FORTRAN II. [*CHum,* May 1969]

L241. Types of German Nouns

Principal investigator: Dr. H. Vater, Seminar für vergleichende Sprachwissenschaft der U. Hamburg, von-Melle-Park 6, 2 Hamburg 13, Germany.

Scope and method: Seit Herbst 1965 wird ein Gebiet der deutschen Wortbildung bearbeitet: "Strukturtypen deutscher Substantive." Aus Texten verschiedener deutscher Autoren und Zeitungen ausgewählte Substantive werden mit Zusatzzeichen versehen, die etwas über ihre Gliederung in Stämme, Affixe usw. aussagen, in die Maschine gegeben und automatisch nach Strukturtypen klassifiziert. Das notwendige Programm wurde im DRZ Darmstadt erarbeitet.

Type of computer: IBM 7094. *Language:* FORTRAN II.

Status: Inactive. [*CHum,* May 1969]

L242. Auxiliary Materials, Experiments, and Analyses toward Computerized Translation

Principal investigators: S. Capobianchi, G. Lustig, S. Perschke, Cetis Euratom C.C.R., Ispra, Italy.

Scope and method: Es laufen 1) automatische Erstellung von Konkordanzen und mehrsprachigen Indices aus amtlichen Texten als terminologische Hilfe für manuelle Ubersetzungen in demselben Fachgebiet; 2) theoretische und experimentelle Untersuchungen über die Darstellungen semantischer Beziehungen zwischen Wörtern mittels Association Factors und deren Anwendung im Information Retrieval; 3) automatische Sprachanalyse und -synthese im Rahmen des Projekts "Automatische Sprachübersetzung aus dem Russischen ins Englische." Das Übersetzungssystem wurde von der Georgetown University Washington entwickelt und wird seit 1962 von Euratom praktisch angewendet und verbessert.

Type of computer: IBM 7094, IBM 1401. [*CHum,* May 1969]

L243. Statistical Analysis of the Words in About 25 Professional Articles

Principal investigator: Dipl.-Ing. S.W. Wagner, Institut für Nachrichtenverarbeitung und Nachrichtenübertragung der TH Karlsruhe, Kaiserstrasse 12, 75 Karlsruhe, Germany.

Scope and method: Seit 1964 laufen Arbeiten auf dem Gebiet der automatischen Inhaltsanalyse von Texten auf Grund rein statistischer Kriterien. Als Experimentiermaterial dienen rund 25 Texte von Fachaufsätzen. Von allen Texten wurden die Worthäufigkeitsverteilungen (im Sinne von Zipf) ermittelt. Betrachtet werden die den absoluten Häufigkeitswerten entsprechenden Rangnummern, ferner die Wortlängen. Gesetzmässigkeiten in den Worthäufigkeitsverteilungen werden auf Grund new eingeführter Textparameter untersucht.–Rückschlüsse auf den semantischen Wert der Wortarten werden auf Grund der Rangnummern gezogen. Dies geschieht durch Vergleich der Rangnummern identischer Wortarten in einer transformierten Häufigkeitsverteilung und in einer sog. Mischtextverteilung, wobei keine vorbereitete Fachwortliste (Thesaurus) benötigt wird. Als Zusatzkriterium wird eine Auswertung von Wortaffinitäten in Erwägung gezogen. Was die Programmiertheorie betrifft, so wurde zue Herbeiführung der Vergleichbarkeit kurzer Einzeltexte mit dem langen Mischtext eine Transformationsvorschrift aufgestellt. Sie liefert Aussagen darüber, wie sich die Verteilungsparameter und die Ränge der Häufigkeitsstufen mit zunehmender Textlänge ändern.

Type of computer: ER 56. *Language:* Symb. Adressierprogramm. [*CHum,* May 1969]

L244. The Relationship Between Deep Structure and Surface Structure in Various Indo-European Languages

Principal investigator: Dr. H. Seiler, Institut für Sprachwissenschaft der Universität Köln, 5 Köln, Germany.

Scope and Method: Es entsteht eine Untersuchung des Verhältnisses zwischen "Deep Structure und Surface Structure in indogermanischen Sprachen im Hinblick auf automatische syntaktische Analyse."

Type of computer: IBM 7094. *Language:* FORTRAN II. [*CHum,* May 1969]

L246. Programming and Codification of Linguistic Data in Various Categories

Principal investigators: Professor Dr. G. Ungeheuer, Dr. H. Schnelle, Dr. D. Krallmann, D. Bünting, Yün-Tin Wang, Institut für Phonetik und Kommunikationsforschung der Universität Bonn, Adenauerallee 98a, 53 Bonn, Germany.

Scope and method: 1) *Organisation sprachlicher Daten.* Es wird an einem System zur Kodierung und Ubertragung von Drucktexten auf Lochkarten und Magnetbänder gearbeitet; ferner an einem Programmiersystem für eine Rechenanlage IBM 7090 zur Verarbeitung sprachlicher Daten auf Lochkarten und Magnetbändern; die Anwendung der Programme soll weitgehend unabhängig sein von der Kodierung und der Menge des Materials. 2) *Statistik.* Es wird eine Systematik der statistischen Methoden zur stilistischen Textbeschreibung entwickelt. Als linguistische Einheiten treten Buchstaben, Silben, Wörter, Sätze usw. auf. An statistischen Massen wurden insbesondere Mittelwerte und die höheren Momente um den Mittelwert verwendet. Weiterhin laufen informationstheoretische Untersuchungen von Texten–Untersuchung der relativen Häufigkeiten von Buchstaben, Berechnung von Entropiewerten, Verbundwahrscheinlichkeiten von Symbolen und Symbolverkettungen–und Untersuchungen zur Systematik der Informationserschliessung, insbesondere innerhalb der stilistischen Textbeschreibung und der automatischen Dokumentation. 3) *Lexikologie.* Es liegen Wortverzeichnisse zu den Werken von Kant vor (Bände 1-9 der Akademie-Ausgabe), ferner zu mittelhochdeutschen Texten (in Zusammenarbeit mit dem Germanistischen Seminar der Universität Bonn). Es wird Programmierhilfe geleistet bei der Aufbereitung von Texten west- und ostdeutscher Zeitungen (in Zusammenarbeit mit dem Institut für deutsche Sprache, Aussenstelle Bonn). 4) *Morphologie.* Es liegen Arbeiten vor zur Systematik, Formalisierung und Programmierung der deutschen Wortbildung. 5) *Semantik.* Untersuchungen zur Testformation im Sinne von G. Kandler sind im Gange.–Ausserdem werden bisher Entwickelte Ubersetzungsverfahren kritisch untersucht.

Status: Completed December 1972.

L247. Machine-Readable Dictionary of Middle High German

Principal investigators: Professor Dr. W. Besch, Professor Dr. Hugo Moser, Germanistisches Seminar der Universität Bonn; Professor Dr. Gerold Ungeheuer, Institut für Kommunikationsforschung und Phonetik der Universität Bonn, 53 Bonn, Adenauerallee 98a, Germany. *Associates:* U. Krumnack, W. Lenders.

Objective: On the basis of machine-readable Middle High German texts an automatic dictionary will be established which shall contain morphological, syntactical, and semantic information in a formal manner. This dictionary may be used in order to generate lemmatized and grammatically classified concordances and indices as described in project L263. It may also be used in order to make accessible a Middle High German data bank.

Type of computer: IBM 370/165. *Language:* FORTRAN IV, PL/I, KOMFOR. *Disks or tapes:* About 45 tapes. Program is available.

References: See L263.

L248. Concordance and Grammar of Spoken German, with Attention to Some Dialects.

Principal investigator: Professor Dr. E. Zwirner, Deutsches Spracharchiv, Studtstrasse 39, 44 Münster, Germany.* *Associates:* H. Richter, W. Bethge, E. Knetschge, A. Ruoff.

Scope and method: Im Mittelpunkt der gegenwartigen Arbeiten steht die Herstellung digitaler Duplikate von Schwerpunktaufnahmen des Deutschen Spracharchivs (Umgangssprache, schlesische und sudwestdeutsche Mundarten). Dies wird im wesentlichen auf folgendem Weg erreicht: 1) Speicherung von phonetisch-hochsprachlichen Konkordanztexten (mit einer Zuordnung von Transkription und hochsprachlicher Entsprechung durch eine Adresskonstante) auf Lochstreifen; 2) Spezifizierung von Zeichenketten des hochsprachlichen Konkordanztextteils als Elemente bestimmter Morphemklassen einer Kodierungsgrammatik; 3) Speicherung digitalisierter physikalisch-technischer Bewertungen des Sprachsignals; 4) Speicherung einer auditiven Beurteilung der suprasegmentell-prosodischen Merkmale; 5) Zuordnung der unter 1-4 beschriebenen Daten und ihre Vereinigung auf Magnetband. Nach Aufbereitung der Konkordanztexte gemass der Kodierungsgrammatik sollen erste, Schritte zur Erstellung einer formalen Grammatik der gesprochenen Sprache unternommen werden. Die Kodierungsgrammatik wird wahrscheinlich eine Mischform von Konstituentenstruktur- und Kategorialgrammatik sein. Ferner ist vor allem eine Statistik von Signalparametern vorgesehen mit dem Ziel, akustische Korrelate sprachlicher Einheiten aufzufinden. [*CHum,* May 1969]

L249. Computer-Assisted Translation and Computerized Lexicography

Principal investigator: Friedrich Krollmann, Leitender Regierungsdirektor, Bundessprachenamt, D-503 Hürth, Bez. Köln, Horbeller Strasse, West Germany. *Associates:* Hans-Jochen Schuck, Erika Hoffmann.

Objective: Speeding up and streamlining of lexicographical work with EDP equipment, semiautomation of the translation process, updating of large data files, and fully automated typesetting of special dictionaries in various languages. *Method:* Data base consists of more than 500,000 entries in English, French, Russian, and German, originating predominantly from technical and scientific fields. The teleprocessing system provides the translator with a text-oriented glossary as an aid in the translation process; computer dialog envisaged for 1974.

Type of computer: IBM 370 equipment is used in time-sharing mode; TP operation.

References: U.Winkler, "Aufbau und Anwendungsmöglichkeiten elektronischer Wörterbücher," *Nichtnumerische Informationsverarbeitung,* Hrsg. R. Gunzenhäuser, (Wien: Springer Verlag, 1968); H.-J. Schuck, "Zusammenarbeit Mensch/Maschine beim Umgang mit elektronisch gespeicherten Wörterbüchern," *Nachrichten für Dokumentation* (Frankfurt/Main) 20, 1, (1969); F. Krollmann, "Linguistic Data Banks and the Technical Translator," *Meta, Journal des Traducteurs* (Montréal) 16, 1-2 (1971).

L250. Collection of Data on Current German, Preparatory to Many Applications

Principal investigator: Professor Dr. H. Moser, Institut für deutsche Sprache, Friedrichsplatz 12, 68 Mannheim,West Germany.

Scope and method: Bei dem 1965 in Angriff genommenen Projekt "Deutsche Textbibliothek" handelt es sich zunächst darum, einen repräsentativen Querschnitt der deutschen Gegenwartssprache mit Hilfe von Lochstreifen (Mannheim) und Lochkarten (Bonn) auf Magnetbänder zu übertragen. Dieses Datenmaterial soll nach den verschiedensten Gesichtspunkten maschinell bearbeitet werden. (Hinsichtlich einer ausführlichen Darstellung der angewendeten Methoden bei der Textaufnahme in den Beiden Arbeitsstellen, des Standes der Bearbeitung und weiterer Planungen, einer Beschreibung des Verfahrens die der Ermittlung einer Querschnitts [qualitative Auswahl] sowie der Darstellung des Verfahrens zur Ermittelung einer [quantitativ] optimalen Repräsentanz bei Zeitungstexten wird auf ein Heft der Mannheimer Schriftenreihe vierwiesen, das in Kürze erscheinen soll.)

Type of computer: IBM 7094, IBM 1401. *Language:* FORTRAN II. [*CHum,* May 1969]

L251. Lexical Statistics to Be Derived from the Texts of Certain Authors

Principal investigator: Professor Dr. W. Fucks, Physikalisches Institut der TH, Abteilung Sprach- und Musikstatistik, Templergraben 55, 51 Aachen, Germany . *Associates:* J. Lauter, W. Reckziegel, D. Wickmann, R.Mix, N. Malmendier.

*Scope and method:*Texte bestimmter Autoren werden untersucht im Hinblick auf Silben, Satzlängen, Wortklassen und Worklassenverbindungen. Sämtliche statistischen Charakteristika werden ermittelt, insbesondere Korrelationen, Markoffketten, Ubergangsmatrizen und informationstheoretische Hilfsmittel. Es existieren Programme zur Berechnung

der statistischen Daten, der Ubergangsmatrizen mit verschiedenen Normierungen usw. für die IBM 7090. [*CHum,* May 1969]

L252. Translation from Hebrew to English

Principal investigator: Dr. P.O. Samuelsdorff, Institut für Sprachwissenschaft der Universität, 5 Köln, West Germany.

Scope and method: Es wird an einem hebräisch-englischen Ubersetzungsprogram gearbeitet. Die Sprachanalyse basiert auf der Konstituentenstruktur. Es wird keine Interlingua erstellt; der zu übersetzende Text wird Wort für Wort in einem Wörterbuch aufgesucht, in dem neben jedem hebräischen Wort syntaktische und semantische Angaben verschlüsselt sind, sowie ein englisches Wort oder eine Routine, ein englisches Wort unter mehreren auszuwählen.

Type of computer: IBM 7094. *Language:* FORTRAN II. [*CHum,* May 1969]

L253. Mathematical Analysis of Egyptian Papyri

Principal investigator: F. Schulte-Tigges, Abteilung Nichtnumerik, Deutsches Rechenzentrum Darmstadt, Rheinstrasse 75, 61 Darmstadt, Germany.

Scope and method: Dem Charakter des Instituts entsprechend, wurden bzw. werden fast alle anfallenden Aufgaben aus dem Bereich der Linguistik, zusammen mit Mitgliedern anderer Institute, bearbeitet. Der Arbeitsanteil des DRZ variiert von einfachen Programmierhilfen (in wenigen Fällen) bis zur Entwicklung von Lösungsverfahren und Programmen (in den meisten Fällen). Als eigenes Forschungsprojekt werden zur Zeit ägyptische Texte bearbeitet. Die Lösung der behandelten Probleme wurde auf der programmiertechnischen Seite durch ein spezielles Programmsystem ermöglicht, welches in den vergangenen Jahren erarbeitet wurde. Es besteht aus einem umfangreichen Satz von Standard-Unterprogrammen im symbolischen Maschinencode FAP und einer Reihe von organisatorischen Rahmenprogrammen in der algorithmischen Sprache FORTRAN (Version II). Je nach Aufgabenstellung werden diese Programme nach dem Baukastenprinzip zusammengestellt. Für eine Reihe von Aufgabenbereichen (Sortierung nach wählbaren Alphabeten, Lexikographie, Sprachstatistik) liegen Programmsysteme (integrierte Haupt- und Unterprogramme, "chain-jobs") vor. Ein Umschreiben dieser Programme in andere Programmiersprachen (etwa MAP und FORTRAN IV) ist ggf. vorgesehen.

Type of computer: IBM 7094. [*CHum,* May 1969]

L254. Lexical Punch Card Treasure of the Mother Tongue and Two Foreign Languages

Principal investigator: Dr. E. Agricola, Deutsche Akademie der Wissenschaften, Arbeitsstelle für mathematische und angewandte Linguistik und automatische Ubersetzung, Mohrenstrasse 39, 108 Berlin, East Germany.

Scope and method: Es existiert ein Lochkartenarchiv der deutschen Gegenwartssprache (Wort-Grundformen mit einigen grammatischen Informationen) zur automatischen Herstellung von speziellen Wortlisten. Ersters grösseres Produkt: E. Mater, *Rückläufiges Worterbuch der deutschen Gegenwartssprache* (Leipzig, 1965). Fortlaufende englische Fachtexte (Naturwissenschaft, Technik) werden zur Feststellung von Umfang, Inhalt und Frequenzen der Wortschätze bestimmter Fachgebiete und einiger syntaktischer Merkmale, einschliessich Kompositionstypen, auf Lochkarten übertragen. Es sind ferner vorhanden Wörterbücher der englischen und der russischen Wortformen des allgemeinen wissenschaftlichen und einiger Fachwortschätze als Grundlage der Analysewörterbücher für die automatische Ubersetzung, mit Informationen über Flexion, syntaktisches Verhalten und Wahl unter den Ubersetzungsäquivalenten. Ein Wörterbuch der deutschen Wortstämme mit Informationen über Flexion, Rektion usw. als Synthesewörterbuch zu den vorgenannten Quellensprachenwörterbüchern. Auf dem Gebiet der Morphologie wird u.a. an der automatischen Trennung von Worststamm und Endung und Identifizierung der Flexionsformen bei der Analyse des Russischen, auf dem Gebiet der Syntax und der kompletten syntaktischen Analyse englischer und russischer Textsätze, Zuordnung von Strukturbäumen (mit Unterordnungscharakteristiken), Konstruktionsumwandlung, Analyse von Idiomen und Komposita und nachfolgender Synthese der funktionell äquivalenten deutschen Strukturen gearbeitet. Ferner wird ein semantischer Thesaurus und ein Operationssystem zur Lösung semantischer und syntaktischer Mehrdeutigkeiten in Textsätzen aufgestellt. Für die automatische Sprachübersetzung werden alle Stufen des bilingualen Ubersetzungsvorgangs Englisch-Deutsch and Russisch-Deutsch erforscht und algorithmisch dargestellt und in experimentellem Umfang auf Rechenanlagen übertragen. [*CHum,* May 1969]

L255. Computer-Assisted Study of Herman Broch's *Tod des Vergil*

Principal investigator: Gudrun H. Wasson, Assistant Professor, Dept. of German, U. of California, 405 Hilgard Ave., Los Angeles, CA 90024. *Associate:* John T. Wasson.

Scope: Initially, study of imagery in this novel. Later, study of various stylistic features. *Method:* Preparation of a selective computer concordance of the novel, with location-identification and extensive context.

Type of computer: IBM 360/75. *Size of storage:* 500k. *Language:* FORTRAN IV. [*CHum,* May 1969]

L256. Lexicographical and Stylistic Analysis of Gower's *Confessio Amantis*

Principal investigator: D.S. Brewer, Lecturer, Dept. of English, U. of Cambridge, Emmanuel College, Cambridge, England. *Associates:* R.A. Wisbey, M.F. Bott, J.G.B. Heal, M.F. Porter.

Scope: Complete concordance with reverse and rhyme indexes and ranking list of the poem, ultimately for publication.

Type of computer: 1) ICT Atlas II and 2) IBM 370/165. *Size of storage:* 1) 128k and 2) 1024k bytes. *Language:* 1) Machine language, and 2) BCPL. *Disks or tapes:* 4 Data Products diskfiles, 6 Potter tape drives. *Special equipment:* Multiprogramming system, multi-access.

Status: Complete except for publication. Early microfilm version sent to Centre for Medieval and Renaissance Studies, 320 Main Library, 1858 Neil Avenue, Columbus, OH 43210.

L257. Lexicographical and Stylistic Analysis of an Early Middle High German Text

Principal investigator: David A. Wells, Lecturer, Dept. of German, Bedford College, Regent's Park, London, NW1 4NS, England.

Scope: Concordance, reverse index, frequency counts, and rhyme index for the Early Middle High German *Vorauer Bücher Moses.*

Status: Project almost complete. Accepted for publication by Cambridge University Press.

L258. Concordance of the Major Works of Joost van den Vondel

Principal investigator: Peter K. King, Lecturer in Dutch, Faculty of Modern and Medieval Languages, Cambridge U., Sidgwick Ave., Cambridge, England. *Associates:* R.A. Wisbey, M. Porter.

Scope: Concordance of all original plays with their forewords, dedications, etc.; also *Koning Edipus, Sofompaneas, Aenleidinge ter Nederduitsche Dichtkunste, Altaergeheimenissen, Bespiegelingen van Godt en Godtsdienst, Heerlyckheit der Kercke, Joannes de Boetgezant* and some shorter poems.

L259. A Concordance to the Early Middle High German Poem *Das Anegenge*

Principal investigator: Brian O. Murdoch, Lecturer, Dept. of German, U. of Stirling, Stirling, Scotland. *Associates:* J.G.B. Heal, R.A. Wisbey.

Scope: A concordance, reverse index, and rhyme index based on the diplomatic transcription of the *Anegenge* (Vienna MS. 2696) published in the edition by Dietrich Neuschäfer, *Das Anegenge* (Munich: Fink, 1966).

Type of computer: ICT Atlas II. *Size of storage:* 128k. *Language:* Machine language. *Disks or tapes:* 4 Data Products disk files, 6 Potter tape drives. *Special equipment:* Multiprogramming system, multi-access.

Status: To be completed in 1973.

References: "The Production of Concordances from Diplomatic Transcriptions of Early Medieval German Manuscripts," *The Computer in Literary and Linguistic Research,*ed. Roy Wisbey (Cambridge: Cambridge University Press, 1971).

L260. A Concordance of Khotanese

Principal investigator: Professor Dr. Ronald E. Emmerick, U. of Hamburg, Seminar für Geschichte und Kultur des Vorderen Orients, Iranian Section, 2 Hamburg 13, den Rothenbaumchaussee 36, Germany. *Associate:* Harold W. Bailey.

Scope: Concordance and reverse index of the entire corpus of Khotanese literature. *Status:* Continuing.

L261. English Morphotactics

Principal investigator: Magnus Ljung, Lecturer, Dept. of English, U. of Göteborg, Lundgrensgatan 7, Göteborg, Sweden.

Scope: Investigation of the morphotactics in high-frequency English words with a view to establish what principles govern the choice of individual elements in word derivation. The present corpus consists of the 8000 most frequent English words according to the Thorndike-Lorge count. The words were segmented into morphemes and then entered on magnetic tape and processed at Göteborg Datacentral. For each word there is the following information: word, word frequency, word class, prefixes if any, base morpheme, status of base, suffixes if any. The entry "status of base morpheme" refers to the fact that each base morpheme is ranked on a scale ranging from 1 to 6, 1 denoting morphemes which are also words, 6 denoting unique rest morphemes. *Method:* The programs developed to date include a frequency count for morphemes giving both lexical and textual frequencies, an order-class listing of the morphemes in the corpus, and a mapping of the tactics of each individual morpheme. There is also a program exploring the commutation potential for the suffix morphemes.

Type of computer: IBM 360/50. *Size of storage:* 256k. *Language:* FORTRAN IV. *Disks or tapes:* 9-track tapes. Program is available.

References: English Denominal Adjectives (Stockholm: Almqvist & Wiksell, 1970).

L262. ABPILD

Principal investigator: Professor Dr.G. Ungeheuer, Institut für Phonetik and Kommunikationsforschung der Universität Bonn, Adenauerallee 98a, 53 Bonn, Germany. *Associate:* D. Krallmann.

Scope: User-oriented programming system for linguistic data processing, called ABPILD (Allgemeine benutzerorientierte Programmbibliothek innerhalb linguistischer Datenverarbeitung, Automatische Bearbeitung von Prozessen aufgrund informeller linguistischer Deskription). *Method:* 1) Analyze linguistic goals and activities. 2) Interpret them as parts of operating programs. 3) Define program parameters. 4) Write final program. The system ABPILD contains up to now 16 main programs and about 60 subroutines.

Type of computer: IBM 7090/1410. *Size of storage:* 32k. Program is available for IBM 7090; an adaptation for IBM 360/50 and IBM 1800 has been prepared.

Status: Completed December 1972.

References: D. Krallmann, R. Gamp, U. Krumnack, H.E. Neuhaus, "ABPILD-Programmierbibliothek zur linguistischen Datenverarbeitung," Teile I-III, *IPK-Forschungsberichte* 23-25 (Hamburg: Verlag Helmut Buske, 1969).

L263. Lexicographical Analysis of Middle High German

Principal investigators: Professor Dr. W. Besch, Professor Dr. Hugo Moser, Germanistisches Seminar der Universität Bonn; Professor Dr. Gerold Ungeheuer, Institut für Kommunikationsforschung und Phonetik der Universität Bonn, Adenauerallee 98a, 53 Bonn, Germany. *Associates:* U. Krumnack, W. Lenders.

Scope: Lemmatized concordances and indices of Middle High German texts including Konrad von Würzburg, *Gesamtwerk;* Heinrich Wittenwiler, *Der Ring;* Oswald von Wolkenstein, *Die Lieder;* Hartmann von Aue, *Der arme Heinrich, Gregorius, Erec, Büchlein.*

Type of computer: IBM 370/165. *Language:* FORTRAN IV, PL/I, KOMFOR. *Disks or tapes:* About 45 tapes. Program is available.

References: W. Lenders, H.D. Lutz, R. Römer, *Untersuchungen zur Maschinellen Indizierung mittelhochdeutscher Texte,* 2. Auflage (Hamburg: H. Buske Verlag, 1973); W. Lenders, "Verfahren zur automatischen Herstellung klassifizierender Indices," *Literatur und Datenverarbeitung,* Bericht über die Tagung im Rahmen der 100-Jahr-Feier der Rheinisch-Westfälischen Technischen Hochschule Aachen, 1970, hg. v. H. Schanze (Tübingen: Niemeyer, 1972), 45-55; H.-D.Lutz, "Probleme der grammatischen Beschreibung an spämittelhochdeutschen Texten," *Literatur und Datenverarbeitung* (S.O.), 56ff.; W. Lenders, "Lexikographische Arbeiten zu Texten der älteren deutschen Literatur," in *Zeitsch.f.dt.Philologie* 90 (1971); U. Krumnack, W. Lenders, "Herstellung und Verwendung eines automatischen Arbeitswörterbuchs des Mittelhochdeutschen"; Beitrag zur Tagung "Probleme der maschinellen Verarbeitung mittelhochdeutscher Texte," Mannheim, Juni 1971 (erscheint Berlin: im Tagungsbericht, Erich Schmidt Verlag, 1973); U. Krumnack, W. Lenders, "Grammatische Beschreibung und Lemmatisierung spätmittelhochdeutscher Texte mit EDV-Anlagen," *IKP-Forschungsbericht* (erscheint Hamburg: H. Buske Verlag, 1973).

L265. Morphological Analysis of German

Principal investigator: Professor Dr. G. Ungeheuer, Institut für Phonetik und Kommunikationsforschung, Adenauerallee 98a, 53 Bonn, Germany. *Associates:* K.D. Bünting, D. Krallmann.

Scope: Analysis of the morphological structure of German words and evaluation and classification of morpheme patterns. *Method:* 1) Morphological analysis. 2) Formulation of generation rules. 3) Machine generation of possible German words based on a complete list of morphemes. 4) Matching against German dictionary. 5) Evaluation procedure.

Type of computer: IBM 7090, 1410. *Size of storage:* 32k. *Language:* FORTRAN II, FAP. *Disks or tapes:* 720 II, 2 tapes. Program is available.

Status: Completed December 1971.

References: K.D. Bünting, "Morphologische Strukturen Deutscher Wörter," Phil. Diss., Bonn, *IPK-Forschungsbericht* 19 (69/4) (Helmut Buske, Hamburg); K.D. Bünting, "Empirical Investigations of German Word Derivation with the Aid of a Computer," paper presented to the International Conference on Computational Linguistics, Stockholm 1969.

L266. Computational Statistics

Principal investigator: Professor Dr.G. Ungeheuer, Institut für Phonetik und Kommunikationsforschung, Adenauerallee 98a, 53 Bonn, Germany. *Associate:* D. Krallmann.

Scope: Statistical analysis of literary texts; statistical description of stylistic features and evaluation of nonparametric measures. *Method:* 1) N.-gram analysis (for N =1 to 7) within and across word boundaries. 2) Frequency-weighted morphological analysis of vocabularies. 3) Word- and sentence-length investigations. 4) Lexicographical statistics.

Status: Completed 1972.

References: G. Billmeier and D. Krallmann, "Bibliographie zur statistischen Linguistik," *IPK-Forschungsbericht* 18 (69/3), Helmut Buske, Hamburg, 1969; G. Billmeier, "Worthäufigkeiten vom Zipf'schen Typ, überprüft an deutschem Textmaterial," *IPK-Forschungsbericht* 21 (69/6), Helmut Buske, Hamburg, 1969.

L267. Cumulative Dictionary

Principal investigator: D. Krallmann, Professor, Institut für Kommunikationsforschung und Phonetik, Adenauerallee 98a, 53 Bonn, Germany. *Associates:* H.G. Soeffner, D. Thomas.

Scope: Construction of a German dictionary based on about four million running words of present-day German prose and newspapers. The cumulative dictionary (CD) is to be built up from the texts. *Method:* 1) Analyze text for morphological and grammatical information. 2) Sort by morphological, grammatical, frequency, and reference patterns. 3) Match new text against dictionary. 4) Insert new material (after steps 1 and 2) into CD.

Type of computer: IBM 370/165. *Language:* FORTRAN IV, PL/I, KOMFOR. Program is available.

Status: Have analyzed texts and dictionaries for morphological segmentation and grammatical information; have reduced to a restricted set of minimal terms, formatives, classifying these terms according to their functional positions within words; have expanded "language experience" during interaction processes.

References: D. Krallmann, "Linguistische Datenbank und kumulatives Wörterbuch," *Kolloquium Maschinelle Sprachverarbeitung* (Mannheim, 1968); W. Lenders, "Static and Dynamic Lexical Systems" in *Papers of the 2nd International Congress of Applied Linguistics* (Cambridge, 1972); D. Krallmann, "Toward the Development of a Cumulative Dictionary within a Dynamic Information System," *Proceedings of ISLIC-Conference* (Tel Aviv, 1971); U. Krumnack and H.G. Soeffner, "Studien zur Morphologie des Deutschen," *IKP-Forschungsbericht,* Buske,Hamburg, 1972.

L268. Complexity in English Writing

Principal investigator: Louis T. Milic, Professor, Dept. of English, Cleveland State U., Cleveland, OH. *Associate:* Jean A. McConochie.

Scope: Stratified samples (1000 sentences each) from a variety of civil and engineering textbooks and from literary controls. *Method:* A. 1) Code text sentences (one to a card) according to grammatical matrix. 2) Search for predominating tagmemes, inserts and other parameters. B.l)Keypunch natural language text. 2) Prepare frequency distributions of various grammatical categories and items in positional classes.

Type of computer: IBM 360/75. *Language:* SNOBOL4. Program is available.

Status: Completed in 1969.

L269. Reverse Index of Dutch

Principal investigator: Elie R. Nieuwborg, Associate Professor, Institute of Applied Linguistics, Louvain U. (Dutch section), Vesaliusstraat 2, Leuven, Belgium, *Associates:* R. Eeckhout, H. Geeraerts, L. De Busschere.

Scope: 192,344 words from van Dale, *Groot Woordenboek der Nederlandse Taal. Method:* 1) Keypunch. 2) Transfer to tapes. 3) Sort. 4) Print in tabular form.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV. Program is available. [*CHum,* May 1969]

L270. Word Order in Dutch

Principal investigator: Elie R. Nieuwborg, Associate Professor, Institute of Applied Linguistics, Louvain U. (Dutch section), Vesaliusstraat 2, Leuven, Belgium.

Scope: Random samples (5000 sentences) from 20 modern writers (1960-1965). *Method:* 1) Preparation of material: analysis of sentences. 2) Automatic classification of pre-edited sentences.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV. Program is available.

References: De Distributie van het Onderwerp en het Lijdend Voorwerp in het Huidige Geschreven Nederlands (The Distribution of Subject and Direct Object in Modern Written Dutch) (Antwerp: Ed. Plantyn, 1968).

L271. Frequency List of the Dutch Language

Principal investigators: Elie R. Nieuwborg, Associate Professor; Romain E. Eeckhout, graduate research assistant, Institute of Applied Linguistics, Louvain University (Dutch section), Vesaliusstraat 2, Leuven, Belgium.

Scope: 40 random samples of 12,500 words from different written sources of contemporary language. *Method:* The words were first lemmatized, counted per text, and then punched on cards, the frequency in each source being mentioned. Then the data were transferred to tapes and total frequencies computed.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV. Program is available. [*CHum,* May 1969]

L272. Indices and Concordances to Modern Dutch Authors

Principal investigator: Willy J. Martin, Assistant Professor, K.U.L. (Katholieke Universiteit Leuven); Institute of Applied Linguistics, Vesaliusstraat 2, Leuven, Belgium. *Associates:* R. Eeckhout, H. Brems, G. Faes, P. Vermeir.

Scope: Indices and/or concordances to modern Dutch authors (nineteenth-twentieth centuries). In the first phase only poetry will be dealt with. The aim is to give literary scholars and linguists instruments which may help them in the search for objectivity. Up to now an index to "Mei" (32,235-word poem) by Herman Gorter is published, including an index, a frequency list of word forms, a frequency distribution, a dictionary of the binary rhymes (in 2 parts) and a reverse word list. *The Verzamelde Gedichten* (Collected Poems) 1948-1963 of Hugo Claus are now being processed. *Method:* 1) Punching of text. 2) Transfer to tape. 3) Index and/or concordance of word forms. 4) Lemmatized (partly manually) index or concordance. 5) Frequency distributions (lemmata and word forms). 6) Lemmatized frequency list (in descending order, eventually specified according to different sources). 7) Reverse list (of lemmata and word forms). 8) Rhyming dictionaries in fitting cases.

Type of computer: 1) IBM 360/40 and 2) IBM 360/44. *Size of storage:* 1) 128k, 2) 128k. *Language:* PL/I on both computers. Program is available.

References: "Nederlandse indexen via computer" (Dutch Indices by Means of a Computer), *Wetenschappelijke Tijdingen* (Gent), 28 (1969), 11-18; R. Eeckhout and W. Martin, *Woordindex op Mei van Gorter* (Word index to "Mei" by H. Gorter), (Antwerp and Utrecht, 1969); R. Eeckhout and W. Martin, *"Mei" van Gorter, een computerbewerking* ("Mei" by H. Gorter: a computer operation) (Antwerp and Utrecht, 1971); W. Martin, "Het gebruik van indexen in het hoger secundair onderwijs" (The Use of Indices in Secondary Schools), *Persoon en Gemeenschap* (Antwerpen), 24 (1972), 459-463.

L273. Syntactic Analysis of Written English

Principal investigator: Leopold K. Engels, Professor, Institute of Applied Linguistics, Louvain U. (Dutch section), Vesaliusstraat 2, Leuven, Belgium. *Associate:* L. De Busschere.

Scope: Language of essays and of the theatre (after 1960). *Method:* Heuristic method based on linguistic algorithms.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV.

References: "Automatische Analyse van het Engels" (Automatic Analysis of English), *Linguistica Antverpiensia* (1968); "Automatic Analysis of the English Language: Word Groups beginning with Prepositions, Deictic Words and Adjectives," *ITL* 2 (1968). [*CHum,* May 1969]

L274. Quantitative Analysis of the English Language of the Theatre (after 1960)

Principal investigator: Leopold K. Engels, Professor, Institute of Applied Linguistics, Louvain U. (Dutch section), Vesaliusstraat 2, Leuven, Belgium.

Scope: Syntactic, morphological analysis of samples of 2000 syntactic structures (sentences and clauses). Frequency lists of words, word groups and syntactic structures. *Method:* 1) Punch cards. 2) Transfer to tapes. 3) Print in tabular form.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV. [*CHum,* May 1969]

L275. Quantitative Vocabulary Aspects

Principal investigator: Willy J. Martin, Assistant Professor (Katholieke Universiteit Leuven), Institute of Applied Linguistics, Vesaliusstraat 2, Leuven, Belgium.

Scope: The investigation was taken up in the framework of my Ph.D. dissertation (cf. References). The original scope was to establish a battery of tests for quantitative vocabulary study. As a test case the evolution of structure and content in two novels of a modern Dutch writer, Ivo Michiels, was tackled. After finishing my dissertation the project developed towards an investigation into quantitative features of the vocabulary in general, such as trying to find an adequate model for word-length frequencies, testing the so-called objective text parameters given by J. Thiele and K.Sasse (*Grundlagen aus Kybernetik und Geisteswissenschaft* 11[1970], 107-110), investigating the relation word-length in relation to genre, etc. *Method:* 1) Punching of texts. 2) Transfer to tapes. 3) Compilation of frequency data. 4) Testing of hypotheses.

Type of computer: 1) IBM 360/40, 2) IBM 360/44. *Size of storage:* 1) 128k, 2) 128k. *Language:* PL/I,FORTRAN IV on both computers. Some programs are available.

References: Analyse van een vocabularium met behulp van een computer (Analysis of a Vocabulary by Means of a Computer) (Brussels, 1970); "Distributie van nieuwe woorden in Mei van Herman Gorter" (Distribution of "New Words" in the Poem "Mei" by H. Gorter), *De Nieuwe Taalgids* (Groningen), 64 (1971), 163-176; "Some Quantitative Vocabulary Aspects of a Dutch Poem," in *The Computer and Literary Studies,* ed. A.J. Aitken, R.W. Bailey, N. Hamilton-Smith (Edinburgh, 1972).

L276 . Quantitative Analysis of the Vocabulary of the Written Language of Twelve-year-old Pupils

Principal investigators: L.K. Engels, Professor; Willy J. Martin, Assistant Professor; Institute of Applied Linguistics, Katholieke U., Leuven, Vesaliusstraat 2, 3000 Leuven, Belgium.

Scope: Alphabetical frequency list of the words used in a corpus of 2400 written compositions (approximately 125 words each) written by twelve-year-old Dutch-speaking pupils in Belgium. The words are lemmatized and their frequency is specified as to word-forms. Several lexicostatistical investigations (such as the relation between variables like social background, spelling mistakes, vocabulary richness, etc.) will accompany the list.

Type of computer: IBM 360/40 and IBM 360/44. *Language:* PL/I, FORTRAN IV.

Status: Will be ready by the end of 1975.

References: L.K. Engels and W. Martin, *Taalbeheersing bij leerlingen uit het M.O.* (Language Skill of Pupils in Secondary Schools) (Antwerp: Ed. Plantyn, 1968).

L277. The Written Dutch of French-Speaking Pupils in Brussels

Principal investigator: Maurits K. Van Overbeke, Graduate Research Assistant, Institute of Applied Linguistics, Louvain U. (Dutch section), Vesaliusstraat 2, Leuven, Belgium.

Scope: The total amount of deviation from the Dutch norm caused by the pupil's first language, French; a resistance scale of French against Dutch according to different areas of the target language (Dutch). *Method:* 1) Compilation of the corpus (900 Dutch compositions, each having approximately 125 words, of eighteen-year-old French-speaking Brussels pupils). 2) Punching. 3) Selection of deviations. 4) Indicating contrastive features by means of code. 5) Transferring to tapes. 6) Comparison and printing of results.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV. [*CHum,* May 1969]

L278. Statistical Analysis of Greek Hexameter Verse

Principal investigator: Frank Pierce Jones, Professor Emeritus, Dept. of Classics, Tufts U., Medford, MA 02155. *Associate:* Florence E. Gray.

Scope: A statistical analysis was made of the distribution of the 32 hexameter patterns in Homer to test the probability of single as opposed to dual or multiple authorship of the *Iliad* and *Odyssey*. With frequency of pattern (taken individually or in various combinations) as the criterion, comparisons were made between single books of the *Iliad* and *Odyssey;* between the *Iliad* and *Odyssey* as wholes; and between the works of Homer (including the *Hymns*) and the works of Hesiod, Callimachus, Aratus, and Apollonius. *Method:* Pattern counts for Homer, Hesiod, Callimachus, Aratus, and Apollonius were made from the tabulations of La Roche in *Wiener Studien* and the data were punched onto IBM cards. Programs for *chi*-square and product-moment correlation (*r*) were used to test significance of differences. The results of the analysis support the hypothesis of single authorship.

Type of computer: IBM 1130. *Size of storage:* 8k. *Language:* FORTRAN IV. *Disks or tapes:* One 512k disk. *Special equipment:* 1132 printer; 1627 plotter. Program is available.

References: "Notes on the Input for an Automatic Scansion Program," *Revue* (LASLA) (Liège, 1968), 4, pp.5-11; "Hexameter Patterns, Statistical Inference, and the Homeric Question," accepted for publication in *Transactions of the American Philological Association.*

L279. Concordance to Céline, *Voyage au bout de la nuit*

Principal investigator: Paul A. Fortier, Assistant Professor, Dept. of Romance Languages, U. of Manitoba, Winnipeg, MB, R3T 2N2, Canada.

Scope: To produce a complete concordance, including all accent, punctuation, and capitalization symbols as they are presented in a normal printed text for use in thematic analysis of the novel. *Method:* Each text line was encoded and punched, one line per card. The concordance was produced by BIBCON, a concordance program developed at the University of Wisconsin by Professor Richard L. Venezky.

Type of computer: CDC 3600. *Size of storage:* 65k. *Language:* FORTRAN. *Disks or tapes:* 3 tapes. Copies of the Concordance available from Professor Fortier. See *Computers and the Humanities* 3 (1969), 129-138.

Status: Completed December 1971.

References: A book-length study in French of thematic structures in *Voyage au bout de la nuit* has been completed with the help of the concordance.

L280. Quantitative Studies in Prosody: Alliteration, Assonance, Consonance—Spenser to Auden

Principal investigator: John D. Allen, Professor (retired), Dept. of English, East Tennessee State U., Johnson City, TN 37601.

Scope: Determining frequencies in occurrence of alliteration, assonance, and consonance in selected verse of 24 major English and American poets, and of frequencies of nine elements of prosody in the blank verse of 10 major English and American poets. *Method:* Random selection of passages, counting of instances (for some sorts of data definitions had to be devised), tabulating of data. The computer was used in tabulation.

Type of computer: IBM 1401. *Size of storage:* 8k. *Language:* AUTOCODER VI-M13. [*CHum,* May 1969]

L281. A Concordance to the Poetical Works of John Keats

Principal investigator: Michael G. Becker, graduate student, Dept. of English, U. of Wisconsin, Madison, WI 53705. *Associate:* Todd K. Bender.

Scope: To prepare a computer concordance to the poetical works of John Keats based on the edition of *The Poetical Works of John Keats,* ed. H.W. Garrod (Oxford, 1958). This work will be one in a series of concordances being prepared at the University of Wisconsin following the format described by Richard L. Venezky in the January 1969 issue of *Computers and the Humanities.* It will include all variants in the Keats text and will utilize both upper and lower case.

Type of computer: CDC 3600. *Language:* BIBCON, written in FORTRAN IV. Program is available. [*CHum,* May 1969]

L282. A Computerized Language Analysis System

Principal investigator: George A. Borden, Associate Professor, Dept. of Speech, Pennsylvania State U., University Park, PA 16802. *Associate:* James Watts.

Scope: A system of programs has been written to analyze a natural language text. Printout includes number of words processed, type/token ratio, mean and standard deviation of word length, rank-frequency ordering of words and percentage of total words in each rank, total sentences contained in the text with breakdown into number of each type of sentence, mean and standard deviation on number of words in these sentences, and the number of words per sentence

in the order of occurrence. The option is then given to print out an alphabetical concordance or index of all the words in the text, or any specifiable subset of these words, and a listing of these words in the order of descending frequency. Rank-frequency distributions, percentage of words in each rank, and number of words in each sentence are also punched on cards for further analysis.

Type of computer: IBM 360/67. *Size of storage:* 280k. *Language:* PL/I. *Disks or tapes:* 3 2311 disks; 3 9-track tapes.

References: "A Computerized Language Analysis System," *Computers and the Humanities* 5 (January 1971),129-141.

L283. Structuring Semantic Space

Principal investigator: George A. Borden, Associate Professor, Dept. of Speech, Pennsylvania State U., University Park, PA 16802. *Associates:* William Nelson, James Watts.

Objective: To develop the theoretical constructs intrinsic to the superordinate concepts defining semantic space. It is focused on determining the abstraction level of terms found to cluster in specified areas of semantic space. *Method:* Involves experimentation with human subjects and word clustering and tree structuring by computer,using a *Roget's Thesaurus* and various dictionaries.

Type of computer: IBM 360/67. *Size of storage:* 280k. *Language:* PL/I. *Disks or tapes:* 3 2311 disks; 3 9-track tapes. *Special equipment:* CalComp plotter.

Status: Cancelled.

References: William Nelson, "Topoi: Evidence of Human Conceptual Behavior," *Philosophy & Rhetoric* 2, 1 (1969), 1-11; George Borden, "Theoretical Constructs: A By-product of Information Retrieval," *Proceedings of the American Documentation Institute* 4 (1967), 180-185. [*CHum,* May 1969]

L284. Computational Linguistics

Principal investigator: H. Brandt Corstius, Dept. of Computer Science, Mathematical Centre, Boerhaavestraat 99, Amsterdam, Netherlands.

Scope: Formal grammars for parts of the Dutch language. *Method:* Tests of programs on newspaper prose.

Type of computer: Electrologica X8. *Size of storage:* 32k + drum. *Language:* ALGOL 60. [*CHum,* May 1969]

L285. Contrastive Analysis of Serbocroat and English

Principal investigator: Rudolf Filipović, Professor, Institut za lingvistiku, Filozofski fakultet, Zagreb U., Djure Salaja bb, Zagreb, Yugoslavia.

Scope: Contrasting English original (Brown Corpus, reduced by 50 percent) and its Serbocroat translation. To be followed by a similar treatment of a smaller control corpus of Serbocroat originals and their English translation. It is hoped to obtain clear patterns usable for a contrastive description of the two languages (eventually to appear in book form). *Method:* 1) Prepare a magnetic tape (via a Flexowriter tape) of Serbocroat text. 2) Produce complete ordinary KWIC concordances of both texts. 3) Produce a contrastive concordance (with alternating lines from either language) of preceding concordances. 4) Produce selective reverse KWIC concordances of both texts. 5) Produce a contrastive concordance of preceding concordance. (Stages 3 and 5 may be implemented largely through manual sorting of Flexowriter-produced sentence slips.) 6) Repeat the same steps for control corpus.

Type of computer: IBM 360/30. *Size of storage:* 32k. *Language:* Assembler. *Disks or tapes:* 3 2311 disks; 2 2415/4 tapes.

References: "Contrastive Analysis of Serbo-Croatian and English," *Studia Romanica et Anglica Zagrabiensia* (Zagreb) July 23, 1967, pp.5-27; Z.Bujas, *Concordancing as a Method in Contrastive Analysis,* ibid., pp.49-62.

L286. Concordance to Beckett, *En attendant Godot*

Principal investigator: Paul A. Fortier, Assistant Professor, Dept. of Romance Languages, U. of Manitoba, Winnipeg, MB, R3T 2N2, Canada. *Associates:* M.R. Day, R.G. Palmer.

Objective: To produce a research tool permitting serious thematic studies. *Method:* Keypunch text on 80-column cards, using end codes to distinguish diacritics; transfer to magnetic tape; produce concordance on tape; print. Different type styles, punctuation, and all capitals are retained; diacritics and significant capitals are retained and distinguished.

Type of computer: IBM 360/50. *Size of storage:* 256k. *Language:* PL/I (OS). *Disks or tapes:* 3 tapes. Program is available; for input contact Professor Fortier; for program, contact Michael Day, Manager of the Computer Centre, University of Saskatchewan, Regina Campus, Regina, SK, Canada.

Status: Completed June 1971. The keyword concordance to *En attendant Godot* has recently been upgraded to a grammatical concordance.

References: "Beckett emule de Gide," *L'Esprit Créateur* 11, 3 (Fall 1971), 55-66.

L287. Concordance du Théâtre de Pierre Corneille

Principal investigator: Bryant C. Freeman, Associate Professor of French and Chairman, Dept. of French and Italian, U. of Kansas, Lawrence, KS 66044. *Associate:* Lynne Huxsaw.

Scope: Complete theatre of Corneille, including all variants; based upon the Grands Ecrivains de la France edition. Word-indices for "common words"; frequency list by descending order; analytical table for each play.

Type of computer: Burroughs B5500. *Size of storage:* 192k-character core memory. *Language:* COBOL and ALGOL. *Disks or tapes:* 19,200,000-character disk; 4 66-KC IBM-compatible tape drives. *Special equipment:* Friden Flexowriter input; Photon output. Program is available.

References: Concordance du Théâtre et des Poésies de Jean Racine, 2 vols (Ithaca, NY: Cornell University Press, 1968). (See L48) [*CHum,* May 1969]

L288. Computer Analysis of Latin Poetry

Principal investigator: Nathan Abraham Greenberg, Professor, Dept. of Classics, Oberlin College, Oberlin, OH 44074.

Scope: Automatic scansion of dactylic hexameter poetry and the production of metrical indices, followed by statistical analysis.

Type of computer: IBM 360/44G. *Size of storage:* 128k. *Language:* FORTRAN IV-D. *Disks or tapes:* 6 Potter 4314 disks; 2 Potter 2405 9-track tapes. Program is available.

References: "Virgil and the Computer: Fourth Foot Texture in *Aeneid* I," *Revue* (Organisation Internationale pour l'étude des langues anciennes par ordinateur) (1967) 1, 1-16;"Scansion purement automatique de l'hexametre dactylique," *Revue* (1967), 3, 1-30; "Words and Syllables: Four Eclogues," *Revue* (1970), 2, 5-49; "Metrical Expectations in the *Ars Poetica,*" *Revue* (1970), 3, 111-129; "The Hexametrical Maze," *Revue* (1970) 4, 17-63; "Epanastrophe in Latin Poetry," *Revue* (1972), 2, 1-17.

L289. VERTEX: A Set of MAD Functions to Facilitate the Processing of Natural-Language Text [Deleted] [*CHum,* May 1969]

L290. Hebrew Computational Stylistics: Author Attribution Problems

Principal investigator: Asa Kasher, Dept. of Philosophy, Bar-Ilan U., Ramat-Gan, Israel.

Scope: Study of Hebraic texts of disputed authorship. The first text analyzed, processed, and studied is *The Book of Isaiah. Method:* Statistical and clustering analyses of morphological features.

Type of computer: IBM 360/20, IBM 360/40, IBM 360/50. *Size of storage:* 8k, 264k. *Language:* Assembler, FORTRAN, SNOBOL. Program is available.

Status: Completed 1970.

References: "The Book of Isaiah: Characterisation of Authors by Morphological Data Processing," *Proceedings of the (Israeli) National Conference on Data Processing, 1969* (Hebrew; English version available from the author).

L291. Research in Computational Linguistics

Principal investigator: Susumu Kuno, Associate Professor, Linguistics, Division of Engineering and Applied Physics, Harvard U., Cambridge, MA 02138. *Associates:* George Lakoff, Sheila Duncan, Tamotsu Mukaii. *Scope:* Research in mathematical models of grammar and the theory of transformational grammar. These two areas of research are closely interwoven. On the one hand, what mechanisms a theory of grammar for linguistic description should have depends upon what features—semantic, syntactic, and phonological—natural languages display. On the other hand, studies in mathematical models of grammar tell us what types of sentences a given model of grammar can account for, whether these types of sentences are those encountered in natural languages, whether a recognition procedure exists for the language of the grammar and so forth. Since the theory of transformational grammar has been undergoing constant revisions, due to the discovery of new types of phenomena in natural languages, and since the mechanisms of the theory have become more and more complex, constant interactions between the linguistic theory and mathematical models of

grammar have become even more important, so that it can be guaranteed that the former is a mathematically feasible model for linguistic description. *Method:* Such interactions are provided by two computer systems which have been developed at Harvard: a transformational grammar tester and a phonological rule tester. In the first system, the linguist may compose a transformational grammar by entering the rules which comprise it, and he may then test these rules by entering sentences. A derived phrase-marker is produced for each of these sentences. As a research tool, the phonological system may be used by a linguist in testing his phonological rules, to see if they perform their intended function. The program accepts the derived surface structure of a sentence as input, and cyclically applies a set of phonological rules, which the user may enter, to succeeding domains of the input structure until the entire structure has become the phonetic representation of the sentence.

Type of computer: Brand on-line time-charged IBM 7094 with attached model 35 or standard 1050 typewriter. *Size of storage:* 32k. *Language:* SNOBOL compiler and MAD-SLIP compiler respectively. Program is available on special request at Harvard. [*CHum,* May 1969]

L292. Glossary of the Middle English Romance *Gamelyn*

Principal investigator: Angela M. Lucas, Lecturer, Dept. of English, St. Patrick's College, Maynooth, Co. Kildare, Ireland. *Associate:* Peter J. Lucas.

Scope: Glossary of the complete text of *Gamelyn* edited from MS.Harley 7334. *Method:* 1) Transliterate original. 2) Keypunch text onto cards. 3) Sort alphabetically. 4) Post-edit typewritten printout.

Type of computer: IBM 1620 Model I. *Size of storage:* 40k. *Language:* SPS II. *Disks or tapes:* 1 paper tape.

Status: Completed September 1969.

L293. A Textual Comparison of Two Manuscripts of *El Libro de Alexandre*

Principal investigator: Louis F. Sas, Professor, Dept. of Romance Languages, City College, CUNY, 139 St. and Convent Ave., New York, NY 10031.

Scope: A textual comparison of two medieval Spanish manuscripts of *El Libro de Alexandre,* one in Paris and one in Madrid, so that only the differences or variants in 0 (Madrid) which do not appear in the P (Paris) manuscript appear on the list. This will help in the compilation of a dictionary of *El Libro de Alexandre* with every word in both manuscripts (and the variants) defined and studied semantically and morphologically. [*CHum,* May 1969]

L294. A Test of Rule Ordering for Phonological Changes in Russian from PIE to the Present

Principal investigator: Raoul N. Smith, Assistant Professor, Dept. of Linguistics, Northwestern U., 621 Foster St., Evanston, IL 60201. *Associate:* Marvin Kantor.

Method: Input is a lexicon of PIE forms which are passed through a set of rules. After each rule the changed forms are then available as input to the next rule in order. The major purpose is to test the rules themselves and their ordering.

Type of computer: CDC 6400. *Size of storage:* 65k. *Language:* SNOBOL4.

References: "A Computer Simulation of Phonological Change," *ITL-Review for Applied Linguistics* 5 (1969), 82-91.

L295. Dictionarium Tetraglotton (1562)

Principal investigator: F. de Tollenaere, Director Institute for Dutch Lexicologie, Section Thesaurus, Stationsweg 39a, Leiden, Holland.

Scope: To make an alphabetic list of the Dutch words of the dictionary.

Type of computer: IBM 360/50. *Language:* PL/I.

Status: Postponed.

L296. Vulgata-Konkordanz

Principal investigator: Bonifatius Fischer, Vetus-Latina-Institut, D7207 Beuron, Deutschland. *Associate:* Wilhelm Ott.

Scope: Wortkonkordanz sur Vulgata. Textgrundlage: Ausgabe von R. Weber u.a., die 1969 in der Württembergischen Bibelanstalt Stuttgart erscheinen wird. Die in der Ausgabe angeführten textkritischen Varianten werden in die Konkordanz mit aufgenommen. *Method:* Texterfassung auf 8-Kanal-Lochstreifen. Der Text wird zweimal

abgeschrieben, um die Korrektur zu automatisieren; Aufnahme beider Lochstreifen auf Magnetband; Aufbereiten des Textes: Ergänzen der kompletten Stellenangabe zu jeder Zeile; Kontrolle der Mss-Reihenfolge in den kritischen Angaben; Durchnumerierung der Zeilen; Ausführen von Korrekturanweisungen. Output: je ein Magnetband für beide Versionen, Protokoll der festgestellten Fehler; Vergleichen der beiden Textversionen. Output: a) Textausdruck, in dem beide Versionen vereinigt sind, als Grundlage für die Korrektur; er enthält nur an den Stellen beide Textfassungen, an denen Abweichungen und mit einer Vorentscheidung über die wahrscheinlich richtige Fassung. b) Magnetband gleichen Inhalts; Erstellen eines Lochstreifens, der Korrektur-Anweisungen enthält. Herstellen eines fehlerfreien Textes aus 4b) und 5). Output: Magnetband; Textzerlegung, Zurückführen der flektierten Wortformen auf die zugehörigen, Grundformen mit Hilfe des Lexicon Electronicum Latinum von R. Busa, Gallarate. Output: a) Schnelldruckerprotokoll, b) Magnetband; Ergänzen fehlender bzw. Korrektur falscher Grundformen; Sortieren; Wörter und Kontext mit Stellenangabe werden in die für den Druck gewünschte Form gebracht. Output: Magnetband für Lichtsatz.

Type of computer: CDC 3300. *Size of storage:* 32k. *Language:* FORTRAN, COMPASS. *Disks or tapes:* 3 854 disks; 4 604 tapes. *Special equipment:* Paper tape reader. Program is available.

References: Bericht 2, Vetus-Latina-Institut der Erzabtei Beuron (Beuron, 1968), 11-14; Bericht 3 (1969), 13-16; Bericht 4 (1971), 14-15; Bericht 5 (1972), 12-13; Wilhelm Ott, "Erfassen und Korrigieren von Texten auf Lochstreifen, Erfahrungen mit der Vorbereitung des lateinischen Bibeltextes für den Computer," in *Electronische Datenverarbeitung* 12 (1970), 132-137; Wilhelm Ott, "Transcription and correction of texts on paper tape—Experiences in preparing the Latin Bible text for the computer," in *Revue* (LASLA) 1970, 2, 51-66.

L297. Programming Routines for Poetic Stylistic Analysis

Principal investigators: R.A. Sencer, Associate Professor; Paul Ochler, Ph.D. candidate; Dept. of Language and Literature, Rensselaer Polytechnic Institute, Troy, NY 12181.

Objective: To attempt to program the computer to "pull out" metaphors from some short poems. We have to date only accomplished what others have done: listings of words, alphabetizing, words in context (which we hope is the key to metaphoric analysis), and a simple concordance analysis.

Type of computer: IBM 360/50. *Size of storage:* 1162k, with 12k as supervisor. *Language:* FORTRAN IV (we hope to translate to PL/I). [*CHum,* November 1969]

L298. Computational Linguistics

Principal investigator: E.B. James, Assistant Professor, Dept. of Computing Science, Imperial College, 48, Princes Gardens, London, S.W.7, U.K.

Scope: 1) To produce a standard package of programs for the usual literary data processor's requirements with emphasis on convenient use by the inexperienced programmer or linguist with basic computer experience. This includes simple means of specifying text units, processing a unit and counting occurrences of any linguistic feature; presenting results in the form of graphs, with full annotation. 2) To investigate the tolerance of computer language processors to errors in statements; to develop more tolerant compilers which will enable the use of "more natural" programming languages; to feed back the experience gained into natural language structure studies.

Type of computer: CDC 6400, CDC 6600. *Size of storage:* 128k. *Language:* FORTRAN IV. *Special equipment:* Kingmatic, CalComp, and microfilm plotters. Program is available for Part 1, but not yet documented for external use.

Status: The individual components of the package (Part 1) are now completed, and a simple-to-use language is now being developed to enable all the components to be brought into action via a remote terminal.

References: E.B. James and Christine Allwright, "A Frequency Count Package for Literary and Linguistic Research: An Example of Literary Program Design," *The Computer in Literary and Linguistic Research,* ed. R.A. Wisbey (Cambridge: Cambridge University Press, 1971); E.B. James and D.P. Partridge, "Adaptive Correction of Program Statements," *Communications of the ACM,* January 1973.

L299. Rhythmical Ordering as an Index to Literary Style

Principal investigator: Ronald C. Turner, Associate Professor and Chairman, Dept. of Modern Languages, Whitworth College, Spokane, WA 99218.

Scope: To measure the degree of rhythmical patterning as a parameter of style in written Spanish. The study utilizes the Shannon expression of relative entropy as a means of determining the degree of metrical freedom, looseness, or unexpectedness. Current findings from samples totaling 10,000 words have revealed that levels of expression (e.g., poetry, novel, essay, journalistic prose) are readily distinguishable by the computer, using the information theory approach. Future studies will examine various modes of written expression by several authors. *Method:* For a given passage, an initial program reads natural Spanish text and performs a phonemic transcription and rhythmical scansion of each phrase. Another program records the occurrences of all permutations of rhythmical syllables in strings of length one

through seven. The seven entropies (one for each length of string) are calculated and compared with the corresponding entropies of other passages to determine the degree of metrical patterning for the passage.

Type of computer: IBM 360/67. *Size of storage:* 200k (partition available to normal job). *Language:* SNOBOL4, Version 2.0. *Disks or tapes:* 1 2314 Direct Access Facility disk; 5 9-track tape drives. Program is available.

L300. Historical Collation of Editions of *A Midsummer Night's Dream*

Principal investigators: R.L. Widmann, Assistant Professor, Dept. of English, U. of Pennsylvania, Philadelphia, PA 19104.

Scope: To facilitate eye-collation of selected (26) editions of *A Midsummer Night's Dream* from one of 1600 to the forthcoming Evans text (Houghton Mifflin). *Method:* All editions are punched onto punch cards, one line of text per card, with caps and italics denoted by special symbols. The computer then compares the various editions, letter by letter, and notes the variations.

Type of computer: IBM 360/75. *Size of storage:* 230k. *Language:* Initially FORTRAN IV, with a text-processing program called "FORMAT" (from IBM), but also PL/I. *Special equipment:* IBM 1403 printer with TN print train.

References: Shakespearean Research and Opportunities 5-6 (1970-71), 53-59.

L301. Concordance to the *Roman de la Rose* of Guillaume de Lorris

Principal investigator: Joseph R. Danos, Doctoral Candidate, Dept. of Romance Languages, U. of North Carolina, Dey Hall, Chapel Hill, NC 27514.

Scope: To prepare a computer concordance to the *Roman de la Rose* of Guillaume de Lorris accompanied by stylistic analysis, with an eye to an eventual second volume on Jean de Meung's portion and a comparative study of the two. *Method:* The text is put into upper and lower case by Format, then a rough concordance is generated by a PL/I program which also includes a proofreading routine that marks low-frequency combinations. The homographic forms are noted manually and then separated by a PL/I program that uses the rough concordance as input and reads the word and line number. For example, "a," the verb, is a separate entry from "a," the preposition.

Type of computer: IBM 360/75.

Status: Was to be completed June 1973.

L302. Collocations as a Measure of Stylistic Variety

Principal investigator: Louis T. Milic, Professor, Dept. of English, Cleveland State U., Cleveland, OH 44115. *Associate:* Peggy I. Haskel.

Scope: The purpose is to compile a dictionary of standard collocations for a limited list of words. For each entry the dictionary will provide a list of the open-class collocating words and show the frequency of each. This dictionary will then provide the standards or norms against which the collocations of modern writers of expository prose will be measured and will thus constitute a measure of departure from the norm, or originality. *Method:* Scan corpus for occurrence of proposed dictionary entries; collect the contextual open-class words; analyze these collocations linguistically and make statistical frequency distributions.

Type of computer: IBM 360/91. *Size of storage:* 512k. *Language:* SNOBOL4. Program is available.

Status: Completed in 1972.

References: Peggy I. Haskel, "Collocations as a Measure of Stylistic Variety," *The Computer in Literary and Linguistic Research*, ed. R.A. Wisbey (Cambridge: Cambridge University Press, 1971), pp. 159-168.

L303. Parallelism in Present-Day American Prose

Principal investigator: Louis T. Milic, Professor, Dept. of English, Cleveland State U., Cleveland, OH 44115. *Associate:* Mary P. Hiatt.

Scope: Investigation of intra-sentence parallelism in present-day American prose to discover what formal signals on rhetorical and linguistic level define parallelism and to what extent its use in present-day American prose is different from traditional norms. The corpus will be a tape of *A Standard Sample of Present-Day Edited American English, for Use with Digital Computers,* prepared at Brown University. *Method:* Scan plain text for presence of signals of parallelism, such as "not only . . . but also" and collect adjacent contexts. Analyze contexts linguistically for parallelism of grammatical construction or of word-form class. Code contexts as consisting of certain types of parallelism and construct frequency distributions of types according to both linguistic and rhetorical criteria.

Type of computer: IBM 360/91. *Size of storage:* 512k. *Language:* SNOBOL4. Program is available.

Status: Completed April 1971.

References: "Computer Analysis of Intra-sentence Parallelism in Present-Day American Prose," *Language and Style* (in press).

L304. Frequency Dictionary, Verbal Index and Concordance of *Candide*

Principal investigator: Pierre R. Ducretet, Associate Professor, Dept. of French, University College, U. of Toronto, Toronto, ON, Canada. *Associate:* Marie-Paule Ducretet.

Objective: To establish a frequency dictionary, a verbal index and a concordance of Voltaire's *Candide*. Also to furnish some quantitative information about this text in the form of tables, charts, or graphs. *Method:* Punching text onto cards, separating text into words, adding a lemma and a grammatical code to each word. Using computers and related equipment for sorting and other manipulations.

Type of computer: 1) IBM 370/165, 2) IBM 360/65. *Size of storage:* 1) 20,000k, 2) 1024k. *Language:* 1) PL/I, Assembly, G, SPITBOL, 2) ATS, CRJE, APL. *Disks or tapes:* 1) 12 (8-2314/4-3330) disks and 6 2400 tape drives; 2) 8 2314 disks and 2 2400 tape drives. *Special equipment:* TN print chain. Program is available.

References: "Computers and Literary Studies: Another View," *Computers and Old English Concordances* (Toronto: University of Toronto Press, 1971); "Quantitative Stylistics—An Essay in Methodology," *Computers and the Humanities* 4,3 (January 1970), 187-191; "Concordance de *Candide*—Introduction (Méthodologie)" (Unpub. doctoral thesis, Sorbonne, 1968).

L305. Frequency Dictionary, Verbal Index (and Concordance) of Montaigne's *Essais*

Principal investigator: Pierre R. Ducretet, Associate Professor, Dept. of French, University College, U. of Toronto, Toronto, ON, Canada. *Associate:* Marie-Paule Ducretet.

Objective: To establish an index of all the words in Montaigne's *Essais* and give their location in the P. Villey edition (1922). Also to furnish some quantitative information, in the form of tables and charts or graphs. *Method:* Punching text onto cards; separating text into words; using computers and related equipment for sorting and other manipulations.

Type of computer: 1) IBM 370/165 and 2) IBM 360/65. *Size of storage:* 1) 30,000k, 2) 1024 k. *Language:* 1) PL/I, Assembly, G, SPITBOL; 2) ATS, CRJE, APL. *Disks or tapes:* 1) 12 (8-2314/4-3330) disks and 6 2400 tape drives; 2) 8 2314 disks and 2 2400 tape drives. *Special equipment:* TN print chain. Program is available.

References: See L304.

L306. Analysis of Essays by Computer

Principal investigator: Ellis B. Page, Professor and Director, Bureau of Educational Research, U. of Connecticut, Storrs, CT 06268. *Associates:* Dieter H. Paulus, Paul Barbuto.

Scope: We have moved from work dealing exclusively with free essays, written for English classes, to consider problems of subject-matter content. John McManus and Donald Marcotte have been working with short identification answers in history finals, and have demonstrated that a number of different strategies may match the performance of the human evaluator. Work is now deepening into problems of data storage and comparison. *Method:* Both statistical and deterministic strategies are currently under study. Past techniques have used multiple regression to optimize simulation. Effort is to reach formal models which can, ultimately, be practically substituted for much human evaluation and comment.

Type of computer: IBM 360/65. *Size of storage:* 512k bytes in core. *Language:* FORTRAN IV, PL/I. *Special equipment:* CPS terminals. Past programs are listed in final report available from us. Others are partially available. [*CHum*, November 1969]

L307. Graphematics of Early Modern High German

Principal investigator: Ilpo T. Piirainen, Researcher of the U. of Helsinki, Institute for Linguistics, U. of Münster, Domplatz 23, 44 Münster, West Germany.

Scope: With the textual material of about 150,000 words by an unknown writer of the sixteenth-seventeenth century in Germany, the researcher demonstrates how far the computer proves successful the linguistic analysis of earlier linguistic periods. Comparing the graph systems of Middle High German and the text under research, the researcher develops a method of historical graphematics. The graphematic oppositions are being grouped as to their positions and analyzed statistically, and their specific functions within a closed system are being investigated. A successive order of the

linguistic periods under comparison is not required. *Method:* The graphs of the text are sorted by the computer and compared with the graphs of Middle High German in the same position and function. The method, besides facilitating the techniques of handling texts, makes possible an exact description of the graphs and thus provides a solid basis for further linguistic analysis.

Type of computer: IBM 7090/7094; Zuse Z 23. *Language:* FORTRAN II/FAP; Machine oriented. For IBM 7090/94 program "Index" at the DRZ (German Computer Centre), Darmstadt, is available.

References: Graphematische Untersuchungen zum Frühhochdeutschen [=Studia Linguistica Germanica No.1] (West Berlin: Walter de Gruyter & Co.); "Glossarharstellung mit Hilfe des Computers: Ein Beispiel aus dem Frühneuhochdeutschen," *Neuphilologische Mitteilungen* (Helsinki, 1968), No.4.

L308. A Computer Concordance of the Works of Geoffrey Chaucer

Principal investigator: T.J. Ray, Professor, Dept. of English, U. of Mississippi, University, MS 38677. *Associates:* Darryl Shreve, John Notaro, Al Trussel.

Objective: To produce a tape, initially only of the F.N. Robinson edition, of the works of Chaucer; to produce first a means to a computer concordance and later a tape to be used in stylistic analysis.

Type of computer: IBM 360/30. *Size of storage:* 32k. *Language:* PL/1. *Disks or tapes:* 4 2311 disk drives.

Status: Project temporarily halted with about 20,000 cards punched.

L309. Modern German Newspaper Language

Principal investigator: Inger Rosengren, Docent, Dept. of German (Institutionen för tyska språket), U. of Lund, Helgonabacken 14, Lund, Sweden.

Scope: A frequency dictionary of modern German newspaper language based on *Die Welt* and *Süddeutsche Zeitung* from November 1966 to October 1967. *Method:* Input: 6-channel tape (linotype) used for newspaper typesetting. Each article is stored on magnetic tape. The running words are sorted alphabetically and numerically with manual and automatic correction of the texts and lists; cumulation of all lists, homograph separation and lemmatization. Output: a numerical and alphabetical frequency dictionary printed as an offset.

Type of computer: SAAB D21. *Language:* DAC, ALGOL. Program is available.

References: "Ein Frequenzwörterbuch der modernen Zeitungssprache: wie und wozu?" *Beiträge zur Linguistik und Informationsverarbeitung* 14 (1968); A. Zettersten, "Current Computing Activity in Scandinavia Related to Language and Literature Research," *Computers and the Humanities* 3 (1968), 59. [*CHum,* November 1969]

L310. Concordance to the Works of Theodore Roethke

Principal investigator: Edward P. Stanulis, Instructor, Dept. of Humanities, Michigan Technological U., Houghton, MI 49931. *Associate:* Richard J. Newell.

Scope: Computer-generated concordance to Roethke's poetry with complete word-frequency list. Output is KWIC. Current state of project: 60 percent complete.

Type of computer: IBM 360/44. *Size of storage:* 128k bytes.*Language:* FORTRAN IV G. *Disks or tapes:* 2 IBM type 2311 disks; 2 IBM type 2400 tapes. [*CHum,* May 1969]

L311. A Concordance to the *Book of Common Prayer* of the Episcopal Church of the United States

Principal investigators: Milton Huggett, Assistant Professor, Dept. of English; James M. Pye, Programmer/Analyst, Data Processing Center, Texas A & M U., College Station, TX 77843.

Objective: To produce a concordance that will be of value not only to the clergy but to literary scholars as well. A line-scanning technique was devised that produced, for the most part, a complete syntactic unit as associated context. *Method:* The text was punched onto cards in line-image form with identifying line and page numbers. A tape built from the cards became input to the scanning routines. Certain function words as well as repetitive rubrics were omitted.

Type of computer: IBM 360/165. *Size of storage:* 512k. *Language:* COBOL F. *Disks or tapes:* 1 2314 D.A.S.F. disk, and 4 2401/2 tapes. Program is available. [*CHum,* November 1970]

L312. A Concordance to the Poetry of Sir Philip Sidney

Principal investigator: Herbert S. Donow, Assistant Professor, Dept. of English, Southern Illinois U., Carbondale, IL 62901.

Scope: To produce an old-spelling concordance of the poetry of Sir Philip Sidney based on the text of the poems in William A. Ringler's edition (Oxford). *Method:* Text is encoded with all variants, keypunched and stored on magnetic tape. Text is then processed by concordance system which also formats the output, right and left justifies text, creates running heads (right and left justified) with page numbers. (System has capability to assist in modernizing the spelling of index if desired.)

Type of computer: IBM 7044, IBM 1401. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 2 1301 disk files; 10 729 Mod.5 drives. The programs are functioning but are in the process of being edited and generalized. Documentation is also being prepared.

References: "Concordance and Stylistic Analysis of Six Elizabethan Sonnet Sequences," *Computers and the Humanities* 3 (March 1969), 205-208. [*CHum,* November 1970]

L313. The Graphic System of American English

Principal investigator: Arne Zettersten, Docent, Dept. of English, U. of Lund, Helgonabacken 14, Lund, Sweden.

Scope: Various statistical studies of the graphic system of English based on the Brown University Standard Corpus of Present-Day American English.

Type of computer: CDC 3600. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 6 tapes.

References: A Statistical Study of the Graphic System of Present-Day American English (Lund: Studentlitteratur, 1969). [*CHum,* November 1969]

L314. On the Theory of Finite State Languages with Application to the Italian Language

Principal investigator: Arnim von Stechow, Institute for Linguistics, U. of Münster, Domplatz 23, 44 Münster, West Germany.*

Scope: A survey of the most important results of linguistic investigation with automata theory; a right linear grammar which covers a part of the Italian language for the generation and recognition of sentences. Algorithms for the elimination of the ambiguity of right linear grammars are realized. This investigation is characterized especially by the trial of the combination of grammars with finite transducers. *Method:* Automata theory and the theory of context-free languages.

Type of computer: IBM 360. *Language:* PL/I. Program is available.

References: Masters thesis (German, mimeo). [*CHum,* November 1970]

L315. Renaissance Fable Literature. Part I. English Fable Books, 1484-1650 [*CHum,* November 1970] [Deleted]

L316. Automatic Retrieval of Dutch Compound Verbs

Principal investigator: Elie R. Nieuwborg, Associate Professor, Institute of Applied Linguistics, Louvain U. (Dutch section), Vesaliusstraat 2, Leuven, Belgium. *Associate:* L. De Busschere.

Objective: Euroterm Index.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV. Program is available. [*CHum,* November 1969]

L317. Automatic Production of Instant Education (Computers, Syllabi, and Lexical Non-Trivia)

Principal investigator: H. Frank, Professor, Institut für Kybernetik, 479 Paderborn, Riemekestrasse 62, West Germany. *Associates:* Lieske, Lehnert, Riedel, Seipp, Arlt, Pietsch, Meder, Richter, Ehmke, Meyer, Simons, Hilgers, Meier. *Method:* Die automatische Bearbeitung sprachlicher Daten steht im Zusammenhang mit der dort versuchten automatischen Erstellung von Lehrprogrammen aus Basaltexten, die den Lehrstoff in konzentrierter Form enthalten. Dabei tritt Sprachbearbeitung auf: 1) bei der Auswahl von Verknüpfern zu Lehrquanten in Lehrprogrammen—es wird auf Ubereinstimmung in möglichst vielen nichttrivialen Wörtern geprüft, 2) bei der Auswahl von Lehrquanten, die gegenüber dem bisherigen Text möglichst wenige neue Nichttrivialwörter bringen, 3) beim Aufbau von Lehrquanten und Verknüpfern von Satzformen mit Leerstellen und Nichttrivialwörten aus dem Basaltext.

L318 . Testing of Frequency Lists

Principal investigators: Leopold K. Engels, Professor; Romain E. Eeckhout, Graduate Research Assistant; Institute of Applied Linguistics, Louvain U. (Dutch section), Vesaliusstraat 2, Leuven, Belgium.

Scope: 3000 most frequent words of the English frequency list (M. West) compared with 10 samples of 1000 words; 3000 most frequent words of the Dutch frequency list (G. Vannes) compared with 10 samples of 1000 words. *Method:* Punching frequency lists and lemmatized texts; transferring to tapes; comparing and printing results.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV. Program is available.

References: L.K. Engels, "The Fallacy of Word Counts," *IRAL* 6, 3 (1968), 213-231; R. Eeckhout, *Woordfrequentie in het Geschreven Nederlands* (Word Frequency in Written Dutch) (Antwerp: Ed. Plantyn, 1968). [*CHum*, November 1969]

L319. Linguistic Analysis of Modern German Texts

Principal investigators: Dr. Ulrich Engel, Assistant Director; Paul J. Wolfangel, M.A., Institut für Deutsche Sprache, 68 Mannheim 1, L 11, 18, Germany.

Scope: Modern German text library on magnetic tape; alphabetical and frequential index registers of the texts. Linguistic programming system with emphasis on compatibility (in preparation); grammatically coded parallel texts (multi-purpose grammar code).

Type of computer: Siemens 4004/35. *Size of storage:* 64k bytes. *Language:* Assembler, ALGOL, FORTRAN IV. *Disks or tapes:* 2 4004/564 disks; 4/Siemens 4004/43-1-9 tapes. *Special equipment:* Paper tape reader/punch, chain printer with capital and small letters. Program is available on exchange basis.

References: Forschungsberichte des Instituts für deutsche Sprache 2 (1969); Bayer and Kurbel, "Maschinelle Textverarbeitung im Institut für deutsche Sprache, Mannheim" (to appear in *Institut für deutsche Sprache, Forschungsberichte*).

L320. A Concordance to the Plays of W.B. Yeats

Principal investigator: Eric Domville, Assistant Professor, Dept. of English, New College, U. of Toronto, Toronto 5, Canada. *Associate:* James Painter.

Scope: A concordance based on the Variorum Edition of Yeats' Plays edited by Russell K. Alspach. It includes all the main text and all variants considered significant. *Method:* Keypunch; transfer to tape; produce concordance and relevant statistics.

Type of computer: IBM 360/65. *Size of storage:* 512k. *Language:* PL/I F. *Disks or tapes:* 1 IBM 2314 disk; 6 IBM 2704 IV tapes. Program is available.

Status: Concordance has been published. [*CHum*, November 1969]

L321. Grammatical and Stylistic Analysis of Selected French Texts (Old and Modern)

Principal investigator: Guy de Poerck, Titular Professor, Dept. of French Linguistics, Arts School, State U. of Ghent, Blandijnberg 2, 9000 Ghent, Belgium. *Associates:* H. Cuyvers, R. Van Deyck, R. Zwaenepoel, A. Boone, D. Willems.

Objective: Drawing up of basic material enabling an exhaustive description of French (old and modern) from syntactic, semantic, and statistical points of view. *Method:* 1) Establishment of continuously improved concordances (with morphological regrouping of forms, frequency of occurrence and context). Perforated texts: Old French—*Chanson de Roland, Charroi de Nimes, Couronnement de Louis, Prise d'Orange,* François Villon; Modern French—J.M.G. Le Clezio, *L'Extase Matérielle;* J. Rostand, *L'Etat présent du transformisme,* etc.; Général De Gaulle, *Allocutions radiodiffusées;* Texts written up by W.Z. Zwanenburg from recorded conversations, as published in *Recherches sur la prosodie française.* 2) From these basic listings, syntactical analysis by more or less complete formalization of the text (replacement of words

by their grammatical symbol). 3) Automatic analysis of the 35,000 examples of the *Essai de grammaire de la langue française* by J. Damourette and E. Pichon. As each example is accompanied by biography, bibliography, and literary data, this material will permit study of specific periods or styles.

Type of computer: IBM 360/30. *Size of storage:* 64k. *Language:* Assembler and PL/I D. *Disks or tapes:* 2311 disks; 3420, 9-track tapes. *Special equipment:* Paper tape, plotter, special Typebar (63 characters). Program is available.

Status: Continuing.

References: "Traitement automatique de textes littéraires en ancien français. Une expérience: *Le Charroi de Nîmes,*" *Proceedings of the Colloquium L'Utilisation des ordinateurs et la recherche en sciences humaines, Bruxelles, 25-27 février, 1971* (Bruxelles, 1971), pp. 103-111; A. Boone, H. Cuyvers-Lippens, G. De Poerck, R. Van Deyck, D. Willems, and R. Zwaenepoel, "Projets et réalisations en traitement automatique dans le domaine du français," *Mélanges Paul Imbs* (in press); H. Cuyvers and R. Zwaenepoel, "De automatische verwerking van teksten: Concordanties," *Documentatiecentrum, Sectie Romaanse,* Rijksuniversiteit te Gent, Belgium (in press); *Textes et Traitement Automatique* (Saint-Aquilin-de-Pacy [Eure]: Librairie-Editions Ph. Mallier): Vol.l, G. De Poerck, R. Van Deyck, and R. Zwaenepoel, *Le Charroi de Nîmes,* 1970; Vol.2, R. Van Deyck, R. Zwaenepoel, and H. Cuyvers, *François Villon* (in press).

L322. Syntax and Stylistic Values

Principal investigator: Juliette Daigneault, graduate student, Dept. of Linguistics and Modern Languages, U. of Montreal, C.P.6128, Montreal 3, PQ, Canada.

Objective: A list of the syntactic features that characterize the stylistic variations in Colette's *Le Blé en herbe. Method:* Abstract symbolic representation of the sentences in the corpus; description of patterns of syntactic structure which constitute stylistic variations; computerized matching of these patterns with the symbolic representation, resulting in a quantification of the relationship between syntax and style.

Type of computer: CDC 3400. *Size of storage:* 32k. *Language:* FORTRAN, SLIP. *Disks or tapes:* 12 602 tapes.

References: Mechanical Translation Project Quarterly Report (Montreal: University of Montreal, Faculty of Letters and National Research Council, July 31, 1966); Ibid., October 31, 1966; D.E.S. Thesis, *Valeurs stylistique et français contemporain,* University of Montreal, 1964. Obtainable from the author. [*CHum,* November 1969]

L323. Stylistic Analysis of the *Obra Poética* of Fernando Pessoa

Principal investigator: Maria Tereza Camargo Biderman, Director of the Chair of Romance Philology, Dept. of Letters, Faculdade de Filosofia, Ciéncias e Letras de Marília, CP 420 Marília, Estade de Sao Paulo, Brazil.

Objective: Elaborate, using the linguistic corpus of Fernando Pessoa's poetic works, the first model of vocabulary structure and distribution in the Portuguese language; analyze the structural unity and diversity of the four different heteronyms—Alberto Caeiro, Ricardo Reis, Alvaro de Campos, and Fernando Pessoa—poets whom the author invented as fragments of his own personality. *Method:* The poet's texts were rewritten, eliminating all the inflections and utilizing the standard lexical form. On IBM cards each word was preceded by a code which identified its grammatical category. The final results were divided into two parts: heteronyms and grammatical categories.

Type of computer: IBM 5081. *Size of storage:* 10,000 words. Program is available. [*CHum,* November 1969]

L324. Analyse linguistique quantitative de la Bible hébraïque

Principal investigators: Gérard E. Weil, Professeur de Langues Sémitiques Comparées; R. de Possel, Lecture optique, Professeur d'Analyse Mathématique à la Faculté des Sciences de Paris; Dept. of Biblique et Massorétique, Institut de Recherche et d'Histoire des Textes et Centre de Recherche et d'Application Linguistique, c/o Faculté des Lettres, et des Sciences Humaines, F-54 Nancy-Désilles, B.P. 33.97, France. [*CHum,* November 1970] For current status report write to Professor Weil.

L325. Linguistics and Automatic Language Processing

Principal investigator: Dr. Alfred Hoppe, Forschungsleiter, LIMAS, 53 Bonn, Colmarstrasse 24, West Germany. *Associates:* Dagmar Blum, Gerhard Engelien, Monika Kolvenbach, Werner Kummer, Günther Schweisthal, Burkard Schöder, Bernard Spillner.

Objective: Development of a linguistic theory ("kommunikative Grammatik") based on a semantic metalingue and a set of transformations simulating the encoding and decoding of sentences and texts. Application of the theory to machine translation, language teaching, information retrieval, etc. *Method:* Working out a set of algorithms which can be used both for analysis and syntheses ("reversible") connecting terms of the metalingue with surface forms, and connecting them to a previously laid out integrated system.

Type of computer: 1)IBM 7090, 2) IBM 360/40. *Language:* 1) FORTRAN II, 2) Basic FORTRAN. *Disks or tapes:* 1) 7-channel, 2) 9-channel tapes. [*CHum,* November 1969]

L326. Computer-Based Readability Formulas

Principal investigators: Ian E. Fraser, Master of Education, student; A.W. Anderson, Research Officer in University Education; Dept. of Education, U. of Western Australia, Nedlands, Western Australia 6009.

Scope: Computer programs were written for the Lorge and the Flesch readability formulas. The greater part of this study is concerned with solving the problems in programming which arose because of the inconsistencies within the English language. For this reason, the study is linguistic and logical rather than statistical. *Method:* The procedure followed was to write programs for both formulas, validate them on a piece of test data, and then check them for accuracy with twenty-five samples which were believed to be representative of six broad areas of written material. The level of accuracy set was a computer result which differed by no more than one percent from a correct manual result. The program achieved results which were 100 percent accurate. It is estimated that the programs save between 80 and 90 percent of the time taken to evaluate manually.

Type of computer: PDP-6. *Size of storage:* 32k. *Language:* FORTRAN IV. Program is available at Computing Centre, University of Western Australia. [*CHum,* November 1969]

L327. A Concordance to the Songs of the Early Tudor Court

Principal investigator: Michael J. Preston, Coordinator, Center for Computer Research in the Humanities, Dept. of English, U. of Colorado, Boulder, CO 80302. *Associates:* Constance Wright, J.D. Schuchter, Samuel S. Coleman.

Objective: To produce a concordance, rhyme index, indices of first lines, ranking frequency list, and reverse index.

Type of computer: CDC 6400. *Size of storage:* 64k. *Language:* FORTRAN IV, SNOBOL, COMPASS.

Status: Concordance is completed. Some programs are available.

References: Compendia 4, gen. ed. R.A. Wisbey (Leeds, Eng.: W.S. Maney, 1972).

L328. A Computer-Assisted Study of the British Folk Play

Principal investigator: Michael J. Preston, Coordinator, Center for Computer Research in the Humanities, U. of Colorado, Boulder, CO 80302. *Associates:* Donald C. Baker, Constance Wright, J.D. Shuchter, Samuel S. Coleman.

Objective: To demonstrate the relationships among some 156 complete texts and 38 fragments, and among the characters of nearly 400 texts. *Method:* Produce a geographical index of characters, and both KWIC and "standard" concordances to the texts.

Type of computer: CDC 6400. *Size of storage:* 64k. *Language:* FORTRAN IV, SNOBOL4. *Disks or tapes:* 1 6603 disk drive; 4 607 7-track tapes. Program is available.

Status: Machine work complete; study in progress.

References: Microfilm copies of concordances are stored in the archives of the Folklore Society, London; the archives of the Sheffield Survey of Language and Folklore, University of Sheffield; and Norlin Library, University of Colorado.

L329. A Concordance to the Middle English Shorter Poem

Principal investigator: Michael J. Preston, Coordinator, Center for Computer Research In the Humanities, U. of Colorado, Boulder, CO 80302. *Associates:* Constance Wright, Donald C. Baker, Samuel S. Coleman.

Objective: To produce a complete concordance, indices of first lines, ranking frequency list, and reverse index.

Type of computer: CDC 6400. *Size of storage:* 64k. *Language:* FORTRAN IV, SNOBOL4, COMPASS. *Disks or tapes:* 1 6603 disk drive: 4 607 7-track tapes. *Special equipment:* The interface from our Librascope (L3055) to an IBM Selectric 731. Some programs are available.

L330. Semantic Language Processing

Principal investigators: Rodger Knaus, Research Assistant, Sven Ohman, Professor, Dept. of Phonetics, U. of Uppsala, Research Group in Psycho-Linguistics, Ovre Slottsgatan 12, 752 35 Uppsala, Sweden.

Objective: To investigate the effectiveness of a semantically based parsing of English in resolving ambiguity, modifying a stored data base, and reducing natural-language information to a form suitable for a resolution principle theorem-proving program. The basic syntactical concepts used are drawn from Fillmore's case grammar, the data base is a property list structure, and the logic is a resolution-principle logic including formulas of all finite orders. Programming has just begun. A description of the logic is available. English is thought of as having a LISP-like structure, in the sense that many relational words will be used to call LISP subroutines to parse the sentence, and the two languages have a similar, recursive structure. *Method:* Trial and error development of a suitable program. The immediate object is to allow the computer to process sentences with a very small but ambiguous vocabulary.

Type of computer: CDC 3600. *Language:* LISP 1.5. Program is in initial stage. [*CHum,* May 1970]

L331. Folk, Leode, Man, Nation, People, Theode

Principal investigator: Göran Kjellmer, Dept. of English, U. of Göteborg, Lundgrensgatan 7, S412 56 Göteborg, Sweden.

Objective: To determine the extent to which distributional criteria are valid in a semantic analysis of a Middle English corpus. *Method:* Correlation of distributional and semantic data.

Type of computer: IBM 360/65. *Size of storage:* 256k. *Language:* COBOL F. *Disks or tapes:* 2311 disk; 2401/II. Program is available.

Status: Published in 1971.

References: Context and Meaning (Stockholm: Almqvist & Wiksell, 1971).

L332. The Syntactic Study of the Finnish Dialects

Principal investigator: Osmo Ikola, Professor, Dept. of Finnish Language, U. of Turku, Henrikinkatu 3, 20500 Turku 50, Finland. *Associates:* Jussi Kallio, Olli Järvikoski, Jorma Rakkolainen, Marja-Berit Malmberg, Yrjö Karjalainen.

Objective: To collect syntactic material for constructing the syntactic archives of the Finnish dialects and finally to compile the descriptive syntax of the Finnish dialects. *Method:* Quantitative, statistical, and functional analysis of the syntax of Finnish dialects.

Type of computer: 1) UNIVAC 1108, 2) IBM 1130. *Size of storage:* 1) 3 x 64k words, 2) 16k words. *Language:* 1) and 2) FORTRAN IV. *Disks or tapes:* 1) 12 tapes, 2) 2 IBM 2315 disks. Program is available.

References: Osmo Ikola, "Vorbereitungen zur maschinellen syntaktischen Analyse der finnischen Mundarten," *Abhandlungen der Akademie der Wissenschaften in Göttingen,* Phil.-hist. Klasse, Dritte Folge Nr. 76, 1970, pp.107-113; Osmo Ikola, *Mechanical Treatment of Syntactic Material of Finnish Dialects,* Teesid I, Tallinn, 1970, p. 46; Jussi Kallio, "Suomen kansankielen lauseopillisten ainesten keruu ja arkistointi," *Seulaset* 1 (1969), 6-7; *Kertomus tutkimussopimuksen perusteella suoritetusta suomen kansankielen muoto–ja lauseopillisen aineiston keruu–ja arkistointityöstä,* 1.6.1967-31.12.1971, Helsinki-Turku, 1972; Y.A. Karjalainen and M. I. Nurminen, "Report of an Information Retrieval System for Linguistic Studies," *NordDATA-72 Konferensföredrag* (Helsinki, 1972), 3, 1998-2010.

L333. Investigations for Explanation of Spectral Variations of Sounds

Principal investigator: Antti Iivonen, Assistant Professor, Dept. of Phonetics, U. of Oulu, Kasarmintie 8, 90100 Oulu 10, Finland.

Objective: Explanation of the spectral variations of sounds by means of free acoustic variation, coarticulation, speaker identity, stress, etc. *Method:* Acoustic analysis by means of the Sound Sona-Graph model 6061 B (frequency-intensity displays = sections); digital representation of sound spectra; storing the digital spectra by a computer; calculations and treatment on basis of a special program; graphical representation of statistics by means of CalComp digital plotter.

Type of computer: Elliott 803. *Special equipment:* Elliott CalComp digital plotter.

References: "Automatic Recognition of Speech Sounds by a Digital Computer," International Conference on Computational Linguistics (1969), Sanga-Säby, preprint No. 33; "Some Implications and Applications of Using Digital Sound Spectra," 7th International Congress of Phonetic Sciences 1972, Montreal.

L334. A Graphemic-Phonemic and Morphemic Study of the Finnish Language Used in J.P. Finno's *Rucous Kiria,* Printed in 1583.

Principal investigator: Esko Koivusalo, mag. phil., Lexicography Foundation (Sanakirjasäästiö), Institute of the Finnish Language, U. Helsinki, Fabianinkatu 33, Helsinki 17, Finland.

Objective: To try to find out which features in the graphemics and the morphemics in J.P. Finno's language deviate from the corresponding features in the language of other sixteenth-century writers. An attempt will also be made to determine whether certain anonymous literary works of the sixteenth century are written by Finno. *Method:* Counting the frequency of the letters and their various combinations and defining their phonographemic relation; counting the frequencies of the grapheme sequences corresponding to the allomorphs. The differences between the books and the authors will be studied using the method of variance analysis.

Type of computer: 1) IBM 360/50, 2) IBM 1620. *Size of storage:* 1) 256k bytes, 2) 49k. *Language:* 1) FORTRAN IV, 2) FORTRAN II D. *Disks or tapes:* 1) 9 2314 disks and 8 2401/5 tapes; 2) 2 IBM 1311 disks. [*CHum,* May 1970]

L335. A Word-Frequency Study of French Novels

Principal investigator: Gunnel Engwall, Fil. mag., Dept. of Romance Languages, Stockholm U., Fack, 104 05 Stockholm, Sweden.

Objective: To produce frequency lists on words in written French. The material consists of some 25 novels published during the period 1962-1968. The results will be analyzed with reference to earlier studies. This includes comparisons between different texts in the material. *Method:* The input consists of punched cards, which represent the entire material. KWIC lists and frequency counts are derived by means of data processing. These lists are used for the subsequent analyses of homographs and lemmas, followed by new sortings and frequency counts.

Type of computer: 1) IBM 1401, 2) IBM 360/75. *Size of storage:* 1) 16k, 2) 208k. *Language:* 1) AUTOCODER, 2) FORTRAN. *Disks or tapes:* 1) and 2) 2 tapes. Program is available.

References: "Etude sur la fréquence des mots dans quelques romans français (1962-1968)," paper presented at the XIIIe Congrès international de linguistique et philologie romanes, Québec, 29 août-5 septembre, 1971, forthcoming in proceedings; "A Concordance Program for Linguistic and Literary Research," *Computer Center Report No.8* (Stockholm: Royal College of Forestry, 1972).

L336. Phonological Statistics in Swedish

Principal investigator: Claes-Christian Elert, Professor, Dept. of Phonetics, U. of Umea, S901 87 Umea, Sweden. *Associate:* Siv SJögren.

Scope: 1) Phoneme frequency in Swedish. 2) Phoneme transition probability in Swedish. 3) List of all Swedish words in phonemic transcription in alphabetical and reverse order. 4) Conversion of Swedish orthography to phonemic transcription. *Method:* The words in the frequency lists made up by the Institute for Computational Linguistics, University of Göteborg (led by S. Allén), are transferred to phonemic transcription by a special program using the magnetic tape version of the official Swedish dictionary. Phoneme frequencies are found by multiplication of the phoneme occurrences in the words by the frequency of each word.

Type of computer: 1) DATASAAB D21, 2) CDC 3600, 3) IBM 7090. *Size of storage:* 1) 24k. *Language:* 1) ALGOL-GENIUS, 3) FORTRAN IV. *Disks or tapes:* 1) 6 tapes. Program for the orthography-phonemic writing conversion is available.

References: S. Allén, "Datamaskinell undersökning av tidningsprosa," *Sprakvard* 3 (1966), 3-7; H. Karlgren, "Svenskt ordforrad pa magnetband," *Sprakvard* 3 (1966), 8-11.

L337. Concordanze Billiane

Principal investigator: Federico Albano Leoni, Lecturer, Dept. of Romance Languages, U. of Göteborg, Lundgrensgatan 7, 412 56 Göteborg, Sweden.

Objective: A complete concordance to the 2279 sonnets (in Roman dialect) of G.G. Belli (1791-1863). Also a frequency list, a reverse order list, and a rhyme list. *Method:* The work is carried out in cooperation with the projects led by Sture Allén (Göteborg) and Jan Thavenius (Lund).

Type of computer: DATASAAB D21. *Size of storage:* 24k.*Language:* ALGOL-GENIUS. *Disks or tapes:* 6 tapes. Program is available.

References: To appear in *Romanica Gothoburgensia.* [*CHum,* May 1970]

L338. Vergleichende Untersuchungen zum öffentlichen Sprachgebrauch, insbesondere an Zeitungstexten aus der Bundesrepublik Deutschland und der Deutschen Demokratischen Republik

Principal investigators: Professor Dr. Hugo Moser, Ordentl. Professor F. Germanistik; Dr. Manfred W. Hellmann, Wiss. Angestellter; Institut für deutsche Sprache, Forschungsstelle Bonn, Adenauerallee 96, 5300 Bonn, West Germany. *Associates:* Günter D. Schmidt, Barbara Marzahn, Michael Kinne.

Scope: Untersuchung der Differenzierung im öffentlichen Sprachgebrauch in der BRD und DDR; Feststellung ihrer Schwerpunkte und Trends. Entwicklung eines vergleichenden Wörterverzeichnisses der ost- und westdeutschen Zeitungssprache unter Angabe von Frequenzen und typischen Verwendungsweisen (geplant bis 1976). *Method:* Ermittlung und Dokumentation von repräsentativen Querschnitten aus je 3 Jahrgängen der Zeitungen *Die Welt* und *Neues Deutschland* und Vergleichstexten; Klassifikation der Texteinheiten. Anfertigung alphabetischer und frequentieller Wortregister. Weiterhin geplant: Ermittlung textsortenspezifischer Vokabularien. Teilmaschinelle Rückführung der relevanten Wortformen auf die Grundformen. Teilmaschinelle Ermittlung typischer Verwendungsweisen aus den Kontexten.

Type of computer: IBM 370/165. *Size of storage:* 128k. *Language:* FORTRAN IV. *Disks or tapes:* Min. 3 disks, 12 IBM 9-channel tapes. Program is available, Teilweise 1972, Rest bis 1975.

References: Manfred W. Hellmann, "Zur Dokumentation und maschinellen Bearbeitung von Zeitungstexten in der Forschungsstelle Bonn," *Forschungsberichte des Instituts für deutsche Sprache,* Nr. 2, Mannheim, Dez. 1968, S. 39-125.

L339. German Subjunctive

Principal investigator: Per Baerentzen, cand. mag., Institute of Germanic Philology, U. of Aarhus, Dk-8000, Aarhus C, Denmark.

Objective: To examine the modern German subjunctive 1) as to the verbal forms used, and 2) as to the function and the semantic values in a context. *Method:* The computer picks out the different forms and sentences from a corpus of modern German texts.

Type of computer: IBM 7090. *Language:* FORTRAN IV. Program is available. [*CHum,* May 1970]

L340. Human Speech Production: Encoding Mechanisms and Development

Principal investigator: Sven E.G. Ohman, Professor, Inst. Fonetiska, Uppsala U., S:t Johannesgatan 22, 752 21, Uppsala, Sweden.

Objective: To extend existing quantitative models of speech production so as to classify some of the more complex patterns of organization encountered in natural connected speech of normal adults and children and to approach the problem of how the articulatory motor-skills of the grownup are gradually approximated by the normal child. To this end close comparisons will be made between systematically collected sets of empirical data and the outputs generated by computer simulated models of the speech production processes that are assumed to be fundamental to the observed facts. Attempts will be made to develop acoustic metrizations of developmental factors in speech. Using accurately controllable analysis and synthesis methods, we will also make correlative studies of productive and perceptual performance of children of various ages. [*CHum,* May 1970]

L341. The Vocabulary of *Hufvudstadsbladet* (daily paper, Helsinki) 1967

Principal investigators: Carl-Eric Thors, Professor, Institute of Scandinavian Linguistics, U. of Helsinki, Porthania, Helsinki, Finland; Märta Malmström, Lecturer, Dept. of Philosophy, Institute of Nordic Philology, Pakkahuoneenkatu 12B, U. of Oulu, Oulu, Finland.

Objective: The project is closely affiliated with the Gothenburg project, led by Sture Allén, of investigating the language of five Swedish dailies; its purpose is similar. It should, however, be noted that one of the aims is to find out differences in vocabulary between *Hufvudstadsbladet* (which is published in Swedish) and the five dailies dealt with by the Gothenburg project and to detect how soon novelties in vocabulary spread to Finland. *Method:* The material is selected according to the same principles as have been established for the Gothenburg project. Corpus: 200,000 running words.

Type of computer: DATASAAB D21. *Size of storage:* 24k. *Language:* ALGOL-GENIUS. *Disks or tapes:* 6 tapes. (See L6 for related information.) [*CHum,* May 1970]

L342. SQAP: Swedish Question-Answering Project

Principal investigator: Jacob Palme, Research Engineer, Research Institute of National Defense, Inst. 820, S-104 50, Stockholm, Sweden. *Associates:* Kalle Mäkilä, Erik Sandewall, Siv Sjögren.

Objective: Programming of a question-answering system which can take facts and questions in a simplified natural language (English, Swedish, or Esperanto) and answer the questions from the facts and from deductions based on the facts. Compared to other similar projects, our system is intended to take a wider range of facts, not limited to a certain

subject area, and will use a special method for internal data storage and for deduction. Palme and Sjögren are working on the input text translation, Sandewall and Mäkilä on the deduction procedures.

Type of computer: IBM 360/75. *Size of storage:* 512k bytes. *Language:* PL 360. *Disks or tapes:* 8 2314 disks. *Special equipment:* 2260 display will be used in the future. Parties interested in using the program should contact the principal investigator.

References: "Making Computers Understand Natural Language," *Artificial Intelligence and Heuristic Programming,* ed. N.V. Findler and and B. Meltzer (Edinburgh: Edinburgh University Press, 1973). Mimeographed reports available from the principal investigator.

L343. The Reverse Dictionary of the Finnish Language

Principal investigator: Tuomo Tuomi, Lexicography Foundation (Sanakirjasäätiö), Institute of the Finnish Language, U. of Helsinki, Fabianinkatu 33, 00170 Helsinki 17, Finland.

Objective: Prepare the reverse alphabetized word list of the Finnish language. The first part of it, published in January 1973, consists of the vocabulary of the standard Finnish language. Included in the dictionary are introduction and comments on the inflectional schemes in English. The second part of the vocabulary will consist of the Finnish dialects, and it will be completed in March 1973. Some studies are already under preparation on the basis of the material transferred on the tapes. In the second phase, the purpose is to prepare the statistical analysis of some inflectional, derivational, and phonotactic features. The third phase consists of studies on the regional usages of the dialect words. *Method:* Methods used in the first and second parts are conventional. Methods for the third phase are under preparation.

*Type of computer:*1) IBM 360/50, 2) Burroughs B6700. *Size of storage:* 256k bytes and 357k bytes. *Language:* ALGOL and FORTRAN. *Disks or tapes:* 9 2314 disks and 8 2401/5 tapes.

Status: Phase 1 completed 1973.

References: Reverse Dictionary of Modern Standard Finnish (Helsinki: The Finnish Literature Society, 1973).

L344. Observations on Change in Language

Principal investigator: Gun Widmark, Docent, Dept. of Scandinavian Languages, Thunbergsvägen 3, Uppsala, Sweden. *Associate:* Jan Trost.

Scope: The speech community studied is the population of Uppsala, more precisely those who were born in the city or in the neighboring county (Uppland). A sample of about 250 informants has been interviewed. The investigation concerns mainly different types of vowel quantity variation, some of which must be innovations in Central Swedish pronunciation. The collected data are correlated both with nonlinguistic factors (such as sex, age, social class, education) and with other words and constructions that are clearly involved in a process of linguistic change.

Type of computer: CDC 3600. *Size of storage:* 32k. *Language:* FORTRAN, almost all high-level languages. *Disks or tapes:* 2 603, 4 604 tapes. Program is available, ROBOT. [*CHum,* May 1970]

L345. Mechanical Processing of Natural Languages

Principal investigators: Toshiyuki Sakai, Dr. Eng'g; Shigeharu Sugita, Dr. Eng'g, Dept. of Information Science, Kyoto U., Yoshida Sakyo-ku, Kyoto, Japan.

Scope: Processing of natural languages by digital computer is studied in order to obtain an effective system in which men and computers take a part of mind-like works according to their abilities. As a first step, mechanical translation program between English and Japanese has been developed. The result of translation can be spoken by synthesized voice. Though this program treats only syntactic structures, the technique can be applied to many other problems which relate to natural languages, such as speech recognition or character recognition by computer. *Method:* A medium-size digital computer, to which several special input/output devices are connected, is used. As a model of English and Japanese, phrase structure grammar and transformational grammar are adopted. A study of semantic information as well as a syntactic one is under way.

Type of computer: NEAC 2200/200. *Size of storage:* Core: 32kc; disk: 9.2 mc. *Language:* Assembler. *Disks or tapes:* 5 7-track tapes. *Special equipment:* Flying spot scanner, vidicon scanner, speech synthesizer, graf-pen, storage type CRT.

References: T. Sakai, "Adaptive System of Pattern Recognition," presented at the International Conference on Methodologies of Pattern Recognition, January 1968; S. Sugita, "A Study on Mechanical Translation from English into Japanese," (Unpub. Ph.D. thesis, Kyoto University, January 1968).

L346. A Formal Analysis of Children's Spoken Language

Principal investigator: Hans Vejleskov, Associate Professor, Dept. of Developmental Psychology, Royal Danish School for Educational Studies, Emdrupvej 101, Copenhagen, Denmark. *Associate:* Elisabeth Hansen.

Objective: Analyzing children's spoken utterances according to a grammatical system constructed for description of written Danish language. *Method:* The children's utterances in five different situations were recorded, typewritten, and grammatically analyzed. Since the system in question is a position system with a finite number of possibilities for the filling in of the various positions, the sentence can be converted to a set of numbers.

Type of computer: IBM 7094. *Language:* ALGOL. [*CHum,* May 1970]

L347. Syntactic Analysis with Minimal Information Added

Principal investigator: Gustav Leunbach, Research Director, Dept. of Statistics, Danish Institute for Educational Research, Hermodsgade 28, DK 2200, Copenhagen N, Denmark.

Scope: To what degree is a text amenable to automated syntactic analysis with only fundamental grammatical information added to the coded text? How much is the percentage of the text which can be analyzed increased by adding syntactic and semantic information on separate words? *Method:* A complete novel in Danish has been sorted into word forms, giving a total of about 5000. These word forms have been analyzed for all occurrences of two or more possible flexions of the same root; about 50 percent of the word forms have been marked as such flexions (including all random coincidences). Each flexion carries information about its syntactic function (e.g., participle, i.e., adjective from verbal root). For a number of highly frequent words, such as pronouns, prepositions, and auxiliary verbs (mainly irregularly flexed), syntactic information is added directly. Then a search may proceed for sentence structures, utilizing points, commas, etc., these highly frequent words, and all word forms labeled as flexed forms.

Type of computer: UNIVAC 1108. *Size of storage:* 192k. *Language:* ALGOL.

References: Preprint CA 3.3, International Conference on Computational Linguistics, Sanga-Säby, September 1969.

L348. Philological Analysis of Ancient Texts

Principal investigator: Donald C. Swanson, Professor, Dept. of Classics, U. of Minnesota, Minneapolis, MN 55455. *Associate:* Dorothy M. Swanson.

Objective: A continuous program of lexical and syntactic analysis and tabulation applied chiefly to certain Greek and Latin texts (but occasionally to other linguistic spheres). Present topics are 1) Caesarian vocabulary, 2) Vorro's "De Lingua Latina: An Index Verborum," 3) clause analysis of some of the text of Cicero's "Letter to Atticus." *Method:* The application of Bloomfieldian (empirical-structural) linguistics to vocabulary and clause-syntax; rigorous analysis accompanied by computer precision. Statistics (data quantification) form part of the results aimed at.

Type of computer: CDC 6600.

References: "Names in Roman Verse" (Madison, 1968); "Vocabulary of Egeria," *Zeitschrift für romanische Philologie* 82 (1966).

L349. Automatic Syntactic Analysis of Modern German

Principal investigator: Hans Eggers, Professor, Germanistisches Institut, U. des Saarlandes, D 6600 Saarbrücken, Germany. *Associates:* R. Dietrich, W. Klein, R. Rath, A. Rothkegel, H.J. Weber, H.H. Zimmermann.

Scope: Automatic analysis of written German sentences on the basis of a grammatical dictionary and an approximate parsing system. The program functions on the following levels: 1) reduction and solution of syntactic homography; 2) recognition of nominal and verbal groups; 3) delimitation and inventorying of (even complex–embedded or discontinuous–) clauses; and 4) reduction of case ambiguity. In principle, all levels are completed. We plan to refine and extend the system, automatically lemmatize German (scientific) texts by using the parsing system, and transfer the program on a CDC 3300 computer (FORTRAN).

Type of computer: X1 Electrologica. *Size of storage:* 32k (1 word = 27 bits). *Language:* Assembly (Kieler Kode). *Disks or tapes:* 3 Sandwich tape, Scotch 789. Program is available, but only on an X1 computer.

References: H. Eggers und Mitarbeiter, *Elektronische Syntaxanalyse: Ein Bericht* (Tübingen: Niemeyer, 1969); Harald H. Zimmermann, *Das Lexikon in der maschinellen Sprachanalyse* (Frankfurt: Athenäum, 1972).

Status: Completed in 1970. Project has been merged with project L30 (Automatic Lemmatization).

L350. A Concordance to the Works of Franz Kafka

Principal investigator: Walter Speidel, Professor, Dept. of Germanic Languages and Literatures, Brigham Young U., 324 McKay Building, Provo, UT 84601. *Associates:* Lance S. Smith, R. Brent Thompson, Mark E. Strong, Erich C. Decker.

Objective: To create a research tool that will allow more serious and definite studies of Kafka's language and style (most importantly maybe tracing of motifs, images, and symbols) by concordancing individual, complete works of Franz Kafka; creating keyword-in-extensive-context format with location-identification, and complete alphabetical indices (root form) with total frequency for nouns, verbs, adjectives, adverbs, proper nouns (names), foreign words, numerals (numbers), and creating a general index of the rest of Kafka's vocabulary; forming an alphabetical list of all deletions, and a list of all variants or additions with location-identification. Deletions are identified by an asterisk when they appear as keywords in the concordance and are not counted in the frequency, rank frequency (descending order), percentages and other statistical information, and reverse index. Umlauts, the double *s* and upper-case letters are coded for proper computer typesetting.

Type of computer: IBM 360/50. *Size of storage:* 1384k. *Language:* 360 Assembler F. *Disks or tapes:* IBM 2314 disks, *Special equipment:* 2741 remote terminal.

L351. A Program Concordancing Prose, Poetry, and Bibliographic References

Principal investigator: Janet Hillier Caulkins, Assistant Professor, Dept. of French and Italian, U. of Wisconsin, Madison, WI 53706.

Objective: To produce and publish individual concordances to 1) Chrétien de Troyes, *Cliegés*; 2) Chrétien de Troyes, *Erec et Enide;* 3) Thomas, *Le roman de Tristan;* 4) Béroul, *Tristan et Iseult;* and 5) Benedeit, *Le voyage de saint Brendan.* The concordances are being employed currently to determine significant variants of meaning in recurring vocabulary. *Method:* The concordances were produced by BIBCON, a concordance program developed at the University of Wisconsin by Professor Richard L. Venezky. Each line was punched, one line per card.

Type of computer: CDC 3600. *Size of storage:* 65k. *Language:* FORTRAN IV. Program is available from Professor Venezky.

References: R.L. Venezky, *Computers and the Humanities* 3, 3 (1969), 129-138; R.L. Venezky, *BIBCON: A 3600 Program for Generating Concordances to Prose, Poetry, and Bibliographic References,* U. of Wisconsin Computer Science Technical Report No. 34, July 1968. [*CHum,* May 1970]

L352. Research on Automatic Syntactic Analysis

Principal investigator: Stephen Braun, Wissenschaftlicher Assistant, Abteilung Mathematik, TH München, 8 München 2, Arcisstrasse 21, West Germany. *Associate:* Gerda Schott.

Objective: Automatic syntactic analysis of Russian scientific texts. *Method:* On the basis of word order and agreement properties, a set of rules—at present only for noun phrases—and a corresponding strategy of application of these rules is devised, programmed, tested, and improved.

Type of computer: Telefunken TR4. *Size of storage:* 30k. *Language:* ALGOL 60 (+ code procedures for string manipulation). Program is available.

References: "Studien zur automatischen syntaktischen Analyse des Russischen," Math. Inst. TH München, Arbeitsbericht No.6910, August 1969. [*CHum,* May 1970]

L353. Transformational French Grammar

Principal investigator: Maurice Gross, Professor, Dept. of Linguistics, U. of Paris 8 and Centre National de la Recherche Scientifique, 2 place Jussieu, U. of Paris 7, 75005, Paris, France. *Associate:* Jean-Paul Boons.

Objective: Construction of a formal transformational grammar of French, and of a syntactic analyzer. *Method:* Transformational analysis of French constructions. Exhaustive study of all parts of speech.

Type of computer: IBM 370/75. *Size of storage:* 512k. *Language:* FORTRAN, PL/I. *Disks or tapes:* 12 disks, 12 tapes. Program is available.

Status: To be completed December 1974.

L354. Mathematical Models in the Study of Language

Principal investigator: Barron Brainerd, Professor, Dept. of Mathematics, U. of Toronto, Toronto, Canada.

Objective: 1) Statistical testing of parameters of literary style; 2) testing a model of glottochronology. *Method:* 1) Gather data to be tested for homogeneity among works of same author, works of different authors, different styles of writing, etc.; 2) test the reliability of parameters in a new model of glottochronology. Program is available.

References: "A Stochastic Process Related to Language Change," *Journal of Applied Probability* 7 (1970), 69-78; "On the Relation between Types and Token in Literary Text," *Journal of Applied Probability* 9 (1972), 507-518; "Article Use as an Indicator of Style among English-Language Authors," *Linguistik und Statistik*, ed. Siegfried Jäger (Vieweg-Verlag, 1972), pp. 11-32.

L355. Book of Filial Piety in Tangut

Principal investigator: Eric Grinstead, Librarian, Scandinavian Institute of Asian Studies, Kejsergade 2, Copenhagen, Denmark.*

Scope: Environment of illegible characters in *Kitaiskaya klassika tangutskom perevode* (Moskva, 1966), pp. (from end) 125-211, ed. E.I. Kychanov and others. [*CHum,* May 1970]

L356. Analysis of Nordic Folk Literature

Principal investigator: Bengt Holbek, Director, Nordic Institute for Folk Literature, Birketinget 6, DK-2300 Copenhagen S., Denmark. *Associate:* Carljohan Rosenquist.

Scope: Development of methods of information retrieval in the field of folk literature—beginning with proverbs, to be continued with anecdotes and jests, riddles, folk tales, legends, and ballads.

Type of computer: IBM 7094. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 16 IBM 729 tapes. [*CHum,* May 1970]

L357. Regional Vocabulary, Syntax, and Usage in English

Principal investigator: Gordon R. Wood, Professor, Dept. of English, Southern Illinois U., Edwardsville, IL 62025.

Objective: To develop frequency counts of all words in 1000-word samples of local Southern (USA) speech and, later, to make comparisons of phrase and sentence structures among the various speakers. The main purpose is to extend the scope of dialectology from the study of selected individual words to the analysis of large units of continuous discourse. *Method:* The original tape recordings, made by using a *Pictorial Linguistic Interview Manual,* were transcribed word for word and punched on data cards. The cards were then processed for type-token ratios and other complex statements of relationships of individual word to total text. The individual frequency lists were merged to form composite subregional lists and a composite master list for the whole area (the lower Tennessee Valley to the Gulf of Mexico). Test runs have been completed on 10,000 words. Work has been started on the larger than single word patterns (e.g., *old timey rail fence*).

Type of computer: IBM 360/40. *Size of storage:* 256k. *Language:* FORTRAN IV. *Disks or tapes:* 5 disks, 4 9-track tapes. For availability of program write to Professor Robert S. Wachal, Department of Linguistics, University of Iowa.

References: Robert S. Wachal and Otfried Spreen, "Some Measures of Lexical Diversity in Aphasic and Normal Language Performance," Technical Report No. 40, ONR Project, University of Iowa, 1971, describes the linguistic variables measured; *Sub-Regional Speech Variations in Vocabulary, Grammar, and Pronunciation* (1967) is available from ERIC (ED 019 263).

L358. Vocabulary Change in Regional English

Principal investigator: Gordon R. Wood, Professor, Dept. of English, Southern Illinois U., Edwardsville, IL 62025.

Objective: To sort and tabulate 1000 sets of vocabulary questionnaires so that the investigator could discover patterns of word distribution within eight states in the southern United States. *Method:* Individual responses were coded according to informant's residence, age, sex, education, and word choice. The coded information, with the response punched on cards, was IBM tabulated by county according to age groups. From these original tabulations, compilations of total occurrences and various percentages of local occurrence were produced. Printouts of the percentile tables became part of a book; maps were hand produced from the original county tabulations; scattergrams were also produced by

hand along with other lists originating in the computer tables. Had time and funds permitted, computer mapping would have been tried.

Type of computer: IBM 1401. *Language:* FORTRAN (but any language that can count and list would have served). *Disks or tapes:* 5 tapes.

Status: Completed.

References: Vocabulary Change, A Study of Variation in Regional Words in Eight of the Southern States (Carbondale & Edwardsville: Southern Illinois University Press, 1971).

L359. Coptic Concordance of Nag Hammadi Gnostic Documents

Principal investigator: Eldon H. Degge, Ph.D. student, Dept. of Religion, Claremont Graduate School, Claremont, CA 91711.

Objective: To provide a concordance in Coptic of the Gnostic documents found at Nag Hammadi, Egypt, in 1945. Concordance form is KWIC. Sorting order conforms to the conventions of Crum's *Coptic Dictionary.* Input is by way of punched cards. A specially devised transliteration scheme is used. Concordance separates Greek loan words and proper names, as well as giving separate entry listings for multiple meanings of the same spelling. All words are listed under basic root as given in Crum rather than in the form of their occurrence. Only the text of the Gospel according to Thomas has been processed to date. *Method:* Program 1 builds a contextual record for each single word and a pseudo-sort key to cause sorting in Crum's order. These records are then sorted with standard sort package program. Once sorted, the records are formatted and printed by a second program which divides the Coptic words, Greek loan words, and proper names into separate sections.

Type of computer: IBM 360/40. *Language:* PL/I. *Disks or tapes:* 1 2311 disk; 2 9-track tapes. [*CHum,* November 1970]

L360. A Concordance to the *Pensées* of Pascal

Principal investigator: Hugh M. Davidson, Professor, Dept. of Romance Languages, Ohio State U., Columbus, OH 43220.

Objective: To produce a concordance to the *Pensées* of Pascal (using the edition with variants by L. Lafuma, in 3 vols., Paris: Editions du Luxembourg, 1951). A printout has been made and is in process of revision. Future publication of the concordance in book form is intended.

Type of computer: 1) IBM 360/75, 2) IBM 360/50, and 3)IBM 7094. *Size of storage:* 1) 1024k, 2) 512k, 3) 32k. *Language:* 1) PL/I, 2) UTILITIES, and 3) SCATRAN, Assembly. *Disks or tapes:* 1) 2314 disk; 1), 2), 3), 8 tapes. *Special equipment:* TN print train (upper/lower case). [*CHum,* November 1970]

L361. A Concordance to the Complete Works of Gil Vicente

Principal investigator: William W. Moseley, Professor, Dept. of Foreign Languages, Colorado State U., Fort Collins, CO 80521.

Objective: Preparation of a concordance to all plays and lesser works of Gil Vicente, with location of word by play abbreviation, page, column and line number, with the word in context, and with a ranking list of frequencies. *Method:* The entire text is punched on cards, one line per card; transfer to magnetic tape follows, with final printout by computer-linked Selectric typewriter.

Type of computer: 1) CDC 6400, 2) Librascope, 3) Athena-Selectric. *Size of storage:* 1) 64k, 2) 32k. *Language:* 1) FORTRAN IV. *Disks or tapes:* 1) 6 7-track tapes. For program contact Center for Computer Research in the Humanities, U. of Colorado, Boulder, CO 80302.

Status: Completed December 1972.

L362. Tentativo di applicazione dell'analisi spettrale alla linguistica

Principal investigators: Pierpaolo Luzzatto Fegiz, Professor; Ida Giovanna Conti Puorger, student, Dept. of Economics and Business Administration, Institute of Statistics, U. of Rome, Via del Castro Laurenziano, 9, Rome, Italy.

Scope: The research is concerned with the problem of finding regularities in linguistic style for Indro Montanelli, Gabriele D'Annunzio by spectral and cross-spectral analysis.

Type of computer: UNIVAC 1108. *Size of storage:* 132k. *Language:* FORTRAN V. *Disks or tapes:* F H 432 N.4-Farstrand disks; 6 Uniservo 6 C tapes.

L363. Analysis of the Aramaic of Onqelos

Principal investigator: Menahem Z. Kaddari, Professor, Dept. of Hebrew and Semitic Languages, Bar-Ilan U., Ramat Gan, Israel. *Associate:* Henya Rappaport.

Scope: Sentence types in the Aramaic of Onqelos are being set up. Constituent elements of the various sentence types will be studied; verbal phrases and noun phrases; correlation between animated/unanimated nouns and syntactic functions will be considered. *Method:* Manual analysis of the poetical parts of Onqelos (Gen. XLIX, Ex. XV, Num. XXIII, Deut. XXXII, XXXIII). Full morphological and syntactic analysis has been carried out.

Type of computer: IBM 360/50. *Size of storage:* 250k.

Status: Semantic analysis is in process.

L364. Automatic Translation: Building a Dictionary (French/English)

Principal investigator: J.P. Vinay, Head, Dept. of Linguistics, U. of Victoria, Victoria, BC, Canada. *Associates:* B. Kallio, M.T. Wilton.

Scope: Investigation on automatic translation from French into English has been going on for three years; this new field (building a dictionary) is meant to define the format of the entries used in several "word banks" in Canada and France, using the "semantic components" technique. We are trying to obtain from information stored in the computer either semantic data or translation units corresponding to a rather limited discipline. If the program works, a wider field will be chosen. *Method:* Investigation deals both with computer programming and semantic analysis. *Future plans:* Parallel with the LEXAUTOM project is the revision of *The Canadian Dictionary/Dictionnaire Canadien* (Research Centre for Bilingual Lexicography). Research is being carried out to employ computerized editing and updating of the dictionary.

Type of computer: IBM 360/44. *Size of storage:* 256k. *Language:* PL/1 F, FORTRAN IV G and H. *Disks or tapes:* 6 2311 disk drives; 4 2401/4 tape-drives.

References: Reports on LECTOR, LEXAUTOM and other M.T. Research, University of Victoria, 1968 (mimeographed); "Codage automatique et dictionnaires," *Mémoires de la Société Royale du Canada* IV, 7 (1969); "Utilisation électronique de la Banque de Mots," *Résumé* (distribué à l'occasion du Colloque sur la linguistique et la traduction, Montréal, septembre 1970); "Utilisation électronique de la Banque de Mots," *META* (Montréal, numéro spécial sur le Colloque de 1970) 16, No.1-2, mars-juin 1971; J.P. Vinay and B. Kallio, "Bilingual Lexicography and the Computer—a report of the first phase of the LEXAUTOM project," University of Victoria, May 1971 (mimeographed).

L365. A Concordance to the Poems of Hafiz

Principal investigator: Alan Jones, Vice-Master, St. Cross College, Oxford, England. *Associate:* Inger Karlsson.

Objective: A full concordance to the *Divan* of Hafiz. A complete line of text will be printed for each reference. Printout in Persian characters.

Type of computer: Atlas. *Special equipment:* SC 4020. Program is available.

Status: Completed.

L366. A Sample Word-Count of Modern Turkish

Principal investigator: Alan Jones, Vice-Master, St. Cross College, Oxford, England. *Associates:* G.L. Lewis, R.C. Reep. *Scope:* Part 1: A word-count of 300,000 words taken from 1968-1969 Turkish newspapers. Part 2: A word-count of 300,000 words taken from 1968-1969 Turkish literary works. Part 3: A word-count of 300,000 words taken from 1973-1974 Turkish newspapers. Part 4: A word-count of 300,000 words taken from 1973-1974 Turkish literary works. Grammatical analyses of the texts are also being made.

Type of computer: ICL 1906A. Program is available.

L367. Research in the Automatic Grammatical Tagging of Words in Simple French Sentences

Principal investigator: Paul Jinot, graduate student, Dept. of French, Brown U., Providence, RI 02912. *Associates:* W.N. Francis, Camille Bauer.

Objective: Intended as a step toward the long-range objective of a complete automatic parsing procedure for French sentences, this project attempts to develop an automatic procedure for tagging the individual words of sample, noncomplex French sentences, in terms of the grammatical categories to which they belong. The emphasis will be on the search for morphological and positional clues, rather than on the use of a dictionary look-up routine. *Method:* About one hundred simple sentences from Camus' *L'Etranger* will serve as the research corpus. They and their components (sorted in various ways) will yield the basic information needed to guide the investigation.

Type of computer: IBM 360/50. *Size of storage:* 512k. *Language:* PL/I F. *Disks or tapes:* 1 2314 disk. [*CHum,* November 1970]

L368. Computer Support of Syntactic Research

Principal investigator: Paul G. Chapin, Assistant Professor, Dept. of Linguistics, U. of California at San Diego, La Jolla, CA 92037. *Associate:* Erin H. Bulman.

Objective: Development of efficient and useful methods for the storage and retrieval of data required by the linguist engaged in syntactic description, particularly by elicitation from informants. Machine-independent algorithms of system COMPANION were fully specified in EULER language but never implemented. Documentation available on request from principal investigator.

Status: In abeyance since summer 1970.

L369. Research in Computational Linguistics

Principal investigator: Paul G. Chapin, Assistant Professor, Dept. of Linguistics, U. of California at San Diego, La Jolla, CA 92037. *Associate:* Walter Olson.

Objective: Implementation of a low-redundancy strategy for automatic transformational syntactic analysis; development of analysis procedures for transformational grammars with output conditions, constraints on variables, and complex symbols at both terminal and nonterminal levels.

Type of computer: CDC 3600. *Size of storage:* 32k. *Language:* LISP 1.5, SNOBOL3. A LISP 1.5 version of the Woods-Kuno CF parser is available.

Status: Terminated. [*CHum,* November 1970]

L370. Ben Jonson, *Volpone*

Principal investigator: Gordon R. Wood, Professor, Dept. of English, Southern Illinois U., Edwardsville, IL 62025.

Objective: The purpose of this stage of the literary investigation is to apply the Ross and Rasche Eyeball programs to selected parts of the card deck of *Volpone* produced at an early stage of a now-abandoned effort to do a computer-based concordance of Jonson's dramatic works. A trial run based on 1000 words from the opening speeches of Volpone and Mosca will determine whether there are quantitative stylistic differences between their ways of discourse. *Method:* The first step in using the Ross and Rasche program will be to examine the ratio between patterns of Volpone's discourse and those of Mosca's. The test sample will be 1000 successive words from their respective conversations in Act I. The basic question: Do personages in this play have measurably distinct styles of speaking?

Type of computer: IBM 360/40. *Size of storage:* 256k. *Language:* FORTRAN IV. *Disks or tapes:* 5 disks; 4 9-track tapes. Program is available. Contact Donald Ross, Jr., Dept. of English, University of Minnesota, Minneapolis, MN.

L371. Development of a Theory of Oral Grammar

Principal investigator: Roger A. Benson, 4314 Valle Vista, San Diego, CA 92103.

Scope: This research builds upon the Chomsky model of linguistic competence: a model consisting of three components—syntactic, semantic, and phonological. An expansion of this model is postulated, consisting of additional components—lexical, cultural, and rhetorical. The lexical component follows the suggestions offered by Humboldt, and exists as a "Vorrath von Wörten" that is accessed randomly, producing data that are processed by a linguistic control center which further processes these data through a cultural component and eventually through a rhetorical component. *Method:* A corpus of data has been taken from orally composed poetry, *Beowulf,* and has been subjected to computer analysis for the purpose of revealing aspects of the poem that indicate oral composition. Use has been made of metrical form and the process of alliteration, as well as other factors.

Type of computer: 1) CDC 160-G, 2) IBM 360/34. *Language:* 1) FORTRAN IV, 2) Assembly. *Disks or tapes:* 1) card input, 2) concordances, 3) various outputs. *Special equipment:* S-C 4020 display. Program is available.

References: "Towards a Theory of Oral Grammar," San Diego State College, 1970; "A Generative Model for the Storage and Retrieval of Information," *ERR* (January 1970), General Dynamics/Convair. [*CHum,* November 1970]

L372. Semantic Componential Analysis of Poetic Texts

Principal investigator: Ellen Spolsky, Assistant Professor, Dept. of English, U. of New Mexico, Albuquerque, NM 87106. *Associates:* Richard Kyrlach, Larry Wilson.

Objective: To determine the usefulness of dictionary definitions in determining the semantic content of a poem. Three dictionaries are being compared; The *NED, The Random House Unabridged Dictionary,* and *Webster's Seventh Collegiate Dictionary. Method:* Concordances are prepared of all the words in all the definitions of the words in a poem. The concordance will show which word groups occur most frequently and if there are patterns in the appearance of the words or word groups in the poem. Note on computer usage: The object programs are under 10k; however, the PL/I compiler will require storage based on the version used. The disk drives, used for output files, need have only about 500,000 bytes available for data, plus sort work areas.

Type of computer: IBM 360/67. *Size of storage:* 10k. *Language:* PL/I F. *Disks or tapes:* 2 2314 disk drives; 2 9-track tapes. Program is available.

References: "Computer-Assisted Semantic Analysis of Poetry," *Computer Studies in the Humanities and Verbal Behavior* 3,3 (October 1970), 163-168.

L373. FASTSCAN

Principal investigator: Randall L. Jones, Assistant Professor, Dept. of Modern Languages, Cornell U., Ithaca, NY 14850. *Associate:* Hale Trotter.

Scope: FASTSCAN is a high-speed scanner of natural-language texts. A specified character string (or several character strings) is given as input, and all occurrences are located and printed out together with a specified amount of context. Context-sensitive constraints may be specified, e.g., only when preceded by a certain character, only when followed by a certain character, etc.

Type of computer: IBM 360/65. *Size of storage:* 512k. *Language:* Assembler, FORTRAN IV. *Disks or tapes:* 1 2314 disk; 1 tape. [*CHum,* November 1970]

L374. Theory and Computational Techniques for Semantic Interpretation of Natural Language

Principal investigator: W. A. Woods, Assistant Professor, Dept. of Applied Mathematics, Computation Laboratory, Harvard U., Cambridge, MA 02138.

Scope: To investigate the use of formal rules and computational techniques for the semantic interpretation of natural language in applications where it is necessary for a computer to "understand" natural-language questions and commands and take appropriate actions. *Method:* Development and testing of experimental parsing and semantic interpretation systems and associated grammars and semantic rules.

Type of computer: XDS 940. *Language:* LISP.

References: "Procedural Semantics for a Question-Answering Machine," *AFIPS Conference Proceedings* 33 (1968 FJCC); "Augmented Transition Networks for Natural Language Analysis," Report CS-1, Computation Laboratory, Harvard University, Cambridge, MA, December 1969. [*CHum,* November 1970]

L375. Bibliography of S.T. Coleridge. An Annotated Bibliography of Scholarship and Criticism

Principal investigators: Richard Haven, Professor, Dept. of English, U. of Massachusetts, Amherst, MA 01002; W.B. Crawford, Associate Professor, Dept. of English, California State College, Long Beach, CA 90801. *Associates:* Maurianne S. Adams, Edward Lauterbach.

Objective: Preparation of a definitive bibliography of Coleridge scholarship and criticism from 1794 to 1970, annotated and with author-title and analytical indices of critical, biographical, and bibliographical subjects. Much of the work we have done with the computer has involved combinations of fairly standard sorting routines. The important "discovery" for us was that we could solve the problem of compiling an analytical index which required an impossibly long thesaurus by designing a schematic outline with coded categories within which variables could be entered. By this means, we reduced a 40-page thesaurus to a 3-page outline with single word variables. *Method:* Materials are necessarily collected by traditional methods, annotated, and coded. The computer makes possible the analytical indexing, continual updating, and specialized printouts.

Type of computer: CDC 3600. *Language:* FORTRAN IV.

L376. Normalization of Natural Language for Information Retrieval

Principal investigators: W.P. Lehmann, Director; Rolf Stachowitz, Research Coordinator; Linguistics Research Center, U. of Texas at Austin, P.O. Box 7247, University Station, Austin, TX 78712. *Associates:* Robert Pendleton, Bary Gold.

Objective: To develop a system to reduce paraphrases of a sentence to a representative form. *Method:* Computational analysis.

Type of computer: CDC 6600. *Size of storage:* 131k words. *Language:* FORTRAN IV. *Disks or tapes:* 2 6638 disks; 6 607 tapes.

Status: Completed.

L377. A Study of Proto-Indo-European Compounds

Principal investigator: W.P. Lehmann, Director, Linguistics Research Center, U. of Texas at Austin, P.O. Box 7247, University Station, Austin, TX 78712.

Objective: Analysis of compounds in Proto-Indo-European. *Method:* Computational analysis.

Type of computer: CDC 6600. *Size of storage:* 131k words. *Language:* FORTRAN IV. *Disks or tapes:* 2 6638 disks; 6 607 tapes.

Status: Completed.

L378. A Concordance to Wordsworth's 1805 *Prelude*

Principal investigator: Karl Kroeber, Professor, Dept. of English, Columbia U., New York, NY 10027. *Associates:* Michael Tinker, Barton Friedman.

Objective: A concordance to the 1805 version of *The Prelude* from the edition of Ernest de Selincour (2nd edition, Oxford, Clarendon Press, 1959—with specific permission of Clarendon Press). Since the handmade concordance to Wordsworth's work, ed. Lane Cooper (New York: Dutton, 1911) uses the 1850 version, the new concordance should aid in comparative studies of the two versions and contribute to our understanding of the development of Wordsworth's style and thought. Separate concordances have been made for Books II, XI, and 328 lines of Book VI. *Method:* Keypunch text on cards; transfer cards to tape and compile the concordance.

Type of computer: CDC 3600. *Size of storage:* 2 32, 768 48-bit word memory units. *Language:* 3600 FORTRAN, version 5.41. *Disks or tapes:* 3 3637/851 high-speed drums; 10 607 tape drives. *Special equipment:* The program BIBCON is available from Richard Venezky, University of Wisconsin. [*CHum,* November 1970]

L379. Comparative Concordances to James Beattie's *The Minstrel* and John Clare's *The Village Minstrel*

Principal investigator: Karl Kroeber, Professor, Dept. of English, Columbia U., New York, NY 10027.

Scope: The two poems, published fifty years apart, are of virtually the same length (123 to 120 Spenserian stanzas), and deal with approximately the same subject, Clare's poem being in part imitative of Beattie's. The total concordance of each work (no words are omitted) permits, therefore, a detailed comparison of style, language, prosody, etc., which, because of the nature and time of appearance of the two works should allow not only a uniquely refined study of "influence" but also a special investigation into the impact of "Romanticism" upon poetry. *Method:* Keypunch texts on cards and compile concordance from cards.

Type of computer: CDC 3600. *Size of storage:* 2 32,678 48-bit word memory units. *Language:* 3600 FORTRAN, version 5.41. *Disks or tapes:* 3 3637/851 high speed drums; 10 607 tape drives. *Special equipment:*The program BIBCON is available from Richard Venezky, University of Wisconsin. [*CHum,* November 1970]

L380. Navajo Reading Study

Principal investigator: Bernard Spolsky, Professor, Dept. of Anthropology and Elementary Education, U. of New Mexico, 1805 Roma NE, Albuquerque, NM 87106. *Associates:* Wayne Holm, Jock Embry, Babette Holliday.

Scope: As part of a study of the feasibility and effect of teaching Navajo children to read their own language first, we have been analyzing the spoken Navajo of six-year-old children. *Method:* Tape-recorded interviews with 200 Navajo children were transcribed and keypunched by a Navajo assistant. The corpus consists of 11,128 sentences (52,008 words [tokens] being 8,775 different words [types]). Computer processing produced a complete concordance with sentence environment, an alphabetic list of all words with frequency and range, a reversed alphabetic list, a frequency list, data on grapheme and letter unit frequency, and a loan word concordance.

Type of computer: 1) IBM 360/40, 2) IBM 360/67. *Size of storage:* 1) 102k, 2) 102k. *Language:* 1) and 2) PL/I F. *Disks or tapes:* 1 2314 disk; 2 800 tapes. Program is available.

References: "Proposal for Language Analysis Study," *NRS Progress Report No. 2,* August-September 1969; "Report of a Conference on the Language Sample, held in Albuquerque on December 4-5, 1969," ed. R. Paul Murphy, *NRS Progress Report No. 4,* December 1969 (ERIC Document); Bernard Spolsky, Wayne Holm, Jonathan Embry, "A Computer Assisted Study of the Vocabulary of Six-Year-Old Navajo Children," *NRS Progress Report No. 9,* August 1971; Bernard Spolsky, Wayne Holm, Babette Holliday, Judy Harvey, Marlene Atcitty, Irene Silentman, and Jonathan Embry, "A Spoken Word Count of Six-Year-Old Navajo Children," *NRS Progress Report No. 10,* August 1971; Wayne Holm, "Navajo Spelling Lists," *NRS Progress Report No. 11,* August 1971; Bernard Spolsky, Wayne Holm, Babette Holliday, and Jonathan Embry, "A Computer-Assisted Study of the Vocabulary of Young Navajo Children," *Computers and the Humanities* 7, 4 (March 1973), 209-218.

L381. Automatic Analysis of French and Spanish

Principal investigators: Josse de Kock, Professor, Dept. of Linguistics, U. of Louvain (K.U.L.), Verbindingsstraat, 10, 9120 Destelbergen, Belgium; Walter Bossaert.

Scope: Segmentation automatique de formes déclinées et conjuguées en morphèmes au moyen d'un critère d'économie. Les segments données ne sont pas indexés et les morphèmes sont supposés inconnus. Le programme est applicable à toutes les langues ayant un système morphologique analogue à celui des langues indogermaniques. Le programme a été appliqué avec succès à plusieurs séries de formes d'un corpus français (plus de 8000 formes) et espagnol (plus de 16.000 formes) en transcription phonétique, sans modification aucune. Le projet est achevé. Le même matériel a donné lieu à diverses exploitations phonologiques, dont un calcul de redondance tendant à définir les éléments redondants et distinctifs dans toutes les formes de l'inventaire français. Ce corpus sera utilisé sous peu en vue d'une confrontation avec un langage aphasique.

Type of computer: IBM 360/30. *Size of storage:* 64k. *Language:* PL/I. *Disks or tapes:* 2 2311 disks; 5 2401 tapes.

References: "Pour une segmentation morphologique formelle et automatique du français," *Preprint No.2; International Conference on Computational Linguistics* (1969), 10-11; "The Determination of Linguistic Structures by Means of a Computer. The Morpheme," *Actes du Colloque, L'utilisation des ordinateurs et la recherche en sciences humaines* (Bruxelles, 1971), 143-150; "De la definición de estructuras lingüísticas con la ayuda de un ordenador. El Morfema," *Introducción a la lingüística automática en las lenguas románicas,* ed. Josse De Kock (Madrid, 1973).

L382. Phonological Calculator

Principal investigators: Jonathan Kaye, Associate Professor, Centre for Linguistic Studies, U. of Toronto, Toronto, Canada; Peter Roosen-Runge, Professor, Department of Computer Science, U. of Toronto, Toronto, Canada.

Objective: Development of a phonological rule tester and parser for research and instructional purposes. Currently used in a graduate course in generative phonology, and is being applied to research problems in French and diachronic phonology of Germanic.

Type of computer: 1) IBM 370/165, 2) PDP-10. *Size of storage:* 1) 120k minimum, 2) 50k. *Language:* 1) SPITBOL Ver 2.2, 2) SNOBOL, Ver. 2. Program is available.

References: "A User's Guide to the Phonological Calculator (Ver.1.2, January 1972)" (mimeographed) available from the investigators.

L383. PHILTEX/ZPG: Philologische Texterschliessung: Text-Zerlegprogramm-Generator

Principal investigator: F. Schulte-Tigges, Leiter der Abteilung Nichtnumerik; W. Schenkel, Doktor, Leiter der Gruppe Linguistik und Literaturwissenschaften, Deutsches Rechenzentrum, D-61 Darmstadt, Rheinstrasse 75, Bundesrepublik Deutschland. *Associates:* I. Barth, A. Mathe.

Scope: Analyse präedierter Texte im Rahmen eines allgemeinen Programm-Systems für die philologische Texterschliessung. *Method:* Problemnahe Definition der Textstruktur in einer speziellen Sprache. Erzeugen eines geeigneten Programms mit Hilfe eines Programm-Generators.

Type of computer: 1) IBM 7094/I, 2) IBM 1401, 3) Telefunken TR440. *Size of storage:* 1) 32k, 2) 8k, 3) 256k. *Language:* 1) FORTRAN IV, MAP. *Disks or tapes:* 13 IBM 729/IV tapes; 1) 1 IBM 1301 disk and 3) WSP 414 disk. Program is available.

References: I. Barth, A. Duda, W. Schenkel, "Philologische Texterschliessung, Teil I: Der Zerlegprogrammgenerator ZPG des Anwendungssystems PHILTEX," *Schriftenreiche des Deutschen Rechenzentrums* S-11 (Darmstadt, 1970); "Philologische Texterschliessung, Teil I, Version 2 (TR440)," *Schriftenreiche des Deutschen Rechenzentrums* S-15 (Darmstadt, 1971); U. Ritrau, W. Schenkel, "Philologische Texterschliessung, Teil 2: Der Häufigkeitsgenerator HPG des Anwendungssystem PHILTEX," *Schriftenreiche des Deutschen Rechenzentrums* S-14 (Darmstadt, 1971); "Philologische Texterschliessung, Teil 2 (TR440)," *Schriftenreiche des Deutschen Rechenzentrums* S-16 (Darmstadt, 1971).

L384. Reduction of Greek Forms to Stems and Dictionary Forms

Principal investigator: James J. Helm, Assistant Professor, Dept. of Classics, Oberlin College, Oberlin, OH 44074.

Objective: Ultimately, the program should do lexical and thematic analyses of Greek drama; it should also be useful for the production of indices and concordances arranged according to stem rather than form. *Method:* Development of a series of subprograms to 1) complete elided forms, 2) separate contracted forms (crasis), and 3) remove inflectional endings, augments, reduplication, and possibly prefixes. A dictionary of irregular verb-stems will also be developed for conversion of stems to dictionary forms.

Type of computer: IBM 360/44. *Size of storage:* 128k. *Language:* FORTRAN IV. [*CHum,* November 1970]

L385. Thesaurus Studies

Principal investigator: Karen Sparck Jones, Professor, University Mathematical Laboratory, U. of Cambridge, Corn Exchange Street, Cambridge, England.

Objective: To investigate the use of a thesaurus, or semantic classification, in discourse analysis and generation, by experiments with an existing thesaurus. *Method:* Interactive computational.

Type of computer: Titan (an ICL Atlas 2). [*CHum,* November 1970]

L386. Inverse Catalan Dictionary

Principal investigator: Dennis K. Shakespeare, graduate student, Dept. of Romance Languages, U. of Massachusetts, Amherst, MA 01002. *Associates:* R. Joe Campbell, Alfred I. Towell.

Objective: To provide an inverse alphabetical listing of Catalan lexicon with grammatical categories included, as an initial goal.

Type of computer: CDC 3600. *Size of storage:* 65k. *Language:* SNOBOL4 and SORT2. *Disks or tapes:* 6 tapes. Program is available. [*CHum,* November 1970]

L387. Study of the Finnish Language

Principal investigator: Pauli Saukkonen, Professor, Suomen kielen laitos, Oulunyliopisto, Kirkkokatu 19, Oulu, Finland.

Scope: 1) The basic vocabulary of the Finnish language, 2) the microgrammars of various styles of the Finnish language, 3) the macrogrammar of the Finnish language. *Method:* About 6000 fragments from 60 hypothetical style classes, altogether 400,000 words transferred to punch cards from the period 1961-1968. [*CHum,* November 1970]

L388. Medieval Manuscript Photographic Information Retrieval

Principal investigator: Thomas H. Ohlgren, Assistant Professor, Dept. of English, Purdue U., West Lafayette, IN 47907. *Associate:* Gary Lelvis.

Scope: Storage of structured abstracts describing Bodleian Library medieval manuscript photographic holdings (18,000 color transparencies) and retrieval of catalog and fifteen separate indices, including three devoted to iconography.

Type of computer: IBM 370/155. *Size of storage:* 768k. *Language:* FORTRAN IV. *Disks or tapes:* 4 IBM 33-30 disks; 7 IBM 3420 tapes. Program is available. Basic indexing and retrieval system from Information Systems Lab, Michigan State University, East Lansing, MI 48823.

Status: 1) A catalog with keyword indices to 331 illuminations of eleven Apocalypse manuscripts in Oxford has been completed. 2) After a year's delay, which resulted from hardware and software changes, the formulation of a catalog with indices to 500 microfilm rolls is under way.

References: "Medieval and Renaissance Manuscript Photographic Information Retrieval," *Computer Studies in the Humanities and Verbal Behavior* (November 1972), pp.228-230; "An Archival Index to Medieval Manuscript Illuminations," a paper presented in the Computer Research section of the 1971 MMLA (available from the author).

L389. Statistical Researches in the Latin Dactylic Hexameter

Principal investigator: Pol Tordeur, Professor, Dept. of Greek and Latin, Rue Wery, 51, B1050 Brussels, Belgium. *Associate:* Edm.Lienard.

Objective: 1) To establish, in the poetic works of about twenty Latin poets, from 200 B.C. to A.D.600, the difference in the use and the places of the different word shapes, in the affirmative case, if this difference is to be linked with the chronology of the style of the authors; particular research about the pyrrhics; 2) to establish and publish a complete word index of late Latin poets. *Method:* Writing each word, with reference, metrical shape and place in the verse, on a separate card; then compiling the cards, first following the metrical order, i.e., the different word shapes; later following an alphabetical order to obtain a complete word index of the work. [*CHum*, November 1970]

L390. Computer Aids to Linguistic Research

Principal investigator: Joyce Friedman, Associate Professor, Dept. of Computer and Communication Sciences, U. of Michigan, 4028 L.S.A. Bldg., Ann Arbor, MI 48103.

Objective: To develop, maintain, and extend a computer system which accepts as input a transformational grammar and produces sentence derivations; to apply the system in research and teaching.

Type of computer: IBM 360/67. *Size of storage:* 300k. *Language:* FORTRAN G. Program is available.

References: Friedman, Bredt, Doran, Martner, and Pollack, *A Computer Model of Transformational Grammar* (American Elsevier, 1971); J. Friedman, "Directed Random Generation of Sentences," *Communications of the ACM* 12 (1969), 40-46; J. Friedman, "A Computer System for Transformational Grammar," *Communications of the ACM* 12 (1969), 341-348; J. Friedman and Y.C. Morin, *Phonological Grammar Tester Description* (Ann Arbor: University of Michigan Press, 1971); Y.C. Morin and J. Friedman, *Phonological Grammar Tester: Underlying Theory* (Ann Arbor: University of Michigan Press, 1971).

L391. Studies in Statistical Linguistics in the Book of Isaiah Concerning Its Unity

Principal investigator: Yehuda Th. Radday, Senior Lecturer, Dept. of General Studies, Technion–Israel Institute of Technology, Haifa, Israel.

Objective: To determine the homogeneity of all chapters in the Book of Isaiah, which is said by some scholars to be written by two or more prophets. The argument both for and against its unity is based on history, theology, and style, but these change "within" the same author. Moreover, their evaluation is necessarily a highly subjective matter. The purpose of the present study is to arrive at an objective, i.e., quantitative yardstick in order to measure by which probability sections II-VI of the book should be attributed to the author (Isaiah) of the undisputedly authentic section I (chaps. 1-12). *Method:* Such involuntary, and therefore inimitable, linguistic criteria were investigated as sentence length, word length, frequencies of parts of speech, text entropy, "special" vocabulary facultative particles, vocabulary eccentricity and richness, etc.

Type of computer: 1) Elliott 503, 2) IBM 360/50. *Size of storage:* 1) Main storage 8k, core backing storage 16k, 2) 256k. *Language:* 1) ALGOL, 2) OS.

Status: Completed 1969.

References: The Unity of Isaiah in the Light of Statistical Linguistics (Hildesheim: Verlag Gerstenberg, 1972); *An Analytical Linguistic Concordance (computerized) of Isaiah,* Biblical Research Associates (Wooster, OH: College of Wooster, 1971).

L392. Wortindizes zu den Briefen Hölderlins

Principal investigators: Alexander Fellmann, Wissenschaftlicher Assistant, Neuropsychiatrische Klinik, Justus-Liebig-U., D-6300 Giessen, Am Steg 18; Niels Sörensen, cand. phil., Forschungsinstitut für Deutsche Sprache, Philipps-U., D-3550 Marburg, Krummbogen 28A, W. Germany.

Scope: Briefe Hölderlins (etwa 130,000 Wörter) zum Studium schizophrenen Sprachzerfalls unter sprachstatistischen und semantischen Gesichtspunkten. Gesamtindex und Einzelindizes zu den Briefen sowie Indizes der Briefe nach biographisch relevanten Zeitabschnitten und nach Adressaten. *Method:* Ubertragung des Textes nach der Grossen Stuttgarter Ausgabe über Lochstreifen auf Magnetband. Erstellung von Wortindizes und Worthäufigkeitsregistern.

Type of computer: Telefunken TR4. *Size of storage:* 32k zu je 48 bits. *Language:* FORTRAN TR4, TEXAS. *Disks or tapes:* 4 disks, 4 tapes. [*CHum,* November 1970]

L393. Concordancias y diccionarios de textos en lenguas semiticas (especialmente árabe, hebreo, arameo)

Principal investigator: Juan Vernet-Gines, Catedrático de Universidad, Dept. de Arabe, Facultad de Filosofía y Letras-U., Barcelona, Espana. *Associate:* Carlos Simó.

Scope: Establecer indices de frecuencias en textos scientíficos y literarios medievales. *Method:* Despoje de textos seleccionados y, en especial, del diván de Ibn Quzman y de textos modernos, para el árabe; textos del Pentateuco para el hebreo y del Targum Neofiti para el arameo.

Type of computer: IBM 360/30. *Language:* PL/I. *Disks or tapes:* IBM 2311 disks.

L394. A Concordance to Leopoldo Panero's Poetry

Principal investigator: Enrique Ruiz-Fornells, Professor, Dept. of Romance Languages, U. of Alabama, P.O. Box 4932, University, AL 35486. *Associates:* Gladys Neggers, Maura Rainez, Jorge Isquierdo.

Type of computer: 1) UNIVAC 1108, 2) IBM 360/50. *Size of storage:* 1) 65k words, 2) 256k bytes. *Language:* 1) FORTRAN V, SLEUTH II, 2) FORTRAN IV, BAL. *Disks or tapes:* 1) 10 VIII C tapes; 4 432 and 1 782 disks; 2) 4 2514 tapes; 2314 disk. *Special equipment:* IBM text chain. Program is available if you have a comparable 1108 configuration.

L395. Aymara Materials Project

Principal investigator: Martha J. Hardman-de-Bautista, Professor, Dept. of Anthropology, U. of Florida, Gainesville, FL 32601.

Objective: Preparation of dictionary for Andean languages. The program allows 21 categories for each entry and output according to any one or combination of these categories. The output will be used for normal dictionary use and also for comparative work in Andean languages. A concordance program is being used jointly for taking items from texts in the languages involved.

L396. D.A.T.U.M.

Principal investigator: Jacques Boucher, Professor, Dept. of Faculty of Law and Computing Center, U. de Montréal, C.P. 6128, Montréal 101, PQ, Canada. *Associates:* Michel Dignard, Claude Fabien, Lise Legeaulty, Ejan Mackaay, Serge Poulard, Pierre Stewart.

Objective; Realization of an on-line retrieval system for legal cases, some of which are in English, the rest in French. System uses full text. Machine-aided development of a French/English synonym dictionary and dictionary of grammatical forms. Queries in one language will retrieve both English and French documents.

Type of computer: CDC 6400. *Size of storage:* 640k characters (6 bits). *Language:* Assembler, FORTRAN IV. *Disks or tapes:* 6638 and 841 disks. *Special equipment:* Mohawk DATARECORDERS 1181 [*CHum,* November 1970]

L397. Project on Linguistic Analysis (POLA)

Principal investigator: William S-Y. Wang, Professor, Dept. of Linguistics, U. of California, Berkeley, CA 94720. *Associates:* S. Chan, J. Ohala.

Scope: Linguistic research: 1) phonological history, 2) contrastive lexicography, 3) experimental phonetics.

Type of computer: 1) CDC 6400, 2) Linc 8, 3) PDP-11. *Special equipment:* Speech synthesizer, electromyography facility.

References: POLA reports, second series, 1-15.

L398. Renaissance and Restoration Prose Style; A Preliminary Computer Sample

Principal investigator: Robert Cluett, Associate Professor, Dept. of English, York U., 234 Winters College, 4700 Keele Street, Downsview 463, Ontario, Canada. *Associates:* G.J. Carpenter, Michael Rehner, Derek Dalton, Tom Greenwald, Colin Rutledge.

Scope: 100 samples from 37 authors, 1570-1715, plus an additional 25 samples from 14 authors, 1715-1870, analyzed for distribution of word classes, sentence openings, sentence closings, anaphora, pattern variety, certain prepositional phrase strings, and certain formulaic word parallelisms. Authors (with number of samples) include: Lodge (1), Lyly (2), Sidney (6), Nashe (2), King James Bible (2), Donne (4), Bacon (6), Jonson (1), Hobbes (3), Burton (3), Clarendon (3), Taylor (6), Raleigh (2), Walton (2), Browne (4), Milton (5), Locke (3), Rymer (2), Savile (2), Bunyan (2), Sprat (6), Glanvill (2), Boyle (3), Marvell (3), Burnet (5), Tillotson (3), Stillingfleet (3), Robert Howard (2), Dryden (4), plus one each from Collier, Cowley, Lestrange, Andrewes, Coxere, Newton, Wilkins, and John Davis. *Method:* Basically the Milic method, with the addition of over 60 word classes to the 24 that he derived from Fries's grammar. These classes have been added by the use of a third digit to the classes of the Fries-Milic code; tapes thus can be interchangeable with those developed in the original code.

Type of computer: IBM 360/50. *Size of storage:* 512k. *Language:* FORTRAN G. *Disks or tapes:* 1 2314 disk. *Special equipment:* 2540 card reader, 2540 card punch, 1403 line printer.

Status: Completed November 1972. Data still being sorted out.

References: Article on Sidney in *Language & Style,* 1972.

L399. A Concordance to the *Libro de Buen Amor* of Juan Ruiz

Principal investigators: Marion F. Hodapp, Assistant Professor; William W. Moseley, Professor; Dept. of Foreign Languages, Colorado State U., Fort Collins, CO 80521.

Objective: Preparation of a concordance to the *Libro de Buen Amor* of Juan Ruiz, with location of word by stanza and line number, with each occurrence in context, reverse index, and ranking list of frequencies. *Method:* The text is punched on cards, one line per card; transferred to magnetic tape, with final printout by computer-linked IBM Selectric typewriter.

Type of computer: 1) CDC 6400, 2) Librascope, 3) Athena-Selectric. *Size of storage:* 1) 64k, 2) 32k. *Language:* FORTRAN IV. *Disks or tapes:* 1 disk, 6 7-track tapes. For program, contact Center for Computer Research in the Humanities, U. of Colorado, Boulder, CO 80302.

L400. Le Langage Poétique de Paul Valéry

Principal investigator: Pierre Laurette, Associate Professor, Etudes françaises, U. Carleton, Colonel By Drive, Ottawa 1, ON, Canada.

Scope: Analyse descriptive, statistique et formelle de la poésie de P. Valéry. Index alphabétique, index de fréquence, index inverse. Concordance alphabétique, concordance syntagmatique distributionnelle. Codage phonétique, sémantique, syntaxique rhétorique. *Method:* Travail de pré-édition: transcription du texte sur formulaire FORTRAN. Lemmatisation. Après chaque mot, filtre de sept cases pour le codage sémantique, syntagmatique, syntaxique, et phonétique. Après chaque phrase, filtre de 160 cases pour codage linguistique et stylistique. Chaque case du filtre donne d'une à 62 informations. Cartes perforées et bandes.

Type of computer: XDS Sigma 7. *Size of storage:* 80k. *Language:* SNOBOL, FORTRAN IV. *Disks or tapes:* 1 tape. [*CHum,* May 1971]

L401. The Old English *Exodus:* A Glossary

Principal investigator: Peter J. Lucas, Lecturer, Dept. of Old and Middle English, University College, Belfield, Dublin 4, Ireland. *Associate:* Angela M. Lucas.

Scope: Full *index verborum* of the complete text of *Exodus* edited from Junius II in the Bodleian Library, Oxford; to be used as the basis for the glossary of a new edition (to be published in Methuen Old English Library series). *Method:* Transliterate original, keypunch text onto cards, sort alphabetically, then print out in tabular form.

Type of computer: IBM 1620/1. *Size of storage:* 40k. *Language:* SPS II. *Disks or tapes:* 1 paper tape.

L402. Statistical Analysis of Old Japanese Phonology

Principal investigator: Gerald B. Mathias, Assistant Professor, Dept. of East Asian Languages and Literatures, Indiana U., Bloomington, IN 47401.

Objective: To find frequencies of various patternings of phonological elements in eighth-century Japanese. Such information should be useful in the reconstruction of the phonology of the period, as well as providing insights in prehistorical word-formation and phonological development. *Method:* Test vocabulary lists and texts of the major sources of Old Japanese (the Man'yoshu, Kojiki, and Nihonshoki) for occurrences of particular patterns or elements.

Type of computer: CDC 6600. *Size of storage:* 160k. *Language:* SNOBOL4. Program is available.

L403. Structural Thematics: Concordances for Pattern Analysis of English

Principal investigator: L.D. Misek, Assistant Professor, Dept. of Computer Science Studies, Vassar College, Poughkeepsie, NY 12601. *Associates:* E.L. Glaser, William Cornwell.

Scope: A library of programs for studying contextual aspects of natural language messages. Primary concern is for standardization of data input conventions across texts drawn from different centuries of production, as well as different native language bases (French and German are now included in the investigation). *Method:* In addition to literal counts on explicit language phenomena at the lexemic level, the programs store expectations as to certain syntactic and semantic pattern norms. Content analysis includes both hand-indexing and automatic recognition.

Type of computer: UNIVAC 1108. *Size of storage:* 131k. *Language:* ALGOL, FORTRAN V. *Special equipment:* 2 FASTRAND, onsite 1004 line printers.

References: Automated Contextual Analysis of Thematic Structure in Natural Language (Cleveland: Case Western Reserve University, Jennings Computing Center Report 1108). [*CHum,* November 1970]

L404. Concordance to Early Tamil Literature

Principal investigator: R. Panneerselvam, Visiting Fellow, Scandinavian Institute of Asian Studies, 2, Kejsergade, 1155 Copenhagen K, Denmark. *Associate:* Donald B. Wagner, R.W. Last.

Objective: To bring out a complete vocabulary and concordance, with English equivalents and grammatical notes, and to make it available to the scholars of literary and linguistic interest for evaluating the authorship or for describing the structure of the language.

Type of computer: 1) and 2) A/S Regnecentralen GIER. *Size of storage:* 1) 1k, 2) 5k. *Language:* 1) and 2) ALGOL 60. *Disks or tapes:* Drums: 1) 1 40,000 words 40 bits, 2) 1 400,000 40 bits (Analex). [*CHum,* November 1970]

L405. Le Cantique des Cantiques (MS. 173. Bibl.Mun. Le Mans)

Principal investigators: C.E. Pickford, Professor, Dept. of French; R.W. Last, Lecturer, Dept. of German, U. of Hull, Yorkshire, England.

Scope: Complete concordance and word-frequency count for publication, to act as companion volume to Professor Pickford's edition of the text.

Type of computer: ICL 1905E. *Size of storage:* 48k. *Language:* ALGOL/PLAN. *Disks or tapes:* 2 EDS disks, 4 tapes. Program is available. [*CHum,* November 1970]

L406. Word Count Research of Newspaper Texts

Principal investigator: Shiro Hayashi, Chief, 4th Research Division, National Language Research Institute, 3-9-14, Nisigaoka, Kita-ku, Tokyo, Japan. *Associates:* A. Tanaka, T. Ishiwata, H. Saito, H. Nakano, A. Turuoka, S. Muraki.

Scope: The descriptive study on the vocabulary of contemporary Japanese through the word counting research on the texts of representative newspapers, *Asahi, Mainiti* and *Yomiuri* in 1966. Sample size, 3 million words, one sixtieth of a universe of 180 million. *Method:* Words are counted by a computer. For the input/output of Kanzi (Chinese character) data, Kanzi teletypewriters and Kanzi teleprinters are used. Major contents of pre-editing are 1) segmentation of sentence into words, 2) determination of Kanzi pronunciation, 3) classification of word (part of speech, etc.).

Type of computer: HITAC 3010. *Size of storage:* 20k characters. *Language:* Assembler. *Disks or tapes:* 6 H-382 tape decks. *Special equipment:* Kanzi teletypewriter. Program is available.

References: Studies on the Vocabulary of Modern Newspapers, vol. 1 (1970), vol. 2 (1971), vol. 3 (1972) (Tokyo: The National Language Research Institute, Syuei Syuppan Company).

L407. Oirat-Mongolian to English Citation Dictionary

Principal investigator: John R. Krueger, Professor, Dept. of Uralic and Altaic Studies, Indiana U., GB-142, Bloomington, IN 47401.

Scope: Oirat is the Western Mongolian literary language used for Buddhist scriptures and translations from 1648 to about 1942. The dictionary proposes to embrace the whole vocabulary of the literary language and its works during these times. It will thus form a basic reference work for any person studying Mongolian history, literature, culture, and civilization. *Method:* Listing of individual words with citations based on usage, meaning, and cognates.

Type of computer: CDC 6600. *Language:* FORTRAN. Program is available.

References: "Oirat Literary Resources and Problems of Oirat Lexicography," in Mid-West Branch Semi-Centennial volume, *American Oriental Society,* Bloomington, IN, 1969, pp. 134-156; "The Application of Computer Processing to Altaic Studies," *Central Asiatic Journal* 11 (1966), 161-186.

L408. Computer Poetry Generation

Principal investigator: Louis T. Milic, Professor, Dept. of English, Cleveland State U., 24th and Euclid Ave., Cleveland, OH 44115. *Associate:* A. Saulino.

Scope: The general goal of this investigation is the search for what is essentially poetic in random, well-formed strings of English words. The specific goal is the reconstitution by synthetic means of an original poem whose syntax and vocabulary have been analyzed. *Method:* A poem written in English is taken as the base. Vocabulary items from the poem constitute the lexical resources of the program. Algorithms are written which incorporate randomness and uncertainty and constitute the matrices into which the lexical material is inserted.

Type of computer: IBM 360/40. *Size of storage:* 256k. *Language:* SNOBOL4 2.0. Program is available. Presently at work on a program to generate monometer poems and another to do heroic couplets.

References: "Computer Programs and the Heroic Couplet," *Think* (May-June 1968), pp.28-31; "The Possible Usefulness of Poetry Generation," *Proceedings of the Symposium on Literary and Linguistic Uses of the Computer* (Cambridge University, March 24-26, 1970), ed. Roy A. Wisbey (Cambridge: Cambridge University Press); *Erato,* (CSU Poetry Booklet No. 4, Cleveland: The Cleveland State University Poetry Center, 1971); "The 'Returner' Poetry Program," *Institute of Applied Linguistics (ITL)* 11 (1971), 1-23; "Computer Poetry," *Princeton Encyclopedia of Poetry and Poetics,* ed. Alex Preminger et al. (in press); "'Returner' Re-Turned," *Midwest Quarterly* 13 (July 1972).

L409. Concordances et étude philologique approfondie de la *Regula Magistri*

Principal investigator: Dom Jacques Froger, Chargé de cours associé; Paul Tombeur, Directeur du CETEDOC, *c*entre de *t*raitement *e*lectronique des *Do*cuments, U. Catholique de Louvain, Bogaardenstraat, 81, 3000 Leuven, Belgium.

Scope: Index, concordances diverses, relevés statistiques divers concernant la *Regula Magistri. Method:* Le texte critique établi par Dom Froger est enregistré sur cartes perforées puis sur bandes magnétiques. L'ordinateur référencie automatiquement chaque mot du texte. Celui-ci est ensuite analysé par la machine et par les philologues. L'ensemble des analyses est enregistré sur la bande texte référenciée. L'ordinateur établit ensuite toutes sortes de concordances et divers relevés statistiques concernant la lexicographie, la grammaire et le style de la *Regula Magistri.*

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* COBOL. *Disks or tapes:* 2311 disk, 2402 et 2403 tapes. [*CHum,* November 1970]

L410. An Author-Title Survey of Cataloged Greek Manuscripts

Principal investigator: Walter M. Hayes, Research Associate, Pontifical Institute of Mediaeval Studies, 59 Queens' Park, Toronto 5, Canada. *Associate:* Lidwine Fitzgerald.

Objective: To provide classical, New Testament, patristic, medieval, and Byzantine scholars with a complete listing of all cataloged Greek manuscripts for which there exist handwritten copies. This will provide the means of evaluating author popularity and culture diffusion as witnessed by the manuscript. *Method:* The manuscript catalogs themselves are numbered in M. Richard, *Repertoire des Bibliothèques et des Catalogues des Manuscrits Grecs.* We proceed to the catalog itself and list: 1) the name of the catalog, 2) source language, 3) volume number, 4) page number, 5) manuscript title, 6) material of the manuscript, 7) total number of folios, 8) date of manuscript, 9) size of manuscript, 10) Greek author's name, 11) title of his work, 12) entry number of this work in manuscript, 13) folio numbers where work is located in manuscript, 14) notes.

Type of computer: GE 215.

L411. Computer-Assisted Tutorial and Testing System (CATTS)

Principal investigator: Frank A. Yett, Director, Center for Computer Assisted Learning and Chairman, Dept. of Computer Sciences, Pasadena City College, 1570 E. Colorado Boulevard, Pasadena, CA 91106. *Associates:* Winifred Acker, William Johns.

Objective: To determine the effectiveness of an individually paced, continuous, computer-assisted training system to bring students in a freshman remedial English course to criteria levels of achievement for correctly spelling a college-level English vocabulary. *Method:* A pre-training period test is administered to an entire cadre of students in a regular training program. A computer-trained and a control (regular program) group, matched on pre-test scores, will be post-tested immediately upon the completion of the computer-assisted course.

Type of computer: Burroughs B3500. *Size of storage:* 90,000k. *Language:* FORTRAN IV. *Disks or tapes:* 2 7-track tapes.

References: Computer-Assisted Tutorial and Testing System (CATTS). [*CHum,* May 1971]

L412. Computer Analysis of Literature and the Teaching of French

Principal investigator: Clodius Willis, Assistant Professor, Dept. of Romance Languages, U. of Vermont, Burlington, VT 05401.

Scope: Programs are developed to count occurrences of lexical and grammatical items in any French text. Results will be used in preparing lexical and grammatical descriptions of literary texts for use at the intermediate level of French. *Method:* Develop the following programs: 1) Word find, and word count (completed). 2) Grammatical analysis.

Type of computer: 1) IBM 360/44, 2) PDP-10. *Size of storage:* 1) 256k, 2) 64k. *Language:* 1) and 2) FORTRAN IV. *Disks or tapes:* 1) 2 tapes. Program is available. [*CHum,* May 1971]

L413. A Study of Grammatical Structure Frequencies in Modern Journalistic German

Principal investigators: Bengt Nilsson, Research Coordinator; Ebbe Lindell, Dept. of Educational and Psychological Research, Malmö Schoôl of Education, Fack, 20045 Malmö, Sweden.

Objective: To explore the primary grammatical structures in modern journalistic German. *Method:* A sample of the German newspaper *Bildzeitung* is analyzed by special program.

Type of computer: UNIVAC 1108. *Size of storage:* 64k. *Language:* FORTRAN IV.

References: B. Nilsson, "Strukturfrekvenser i tyskt tidningssprak ("Bildzeitung"): Metodkonstruktion och nagra förstudier" [A study of grammatical structure frequencies in modern journalistic German ("Bildzeitung"): Method construction and pilot studies], *Pedagogisk-psykologiska problem* 88 (1969), Malmö, Sweden: School of Education. (In Swedish) [*CHum,* May 1971]

L414. Computer Studies on Natural Languages as Examples of Information Systems

Principal investigators: Margaret Masterman; Yorick Wilks, Cambridge Language Research Unit, 20 Millington Road, Cambridge, England. *Associate:* Robin Anderson.

Objective: To make automatic abstracts of paragraphs of running text. *Method:* Using a phrasing program, the text is first broken up into phrasings. A simplified syntax analysis is then applied to the phrased text. That is followed by a semantic abstracting algorithm based on a notion of "semantic redundancy." [*CHum,* May 1971]

L415. Computerizing Serial Surveys

Principal investigator: Jane Collier, Cambridge Language Research Unit, 20 Millington Road, Cambridge, England. *Associates:* Bernard Adamczenski, Andrew Barbour.

Objective: To enable the computer to assign codes to natural language items in social surveys of various kinds.

Type of computer: Titan. *Language:* SNOBOL4. Program is available. [*CHum*, May 1971]

L416. Redefinition and Retrieval Algorithms to Handle Technical Terms in the Proposed Canadian On-Line Banque-des-Mots

Principal investigator: David Shillan, Norfolk Lodge Cottage, Terrace Lane, Richmond TW10 6NF, Surrey, England. *Associates:* Margaret Masterman, A.F. Parker-Rhodes, Jane Collier.

Objective: 1) To examine the general role of a reactive computerized data bank of technical terms for the use of translators; 2) to develop algorithms for handling such a data bank; 3) to investigate the possibility of automatic insertion of new dictionary entries from text. The analogy is being explored between algorithms which are being developed at CLRU for the automatic processing, coding, and cataloging of technical descriptions and algorithms which are needed to handle a Banque-des-Mots. [*CHum*, May 1971]

L417. Machine-Aided Translation

Principal investigator: David Shillan, Norfolk Lodge Cottage, Terrace Lane, Richmond TW10 6NF, Surrey, England. *Associates:* Margaret Masterman, J.E. Dobson, R.A. McKinnon-Wood.

Objective: To develop a system of machine-aided translation from English to French on a "phrasing-for-phrasing" basis, producing low-level but message-preserving output. *Method:* The text is segmented by computer program into natural "phrasings." The text is processed phrasing by phrasing, using a closed list of structure-words (called "links"), which are translated as part of their "phrasing-frame." Content-words (called "signals") are treated by separate dictionary only after they have been replaced by semantic markers, reducing the number of phrasing-frames required. The result aimed at is a French output that is comprehensible and semantically accurate, easily capable of being polished by a trained translator to publishable standards. It is assumed that large amounts of text can be prepared in this way, and the translator is relieved of much slow preparatory work.

Type of computer: 1) ICT 1202, 2) Titan, 3) Modular One. *Language:* 1) and 2) Machine Code, 3) TRAC.

References: Margaret Masterman, "Man-aided Computer Translation from English into French using an On-line System to Manipulate a Bi-lingual Conceptual Dictionary or Thesaurus"; David Shillan, "Segmenting Natural Language by Articulatory Features," Second International Conference on Computational Linguistics, Grenoble, France, 1967. [*CHum*, May 1971]

L418. Interactive Lexicon Construction

Principal investigator: Raoul N. Smith, Assistant Professor, Dept. of Linguistics, Northwestern U., Evanston, IL 60201. *Associate:* Martha W. Evens.

Objective: Updating of syntactic and semantic features in a lexicon. *Method:* Interactive scheme using informant responses to generated sentences as indicator of feature marking.

Type of computer: CDC 6400. *Size of storage:* 40k. *Language:* SNOBOL4. *Disks or tapes:* 1 6603 disk. Program is available.

References: Raoul N. Smith, "Interactive Lexicon Updating," *Computers and the Humanities* 6, 3 (1972).

L419. Phonological Contrasts Between Verbal Art and Verbal Non-Art

Principal investigator: John B. Lord, Professor, Dept. of English, Washington State U., Pullman, WA 99163. *Associate:* Barry Abrahamsen.

Objective: Determination of phonological norms (distinctive acoustic features) in utilitarian American English, for comparison with poetry, for genre differences and stylistic variations. *Method:* Expansion of a sample into its full matrix (including all redundancies) of distinctive acoustic features; determining and counting all string lengths for each feature; and finding average concentration per 100 letters of each feature. Comparison of such counts in verbal non-art to corresponding counts in verbal art; comparison of non-art to corresponding counts in verbal art; comparison of given poets to non-artistic norms, or to each other, or of several poems by one poet, in these phonological respects. Discovery of particular acoustic patterns in particular poems or poets.

Type of computer: IBM 360/67. *Size of storage:* 1800k. *Language:* FORTRAN IV, G and H. *Special equipment:* Data cell. Program is available. [*CHum*, May 1971]

L420. Mathematical Analysis of Lyric Poetry

Principal investigator: B.H.D. Dangerman, Research Supervisor, Institut für mathematisch-empirische Systemforschung (MESY), 65 Theaterstrasse, 51 Aachen, West Germany. *Associates:* H.M. Dannhauer, B.B. Rieger, H.D. Wickmann.

Objective: Quantitative analysis of lyrics (German) in view of a theory of descriptive esthetics. Investigation in semantic characteristics of textual elements: "word and motive complexes," their affinities and repugnancies, compilation of a synoptic dictionary (mapping of their historic changes). *Method:* Derivation of a new probability distribution based on a model analogous to that of Bose-Einstein in quantum statistics.

Type of computer: CDC 6400. *Size of storage:* 96k. *Language:* FORTRAN IV, CDC 6400 Assembly language. *Special equipment:* Extended core (ECS).

References: B.B. Rieger, "Uber die Erstellung eines Synoptischen Affinitäts-Lexikons (SAL)," *Literatur und Datenverarbeitung,* ed. H. Schanze (Tübingen: Niemeyer, 1972).

L421. Computer Count and Analysis of Million-Character Corpus of Modern Chinese

Principal investigator: Frank A. Kierman, Senior Research Sinologist, Chinese Linguistics Project, Princeton U., Princeton, NJ 08540.* *Associates:* E.J.W. Barber, Jerry Norman.

Objective: To collect in machine-readable form a million-character corpus of modern vernacular Chinese literature (since 1920) for purposes of general research; to produce frequency counts of the individual characters and of the more frequent multi-character groups; to produce some sample (partial) concordances and provide the materials for scholars to make other concordances from the corpus. *Method:* Text coded onto punched paper tape, transferred to magnetic tape, converted to Standard Telegraph Code, corrected, and run through counting and concording programs (by Device Development Corp., Lawrence, MA, using programs previously written for the American Mathematical Society). Other concordances prepared by tape conversion using an IBM 360 version of TRICON.

Type of computer: 1) Honeywell 200, 2) IBM 360/91. *Size of storage:* 1) 24k, 2) 160k, 225k. *Language:* 1) EASYCODER, 2) FORTRAN IV, Assembler, SNOBOL4. *Special equipment:* Modified teletype (Chinese Teleprinter Corp. Model 57b). TRICON program is available.

References: F.A. Kierman and E. Barber "Computers and Chinese Linguistics," *Unicorn* 3 (1969), 29-73; W.G. Boltz, E. Barber, and F.A. Kierman, "Progress Report on Pai-hua Wen Computer Count and Analysis," *Unicorn* 7 (1971), 94-138 (including manual for the IBM 360 version of the TRICON concordance program, by E. Barber and M. Eads). [*CHum,* May 1971]

L422. Chinese Dictionary Link-up

Principal investigator: Eric Grinstead, Lektor amanuensis, Ostasiatisk Institut, Copenhagen U., Grabrodre Torv, 1155 Copenhagen, Denmark.

Objective: Manipulation of sound changes in Chinese. *Method:* The standard character is defined by the number in Morohashi's dictionary and the standard pronunciations by *Grammata serica recensa* (Karlgren). The first aim is to provide a combined index. Karlgren 1-500 is already typed (approximately half of the typing effort).

Type of computer: GIER 2. *Language:* ALGOL 4. [*CHum,* May 1971]

L423. Analisis Lingüístico Automatica del Espanol

Principal investigator: Josse De Kock, Professor, Dept. of Linguistics, U. of Louvain (K.U.L.), Verbindingsstraat, 10, 9120 Destelbergen, Belgium. *Associate:* Dirk Geens.

Objective: Linguistic and stylistic study of modern Spanish literary texts by limited but strictly formal algorithms that do not require pre-indexing or dictionary look-up. 206 articles from 1931-1936 and the *Cancionero* of M. de Unamuno (more than 300,000 words) are presently being studied. A computer-aided quantitative confrontation of both texts has been initiated. A selection of prose texts of modern Hispanic authors (non-literary texts of M. de Unamuno, A. Machado, Azorín, P. Salinas, J.R. Jiménez, J.L. Borges, M. Asturias) are now in process.

Type of computer: 1) IBM 360/40 and 2) IBM 360/44. *Size of storage:* 1) and 2) 128k. *Language:* 1) PL/I. *Disks or tapes:* 1) 3 9-track (EBCDIC) tapes and 2) 1 9-track (EBCDIC) tape. Program is available.

References: "Pour une définition de la rareté de certains tours syntaxiques en Espagnol, ou de l'emploi de l'ordinateur dans les recherches syntaxiques," 3rd International Congress of Applied Linguistics, Copenhagen, 1972.

L424. Soviet Academy *Grammar:* Phonology and Morphology. A Computer-Aided Index

Principal investigator: Demetrius J. Koubourlis, Assistant Professor, Dept. of Foreign Languages and Literatures, U. of Idaho, Moscow, ID 83843.

Objective: To make the Soviet Academy *Grammar,* a large reference work without an index, more accessible to Russian language specialists; to develop techniques for manipulating non-Roman alphabet data and preparing camera-ready copy. *Method:* Entries were created in Roman alphabet transliteration, keypunched, and stored on tape. All subsequent handling, including correcting, cross-referencing, converting for Cyrillic alphabetic order, sorting, decapitalizing, and "typesetting" were accomplished via computer.

Type of computer: IBM 360/75. *Size of storage:* 250k. *Language:* PL/I F. *Disks or tapes:* 6 disks, 4 9-track tapes. Program is available.

Status: Completed spring 1972.

References: Soviet Academy Grammar: Phonology and Morphology. A Computer-Aided Index (Moscow, ID: University of Idaho Research Foundation, 1972); "Computer Sequencing and Non-Alphabetic Interference in Language Data Processing," *Computers and the Humanities* 7, 3 (January 1973), 149-155.

L425. Phonemic and Phonetic Frequencies in Latin

Principal investigators: Ladislaus J. Bolchazy, Assistant Professor, Dept. of Classics, La Salette Seminary, Altamont, NY 12009; H. Berkley Peabody, Professor, Dept. of Comparative Literature, SUNY, 1400 Washington Avenue, Albany, NY 12203. *Associates:* Chin Shih Yu, Lorraine Noval.

Objective: To establish phonemic and phonetic relative frequencies in: *Aeneid* I, IV, XII; *Eclogues; Culex, Moretum;* Lucan I, X; Ovid, *Metamorphoses* I, XII: Lucretius I, III; Cato, *De Agri Cultura;* Sallus, *Catiline, Jugurtha;* Varro, *De Lingua Latina* V; Ammianus XIV, XV; and in random passages in Caesar, Cicero, Catullus, Horace, Valerius Flaccus. The goals are to derive a statistical formula for alliteration and assonance, to establish stylometric criteria based on preferred sounds, to see if significant differences in phonemic and phonetic frequencies exist between prose and verse, and between early and late Latin. *Method:* The phonemic and phonetic distinctions are based on Allen and Sturtevant. 1) Program A produced absolute and relative frequencies for the letters of the alphabet in some of the works mentioned. (These tables are now available.) 2) Program B produced absolute frequencies of the letters of the alphabet in their specific environments (phonemic and phonetic elements). Long and short vowels will be counted by hand. 3) Averages and statistical probabilities will be established.

Type of computer: UNIVAC 1108. *Size of storage:* 131k. *Language:* COBOL. *Special equipment:* FASTRAND. Program is available.

L426. A Reverse Frequency List of Dutch

Principal investigator: Willy J. Martin, Assistant Professor, Katholieke U. Leuven, Dept. of Linguistics, Institute of Applied Linguistics, Vesaliusstraat 2, Leuven, Belgium.

Objective: To provide the linguist and literary researcher with an instrument by which to build a probability model for suffixes and other end-morphemes in language usage (in contrast with reverse dictionaries which are an approximation of language). *Method:* The Dutch frequency-count of De la Court (1937) was divided into three parts: 1) words with a frequency of 25 or more; 2) words with a frequency between 2 and 25; 3) words with a frequency of one. These were then punched, transferred to tapes, sorted in reverse order, and printed (giving the frequency totals per column). To establish the expected frequency of an end-morpheme in a text, it is sufficient to collate the words ending with that morpheme in our list and adapt the number obtained to the size of the text under investigation.

Type of computer: 1) IBM 360/40 and 2) IBM 360/44. *Size of storage:* 1) and 2) 128k. *Language:* 1) and 2) PL/I. Program is available.

Status: Completed 1971.

References: "A Reverse Frequency List of Dutch (with an investigation into the lexical and real frequency of adjectival suffixes), *ITL* (Leuven), 2, 1968, 37-62; "Les mots spéculaires en néerlandais (Mirror-words in Dutch)," *Cahiers de Lexicologie* 17 (Paris, 1970), 64-73; *Inverte frequentielijst van het Nederlands* (Reverse Frequency List of Dutch) (Leuven, 1971).

L427. Inflexional Endings in Contemporary Dutch

Principal investigator: Geerte de Vries, Amanuensis, Institut for germansk filologi, Kobenhavns U., Vesterbrogade 16, 1620 Kobenhavn V, Denmark. *Associate:* R. Lilje-Jensen.

Scope: Einige morphologische untersuchungen des modernen Niederländischen. *Method:* Based on "Woordenlijst der Nederlandse Taal."

Type of computer: 1) IBM 7094, 2) C 360. *Language:* FORTRAN IV.

References: Kopenhagener Beiträge zur germanistischen germanistik, Arbeitsberichte 1971-1972 (Kopenhagen, 1972).

L428. Sprogstatistisk undersokelse (Statistical Methods in Linguistics)

Principal investigator: Hans Olav Egede Larssen, Stipendiat, Lingvistisk institut, U. of Oslo, Blindern, Oslo 3, Norway.

Objective: Investigation of Norwegian "double forms": to prove, if possible, that if two "synonymous" words, one Norwegian Riksmal word and one Nynorsk (New Norwegian) word are used by the same person(s), there will be some differences in meaning or style. *Method:* Text samples from Norwegian newspapers stored on magnetic tape. Using program SPROG, the computer searches for two specified words, all occurrences of which will be printed in context (context = 3-29 lines). The word may be deleted. Context is classified. The hypothesis to be verified is that choice between the two words *does* depend on context class. Statistical tests: Chi-square, Fisher-Irwin. Computer program: Program CHI, author Tore Olafsen.

Type of computer: CDC 3300. *Language:* FORTRAN. *Disks or tapes:* Part of 1 disk, 2 tapes. Program is available. [*CHum,* May 1971]

L429. A Word Frequency Dictionary of *Minnesang des 13.Jahrhunderts*

Principal investigator: Christian Gellinek, Professor, Dept. of Germanic and Slavic Languages and Literatures, College of Arts and Sciences, U. of Florida, Gainesville, FL 32601. *Associate:* Martin Saunders.

Objective: To establish an accurate word count of the words used by the Middle High German poets of thirteenth-century Minnesong. *Method:* Typing, packing, sorting, and alphabetizing the poems in v. Kaus-Kuhn's edition, *Minnesang des 13. Jahrhunderts* (Tübingen, 1962).

Type of computer: IBM 360. *Language:* FORTRAN IV. Program is available.

Status: Completed August 1972.

References: A Word-Frequency Dictionary on IBM 360/65; Christian Gellinek and Heidi Rockwood, *Häufigkeitswörterbuch zu den Denkmälern deutscher Prosa des 11.und 12. Jahrhunderts nach der Ausgabe von Friedrich Wilhelm programmiert von Martin Saunders und Andy Olivenbaum* (Tübingen: Max Niemeyer Verlag, 1973).

L430. Middle Low German Dictionary, Grammar, Word Atlas

Principal investigators: Karl Hyldgaard-Jensen, Professor, Institut for germansk filologi, Kobenhavns U., Vesterbrogade 16, 1620 Kobenhavn V, Denmark; Gisbert Keseling, Professor, U. Marburg, West Germany; Gerhard Cordes, Professor, U. Kiel, West Germany. *Associates:* Flemming Schroller, Henrik Borg Jense, Ute Langenbucher, Jorgen Olsen.

Scope: Investigation of all Middle Low German texts with a view to create a new dictionary for Middle Low German. All inflexional forms are stored, provided with signs for grammatical criteria, dated, and located so as to enable the machine to generate a grammar and a combined atlas of grammatical forms and words.

Type of computer: 1) IBM 7094, 2) C360. *Language:* FORTRAN IV. Program is available.

References: Germanistische Linguistik 2 (1970).

L431. Middle Low German Influence on a Norwegian Dialect of the Sixteenth Century

Principal investigator: Anne Brautaset, Research Assistant, Nordisk Institut, U. of Oslo, Oslo 3, Norway.

Scope: An investigation of the phonemic integration of transferred Middle Low German morphemes in the dialect of Bergen in the sixteenth century. The investigation is based upon a text from the same period. *Method:* 1) Analysis of the graphemic system of the text. 2) Comparison of the model-morphemes with their reproductions in the text.

Type of computer: IBM 360/50. *Language:* STRIL, PL/I. Program is available. [*CHum,* May 1971]

L432. Textkritik

Principal investigator: Günter Kochendörfer, Akad. Rat, Deutsches Seminar, Alte Abteilung, Albert-Ludwigs-U., D78 Freiburg i. Br., Werthmannplatz, Germany.

Scope: Entwicklung eines Verfahrens zur kritischen Edition mittelalterlicher Viel-Handschriften-überlieferungen. Sieben maschinelle Arbeitsgänge: 1) Kollation der (z. B. auf Lochkarten) transkribierten Textzeugen. Ableitung aller durch Lesarten belegten Zeugengruppierungen. 2) Ermittelung einander widersprechender Zeugengruppierungen. 3) und 4) Erstellen von Lesartenlisten zu diesen Konflikten, zur Ermittelung von Kontaminationen bzw. zufälligem Zusammentreffen von Textzeugen in derselben Lesart. 5) und 6) Ermittelung der auf Grund der Gruppierungen unter Berücksichtigung von Kontaminationen und Zufällen möglichen Stemmata. 7) Rekonstruktion des Archetyptestes mit Hilfe ausgewählter Stemmata, einschliesslich kontaminierter oder sonst gestörter Partien. Erstellung eines negativen Lesartenapparats.

Type of computer: 1) UNIVAC 1106, 2) IBM 360. *Size of storage:* 128k. *Language:* 1) ANSI COBOL, 2) PL/I F. Program is available in COBOL version.

References: "Teilautomatisierung der Textkritik bei mittelalterlichen handschriftlichen überlieferungen," *Zeitschrift für deutsche Philologie* 90 (1971), 356-376.

L433. English-Old Church Slavonic Dictionary

Principal investigator: John S. Bross, Associate Professor, Wilmington College, Wilmington, OH 45177. *Associates:* Judith Bares, Robert Dempsey, Thea Iberall, George Klein, Carol Patch.

Objective: Compilation of a new unabridged dictionary of Old Church Slavonic with English glosses and compilation of the first dictionary with Old Church Slavonic as the target language. *Method:* The computer phase consists of keypunching an Old Church Slavonic-English composite glossary and alphabetizing according to the English entry.

Type of computer: IBM 360/50. *Size of storage:* 256k. *Language:* PL/I.

Status: 386 pages of printout arranged in alphabetic order for Old Church Slavonic and for English. There are approximately 14,000 English entries. Output is also available on a standard IBM magnetic tape. Input data available on cards. Project completed September 1972.

L434. Reverse Alphabetized Dictionary of the Slovak Language

Principal investigator: August R. Vavrus, Assistant Professor, Chairman for Russian, Dept. of Modern Languages, Purdue U., West Lafayette, IN 47907. *Associate:* Charles Rieger.

Objective: A reverse reading dictionary for Slovak. *Method:* The five-volume dictionary of the Slovak Academy of Sciences was put on punch cards, which were then sorted into reverse alphabetical order.

Type of computer: IBM 7094. [*CHum,* May 1971]

L435. "Have" and "Be" in the Present Perfect in Norwegian

Principal investigator: Svein Lie, Universitetet I Oslo, Institutt for Nordisk Sprak/Lit, Blindern, Oslo 3, Postboks 1013, Norway.

Scope: Investigation into the use of "have" and "be" as an auxiliary in present perfect in modern Norwegian. Attempts to find differences between the main verbs used in newspapers and by authors. *Method:* Using texts from newspapers and novels stored at Projekt for datamaskinell sprakbehandling, University of Bergen, listing all occurrences of "have" and "be," looking for all their occurrences as an auxiliary with an active main verb, comparison of frequency between the texts and between different main verbs. The programming system STRIL, made in Bergen, has been used.

Type of computer: IBM 360/51. *Language:* PL/I (STRIL). Program is available.

Status: Completed.

L436. Vocabulary and Style in Russian Symbolist and Acmeist Poetry

Principal investigator: Steinar Gil, Slavonic Dept., U. of Oslo, Blindern, Oslo 3, Norway. *Associate:* Ivar Fonnes.

Scope: A study of vocabulary and stylistic features in the poetry of the Russian symbolists and acmeists. At present the investigation includes the poetry of Annenskij, Achmatova, Gumilev, Bal'mont (only "Budem kak solnce") and Blok (only "Gorod"). *Method:* Analysis of concordances and word-frequency lists, search for keywords, investigation of semantic fields, characteristic images and syntactic constructions.

Type of computer: CDC 3300. *Language:* FORTRAN IV. Program is available.

L437. A Concordance to T.S. Eliot, *Complete Poems*

Principal investigator: Peter D. Holand, Faculty of English, U. of Cambridge, Cambridge, England. *Associates:* J.G.B. Heal, R.A. Wisbey.

Objective: A concordance to T.S. Eliot based on the *Complete Poems and Plays of T.S. Eliot* (London: Faber & Faber, 1969).

Type of computer: ICT Atlas II. *Size of storage:* 128k. *Language:* Machine language. *Special equipment:* Multiprogramming system, multi-access. [*CHum,* May 1971]

L438. Gerard Manley Hopkins and Dylan Thomas: A Study in Computational Stylistics

Principal investigator: Patricia Chandler, Assistant Professor, Dept. of Sociology/Anthropology, SUNY, Plattsburgh, NY 12901.

Objective: To arrive at a stylistic comparison between the sonnets of Hopkins and those of Dylan Thomas by a statistical analysis of basic features of syntax and word choice. These data are also compared with those derived from a set of fifteen sonnets by different authors of the late nineteenth century. Significant co-occurrences among the variables and any significant differences in correlations among the three sets of data are found. *Method:* Counting of the thirty variables (which consist of such items as phrase and clause types, parts of speech, instances of alliteration, etc.) and tabulating; use of multiple correlation coefficients; analysis of results.

Type of computer: IBM 360/65. *Size of storage:* 80k. *Language:* FORTRAN IV. *Disks or tapes:* 1 2314 disk. Program is available. [*CHum,* May 1971]

L439. Complete Lexical Frequency Lists and Concordances of Samuel Beckett's Work in English and French

Principal investigator: E. M. Goldstein, Professor, Faculty of Education, U. of Ottawa, Ottawa 2, ON, Canada K1H6K9.

Objective: To produce a complete machine-readable text as source material for the doctorate thesis/manuscript *La négation chez Beckett: forme et signification.*

Type of computer: IBM 360/65. *Size of storage:* 750k, Progr 230k. *Language:* FORTRAN IV H. *Disks or tapes:* 1 2314 disk, 2 9-track tapes. Program is available.

Status: Project in second phase: the computation of 136 coefficients of stylistic correlation with a view to attempting a definition of the concept of style. To be completed in 1975.

References: La recherche littéraire par ordinateur (Un modèle possible d'analyse littéraire et linguistique), ARIES, Edition de recherches linguistiques, Montreal, 1970, being a presentation of a computer analysis of *Acte sans paroles* by Samuel Beckett.

L440. Mathematical Analysis of German Student Poetry

Principal investigator: B.B. Rieger, Lecturer, Institut für mathematisch-empirische Systemforschung (MESY), 65 Theaterstrasse, 51 Aachen, West Germany.

Scope: Description (synchronic-diachronic) of German nineteenth- and twentieth-century minor poems. Quantified analysis of their morphological and lexematic structures to detect characteristic features of literary mass phenomena. *Method:* Development of several sets of statistical characteristics and their dependencies.

Type of computer: CDC 6400. *Size of storage:* 96k. *Language:* FORTRAN IV, CDC 6400 Assembly language. *Special equipment:* Extended core (ECS).

References: Poetae Studiosi: Analysen studentischer Lyrik des 19. u.20. Jhs.—: ein Beitrag zur exaktwissenschaft lichen Erforschung literarischer Massenphänomene (Frankfurt/M: Thesen-Verlag, 1970); "Wort- und Motivkreise als Konstituenten semantischer Umgebungsfelder," *LiLi* 4 (1972); "Literarische Massenphänomene und mengenorientierte Textanalyse," *Das Triviale in Musik, Literatur und Kunst,* ed. H.de la Motte (Frankfurt/M: Klostermann, 1972).

L441. A Concordance to the Poems of Samuel Johnson

Principal investigator: Helen H. Naugle, Associate Professor, Dept. of English, Georgia Institute of Technology, Atlanta, GA 30332.

Objective. A concordance to the *Poems* of Samuel Johnson based on the Clarendon edition with the variations in the Yale edition. *Method:* Alphabetical listings of Johnson's vocabulary (English, Greek, and Latin) with the words in context, the title of the poems in which they appear, and page references.

Type of computer: IBM 360. *Language:* PL/I. Program is available.

Status: Completed May 1973.

References: A Concordance to the Poems of Samuel Johnson (Ithaca, NY: Cornell University Press, 1973).

L442. Non-finite Verb Clause in Shakespeare

Principal investigator: Robert D. Eagleson, Senior Lecturer, Dept. of English, U. of Sydney, Sydney, New South Wales 2006, Australia.

Objective: An examination of the types, grammatical characteristics, interrelationships, frequencies, and stylistic uses of the non-finite verb clause in Shakespeare's plays; a comparison with the practice of his contemporaries is also planned.

Type of computer: IBM 7040. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 6729 tape. Program is available. [*CHum,* May 1971]

L443. Grammatical and Rhetorical Analysis of Sidney's *Astrophel and Stella*

Principal investigator: William Cherubini, Professor, Dept. of English, Cleveland State U., Cleveland, OH 44115.

Scope: Study of Sidney's cycle of 108 Petrarchan sonnets. *Astrophel and Stella* (entire) and two randomly chosen samples of 30 sonnets each from Spenser's *Amoretti* and Shakespeare's sonnet cycle are being analyzed for distribution of word-classes (Fries-Milic code, 24 classes), sentence openings, and various patterns. The same samples are being searched for the occurrence of about 40 rhetorical figures; all tropes are coded, according to the area of experience alluded to ("imagery"), in terms of the scheme developed by C. Spurgeon in *Shakespeare's Imagery* (Cambridge, England, 1936).

Type of computer: IBM 360/50. *Size of storage:* 384k. *Language:* SNOBOL4. Program is available. [*CHum,* May 1971]

L444. Concordance to the Poets of the Dolce Stil Nuovo

Principal investigator: Walter J. Centuori, Associate Professor, Dept. of Romance Languages, U. of Nebraska, Lincoln, NB 68508.

Scope: Complete concordance to the poets of the Dolce Stil Nuovo (Dante excluded). *Method:* The text is initially punched on cards and read to magnetic tape. It is then inverted, sorted, and stored on direct access devices. The concordance is generated using the vocabulary file. The program enables the user to form word-groups, make deletions and concordances of partial or complete text. The final output is in upper/lower case.

Type of computer: IBM 360/65. *Size of storage:* 2000k. *Language:* COBOL. *Disks or tapes:* 3 2314 disks, one 2321; 6 2402 tapes. (The concordance program uses 160k, one tape drive, a bin of the data cell, part of one of the 2314 disks.)

References: M. Marti, *Poeti del dolce stil nuovo* (Firenze: Lemonnier, 1969).

L445. Etude Stylo-Statistique du *Narcisus*

Principal investigators: Serge Lusignan; Bruno Roy, Institut d'Etudes Médiévales, U. de Montréal, C.P. 6128, Montréal 101, PQ, Canada.

Scope: Etude de la spécificité des vocabulaires lyrique et narratif et analyse lexicographique des rimes du *Narcisus* (poéme du XIIe siecle), édité par M.Pelan et N.C.W. Spence (Paris: Les Belles Lettres, 1964). *Method:* 1) L'ensemble du texte, chaque mot étant réduit à la forme du dictionnaire; 2) un lexique et diverses données statistiques; 3) regroupement des rimes selon leur forme; 4) concordance des principaux termes lyriques et narratifs.

Type of computer: CDC 6400. *Language:* FORTRAN IV. [*CHum,* May 1971]

L446. Medieval and Renaissance Scientific Manuscripts of the *Ambrosiana*

Principal investigator: A.L. Gabriel, The Mediaeval Institute, U. of Notre Dame, Notre Dame, IN 46556.

Objective: To catalog (incipits, author, subject, index, date, formats) a hundred Medieval and Renaissance scientific manucripts of the *Ambrosiana.*

Type of computer: UNIVAC 1107. [*CHum,* May 1971]

L447. On-line Information Retrieval in Bibliography: *Gawain and the Green Knight*

Principal investigator: James Joyce, Computer Archives Project, 47 Potomac Street, San Francisco, CA 94117.

Scope: By conventional research methods all printed material (editions, translations, criticism, reviews) regarding *Sir Gawain and the Green Knight* is being recovered. This information is abstracted and entered via a terminal. Scholars may then search for work done on a topic through natural-language-like inquiry commands. *Method:* Files are organized QISAM. Keywords from sets and responses from the total intersection are reported first, from other intersections second, and so forth to a maximum of 30 responses. Tapes are used as back-up for files.

Type of computer: IBM 370/55. *Size of storage:* 2000k. *Language:* PL/1 F. *Disks or tapes:* 2 3330 disks; 2 2400 tapes. *Special equipment:* IBM 2741-compatible terminal.

L448. Encoding and Concordance System for Old English Texts

Principal investigator: Richard L. Venezky, Associate Professor, Dept. of Computer Sciences, U. of Wisconsin, Madison, WI 53706.

Objective: To develop encoding standards and concordance programs for Old English prose, poetry, interlinear glosses, and glossaries, as part of the *Dictionary of Old English* project. *Method:* Conversion of BIBCON (a concordance program derived from CDC KWIC) to the UNIVAC 1108 and to the CDC 6600; testing of encoding techniques in cooperation with the *Dictionary of Old English* editors and associates; design of a file maintenance system for storage and retrieval of concorded data.

Type of computer: 1) UNIVAC 1108, 2) CDC 6600. *Size of storage:* 1) 30k, 2) 30k. *Language:* 1) FORTRAN V, 2) Extended FORTRAN.

Status: Concluded in 1972 for UNIVAC 1108; CDC 6600 version tested in 1971.

References: BIBCON: An 1108 Program for Producing Concordances to Prose, Poetry, and Bibliographic References (Computer Sciences Department Technical Report, University of Wisconsin, 1971); "Computer Processing of Old English Texts," *Computers and Old English Concordances,* ed. A. Cameron, R. Frank, and J. Leyerle (Toronto: University of Toronto Press, 1970).

L449. A Concordance to the Rushworth Matthew

Principal investigator: Richard L. Venezky, Associate Professor, Dept. of Computer Sciences, U. of Wisconsin, Madison, WI 53706.

Objective: To analyze the orthography of Matthew as part of a larger study of the history of English orthography.

Type of computer: CDC 3600. *Language:* BIBCON (a concordance program written mostly in FORTRAN). *Disks or tapes:* 7-track tapes.

Status: Concordance completed 1969. Study of orthography is continuing.

L450. A Concordance to the *Homilies of Aelfric*

Principal investigators: A.F. Cameron, Professor, Centre for Medieval Studies, U. of Toronto, Toronto 181, Canada; Philip H. Smith, Jr., Associate Professor, U. of Waterloo, Waterloo, ON, Canada.

Objective: To produce a KWIC style concordance of J.C. Pope, ed., *The Homilies of Aelfric, a Supplementary Collection* (London: Oxford University Press, 1967-1968) as a test concordance for work on the *Dictionary of Old English.*

Type of computer: IBM 360. *Language:* PL/I. [*CHum,* May 1971]

L451. Analysis of Language by Mathematical Methods

Principal investigators: Y.T. Radday, Senior Lecturer, Dept. of General Studies, Technion, Israel Institute of Technology, Haifa, Israel; D. Wickmann, Lecturer, Institut für Systemforschung, 51 Aachen, Theaterstrasse, West Germany.

Scope: Authorship and authenticity of various Hebrew Biblical texts. (Isaiah completed, L391, *CHum* 5, November 1970, 103; Zacharias completed, Judges in progress.) *Method:* Statistical tests combining a multitude of textual characteristics (see L391).

Type of computer: 1) CDC 6400, 2) Elliott 503, 3) IBM 360/50. *Size of storage:* 1) 32k, 2) 16k, 3) 256k. *Language:* 1) FORTRAN IV, 2) ALGOL, 3) OS. Program is available, ASA compatible.

Status: Completed 1972.

References: D. Wickmann, *Eine mathematisch-statistische Methode zur Untersuchung der Verfasserfrage literarischer Texte* (Köln: Westdeutscher Verlag, 1969); Y.T. Radday, "Isaiah and the Computer," *Computers and the Humanities* 5 (November 1970), 65-73; Y.T. Radday, *The Unity of Isaiah in the Light of Statistical Linguistics* (Hildesheim: Verlag Gerstenberg); Y.T. Radday, *An Analytical Linguistic Keyword-in-Context Concordance of Haggai, Zacharias and Malachi* (computerized) (Wooster, OH: Biblical Research Associates, College of Wooster, 1973).

L452. Concordance to Caesar

Principal investigator: Cordelia M. Birch, 3018 Fourth Avenue, Beaver Falls, PA 15010.

Scope: Concordance for the complete Caesar, based on Oxford Classical Texts, 2 volumes, probably 13,400 cards. *Method:* Keypunching complete text, then alphabetizing.

Type of computer: IBM 1130. *Size of storage:* 8k. *Language:* Basic FORTRAN IV. *Disks or tapes:* 1 removable disk. Program is available.

Status: Keypunched, proofread, and ready for publication.

L453. Recherches sur le latin préclassique et, en particular, sur Plaute

Principal investigator: Albert Maniet, Professeur titulaire, Département de Langues et Linguistique, U. Laval, Québec 10, PQ, Canada. *Associate:* Annette Paquot.

Scope: Fournir une base à des recherches quantitatives sur le domaine du latin préclassique. *Method:* Etablissement d'un fichier analytique de l'oeuvre de Plaute grâce à l'analyse automatique par ordinateur complètée par des indications manuelles.

Type of computer: 1) IBM 1620, 2) IBM 360/201V. *Size of storage:* 1) 20k, 2) 16k. *Language:* 1) SPS, 2) PL/I. *Disks or tapes:* 4 1311 disks. Program is available.

References: A. Paquot, "Détails sur une nouvelle entreprise: index verborum de Plaute avec relevés statistiques," *Revue de l'organisation internationale pour l'étude des langues anciennes par ordinateur* 2 (1967), 35-47; A. Maniet, *Plaute, Lexique inverse, listes grammaticales relevés divers* VIII (New York, Hildesheim, 1969); A. Maniet et A. Paquot, *Plaute, Amphitryon. Index verborum, lexiques inverses, relevés lexicaux et grammaticaux* V (New York, Hildesheim, 1970).

L454. Prosodical Investigation of the Pindaric Odes

Principal investigators: John C. Rouman, Professor; Richard A. Wilson, graduate student; Dept. of Spanish and Classics, U. of New Hampshire, Durham, NH 03824.

Scope: An automatic scansion of the Pindaric Odes including resolution of apparent hiatus by analytical determination of epic correption and digamma, and an analysis of the use of the mute-liquid combinations. Diaresis and caesura are also noted along with an extensive search for three types of bridge: single bridge as represented by Maas's Law, double bridge as represented by Wilamowitz and Knox's Bridge, and bridge as a function of word-beginning as in the case of Giseke and Meyer's Bridge.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV G.

L455. Concordances to the Tragedies of Aeschylus

Principal investigator: Henrik Holmboe, lektor, mag. art., Dept. of Linguistics, U. of Aarhus, DK 8000 Aarhus C, Denmark.

Objective: To make concordances of the tragedies of Aeschylus. *Method:* Texts are being punched on 8-channel paper tape and proofread.

Type of computer: GIER. *Language:* ALGOL. *Disks or tapes:* 1 tape.

Status: The following concordances have been published so far: Aeschylus' *Prometheus Vinctus, Persae, Septem contra Thebas, Supplices, Agamemnon.* The concordances are all available from Akademisk Boghandel, Aarhus, Denmark. Copies of the paper tapes have been delivered to the American Philological Association's data bank. [*CHum,* November 1972]

L456. Concordance to the Poetic Works of Dámaso Alonso

Principal investigator: José Antonio Valverde, Associate Professor and Head, Dept. of Spanish, Acadia U., Wolfville, NS, Canada. *Associate:* Paul A. Fortier.

Objective: To facilitate study of the evolution of stylistic and thematic characteristics of the different periods in *Poemas puros: Poemillas de la cuidad,* plus contemporary poems from the following collections: *Oscura noticia, Hijos de la ira, Hombre y Dios,* and *Gozos de la vista* (the latter collection is unpublished in book form; it is a collection of poems published separately in various literary magazines).

Type of computer: 1) IBM 360/40, 2) IBM 360/65. *Size of storage:* 1) 154k, 2) 1000k. *Language:* 1) and 2) PL/I F. *Disks or tapes:* 1) and 2) 1 2314 disk each; 1) 2 9-channel tapes, 2) 1 9-channel tape. Program is available from Professor Fortier, Dept. of Romance Languages, U. of Manitoba, Winnipeg 19, MB, Canada. [*CHum,* November 1971]

L457. Concordance to Cicero

Principal investigator: Cordelia M. Birch, 3018 Fourth Avenue, Beaver Falls, PA 15010.

Method: Texts have been keypunched using a curved capital U. The eventual plan is a concordance to Cicero.

Type of computer: IBM 1130. *Size of storage:* 8k. *Language:* FORTRAN IV. *Disks or tapes:* 1 removable disk. Program is available.

Status: All the letters (OCT), vols. 1, 2, 5, and 6 of the orations (OCT), and the following from the Loeb series: *De Inventione, De Amicitia, De Divinatione,* and *De Senectute* are keypunched.

L458. Associative Thesaurus

Principal investigator: George R. Kiss, Scientific Staff Member, Speech and Communication Unit, Medical Research Council, 31 Buccleuch Place, Edinburgh 8, Scotland. *Associates:* Christine Armstrong, Keith Farvis, Robert Milroy, James Piper.

Scope: Collection of data and computer processing of a large "Associative Thesaurus." Structural analysis of the resulting large network of associative connections for semantic fields, using various methods, from graph theory and numerical taxonomy. Data are to provide baseline information for empirical work in psycholinguistics, content analysis, education, and information retrieval. *Method:* Each of the 8400 stimulus words is presented to 100 different subjects who give a single response word—the first word which comes to mind. Subjects are a systematic sample of British undergraduate students, stratified according to geographical origin and faculty of studies.

Type of computer: 1) IBM 370/155, 2) PDP-10, 3) ICL System 4. *Size of storage:* 1) 512k bytes, 2) 32k words, 3) 450k bytes. *Language:* 1) COBOL, FORTRAN, 2) POP-2, 3) IMP, FORTRAN. *Disks or tapes:* 1) 1 3330 (shared) disk, 2) RPO2 (shared) disk, 1 700M byte disk; 1) 2 9-track tapes, 3) 2 9-track tapes. *Special equipment:* Datagraphix Computer Output Microfilm. Program is available.

Status: Programs have been written for extracting the associative environment of any word, for calculating measures of connectivity between any node pair in the graph, for calculating a set of local coordinates for each word in an environment and applying clustering techniques, and for deriving "global" clusters on the basis of connectivity.

References: G.R. Kiss, C. Armstrong, R. Milroy, *An Associative Thesaurus of English* (microfilm version), (East Ardsley, Wakefield, Yorkshire: EP Microforms, Ltd, 1972); G.R. Kiss et al., "An Associative Thesaurus of English and Its Computer Analysis," *The Computer and Literary Studies,* ed. A.J. Aitken, R.W. Bailey, and N. Hamilton-Smith (Edinburgh: The University Press, 1973); G.R. Kiss et al., "An Introduction to the Microfilm version of the Associative

Thesaurus," MRC Speech and Communication Unit, 1972 (mimeo); K. Farvis, ed., "Index of Stimulus Words Used in the Associative Thesaurus," MRC Speech and Communication Unit, 1972 (mimeo); G.R. Kiss, "An Associative Thesaurus and Its Structure," MRC Speech and Communication Unit, 1973 (mimeo).

L459. Testing of Transformational Grammars by Computer

Principal investigator: Marguerite Van der Borght, Senior Teaching Fellow, Dept. of French, Monash U., Clayton, Victoria 3168, Australia.

Scope: Construction of a French transformational grammar. *Method:* Rules of grammar are written within the transformational framework. These are then codified and tested with the help of a suitable computer program written in FORTRAN. The rules are progressively enlarged by repeated application of this method.

*Type of computer:*CDC 3200. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 2 disks, 4 tapes. [*CHum,* November 1971]

L460. The Positioning of Preverbs in American English

Principal investigator: Sven Jacobson, Docent, Dept. of English, Stockholm U., Strandprom. 25B, S-13011 Saltsjö-Duvnäs, Sweden.

Method: Statistical analysis of quantified factors.

Type of computer: 1) CDC 3690, 2) CDC 8090. Program is available. [*CHum,* November 1971]

L461. Grammatical Analysis of the Standard Corpus of Present-Day American English

Principal investigator: Henry Kucera, Professor, Dept. of Linguistics and Slavic Languages, Brown U., Box E, Providence, RI 02912. *Associates:* W. Nelson Francis, Gerald M. Rubin.

Scope: The grammatical annotation ("tagging") of all the words in the one-million-word Standard Corpus of Present-Day American English. Labeled bracketing of the principal syntactic structures. Eventually, the development of a theory of linguistic performance and of grammatical style based on the results. *Method:* 1) Automatic tagging. 2) Disambiguation of ambiguous tags through context rules. 3) Adaptation of development of phrase structure parsing procedures. 4) Manual disambiguation of multiple parsings. 5) Retrieval of phrase structure rules underlying the bracketed structures. 6) Retrieval of relevant statistical information pertaining to linguistic performance and style. [Note: Only steps 1 and 2 have been attempted so far, with 3 in initial stages, and 6 tried on limited portions of the Corpus.]

Type of computer: IBM 360/67. *Size of storage:* 350k. *Language:* PL/I F. *Disks or tapes:* 1 tape or 1 disk. "Tagging" program only is available.

References: Henry Kucera and W. Nelson Francis, *Computational Analysis of Present-Day American English* (Providence: Brown University Press, 1967); Barbara B. Greene and Gerald M. Rubin, *Automatic Grammatical Tagging of English,* Dept. of Linguistics, Brown University, 1971. [*CHum,* November 1971]

L462. Automatic Grammatical Tagging of English

Principal investigators: Gerald M. Rubin, Programmer; Barbara B. Greene, Dept. of Linguistics, Brown U., Providence, RI 02912.

Objective: To assign grammatical word-class annotations to English words quickly and without extensive dictionary look-up as a first step in a syntactic system. *Method:* Words are assigned multiple tags on the basis of ending inspection—90 word class distinctions are determined by examining 450 word endings and a dictionary of 2800 exceptions to ending rules. Ambiguities are resolved by context examination.

Type of computer: IBM S/360 OS. *Size of storage:* 256k. *Language:* PL/I F. *Disks or tapes:* 1 disk, but tape can be used instead. Program is available. It is reproduced in publication described below.

References: B.B. Greene and G.M. Rubin, *Automatic Grammatical Tagging of English,* Department of Linguistics, Brown University, 1971. [*CHum,* November 1971]

L463. A Computer-Generated Gothic Word Index and Reverse Word List

Principal investigator: Randall L. Jones, Assistant Professor, Dept. of Modern Languages, Cornell U., Ithaca, NY 14850. *Associates:* Frans van Coetsem, Philip H. Smith, Jr., F. de Tollenaere.

Objective: To generate a complete word index and reverse word list of the Streitberg Gothic Bible. The word index will contain all words in the text together with a reference code to indicate where they occur. The reverse word list will contain all words (types) listed in reverse alphabetic order. *Method:* The magnetic tape of the Gothic Bible has been prepared under the direction of Philip H. Smith, Jr. It has been corrected by three separate groups of readers at Cornell University and at Leiden, The Netherlands. The word tokens were isolated and tagged with a reference code identifying their location in the text. The list was sorted and printed out. A second list was prepared by reversing the characters in each word and sorting the reversed list. After sorting, the list was again reversed and printed out.

Type of computer: IBM 360/65. *Size of storage:* 512k. *Language:* PL/I, SNOBOL4. *Disks or tapes:* 1 2314 disk; 2 9-channel tapes. *Special equipment:* Print train which includes special Germanic characters. [*CHum,* November 1971]

L464. Gothic Grapheme Clusters

Principal investigator: Randall L. Jones, Assistant Professor, Dept. of Modern Languages, Cornell U., Ithaca, NY 14850. *Associates:* Frans van Coetsem, Philip H. Smith, Jr., F. de Tollenaere.

Objective: To scan the Streitberg Gothic Bible for occurrences of words containing a cluster of two or more consonant or vowel graphemes. *Method:* Using the magnetic tape prepared by Philip H. Smith, Jr. (see L463) the program scanned the text for all occurrences of words containing consonant or vowel grapheme clusters. The words containing such clusters were isolated and sorted. Any word containing more than one cluster was duplicated as many times as needed. The words were arranged according to initial, medial, and final clusters, and within each group according to whether the cluster contained two, three, or four characters. Vowels were kept separate from consonants.

Type of computer: IBM 360/65. *Size of storage:* 512k. *Language:* PL/I, SNOBOL4. *Disks or tapes:* 2 9-channel tapes. *Special equipment:* Print train which includes special Germanic characters. [*CHum,* November 1971]

L465. *The Computer Bible* and International Library of Ancient Texts

Principal investigators: J. Arthur Baird, Professor, Dept. of Religion, College of Wooster, Wooster, OH 44691; David Noel Freedman, Professor, Dept. of Religious Studies, U. of Michigan, Ann Arbor, MI 48104. *Associates:* A.Q. Morton, Peter Morris, Yehuda T.Radday, John Hurd, Peder Borgen.

Objective: To produce a series of computer-generated concordances of Biblical and related ancient texts. These are critical concordances for the purpose of content, grammatical, linguistic, morphological, semantic, and historical study. They will be published and made generally available. *Method:* These concordances are being done by various scholars around the world, in the USA and Canada, Scotland, England, Wales, Norway, Israel. They are being edited and published under the auspices of Biblical Research Associates, Wooster, Ohio. J. Arthur Baird and David Noel Freedman are the editors. Method of use is mostly pattern recognition of the various categories.

Type of computer: CDC 6400. *Size of storage:* 438k. *Language:* SNOBOL4, FORTRAN IV. *Disks or tapes:* 10 tapes. Program is available.

Status: Three volumes of the *Computer Bible* have been published: Vol. 1, The Synoptic Gospels; Vol. 2, Isaiah; Vol. 3, Johannine Letters. Future volumes in process: The Hexateuch, Gospel of John, and Minor Prophets. The International Concordance Library is a series of computer-generated KWIC concordances of ancient texts. In process: The Apostolic Fathers, Philo of Alexandria, The Gospel of Thomas, The Oxyrhynchus papyri. Proposed texts: The Apocryphal New Testament, The Dead Sea Scrolls, Josephus, and Eusebius.

References: Audience Criticism and the Historical Jesus (Westminster, 1969). Reviewed by John W. Ellison, *Computers and the Humanities* 4, 3 (1970), 199-205. [*CHum,* May 1972]

L466. Computer Analysis of Printed Chinese Characters

Principal investigator: William Stallings, Principal Scientist, Mail Station 425C, Honeywell Information Systems, 200 Smith Street, Waltham, MA 02154.

Objective: To recognize printed Chinese characters by computer. *Method:* Recognition of a character is based on an analytical description of its pictorial structure. The input to the program is a black-white matrix of points depicting a single character. The program produces a description of the character on two levels: 1) the internal structure of each connected part of the character, 2) the arrangement of two dimensions of the connected parts. A unique numeric code is generated which serves to identify the character.

Type of computer: PDP-9. *Size of storage:* 32k. *Language:* FORTRAN IV. *Special equipment:* Flying-spot scanner. Program is available.

References: "Computer Description and Recognition of Printed Chinese Characters," *Proceedings of the 1972 Spring Joint Computer Conference,* pp. 1015-1025.

L467. Machine Translation of Mandarin Chinese

Principal investigator: Walter J. Stutzman, Dept. of Linguistics, Yale U., New Haven, CT 06520.

Scope: Investigation of Mandarin grammar construction and representation. *Method:* Current models based on Chao's *Mandarin Primer* and *Grammar of Spoken Chinese.* Specification of language using Floyd-Evans productions.

Type of computer: PDP-10. *Size of storage:* 64k. *Language:* SNOBOL4. *Disks or tapes:* 1 RPo2 disk; 2 TU2o tapes.

Status: Ended August 1972.

L468. Concordance of the Complete Theater of Ramón del Valle-Inclán

Principal investigator: Albert H. LeMay, Assistant Professor, Dept. of Romance Languages, Colgate U., Hamilton, NY 13346. *Associate:* Nicolas Vituli.

Objective: To compile a concordance of 21 plays written in both verse and prose by the twentieth-century Spanish author Ramón del Valle-Inclán (1866-1936). The plays were published between 1899 and 1930. The concordance is to include five dictionaries: 1) list of all words according to frequency, 2) alphabetical list of all words including frequency, 3) list of keywords and themes given in context, 4) dictionary of prepositions, conjunctions, etc., 5) dictionary of characters and their location in texts. *Method:* The entire texts including stage directions and character appearances and exits will be keypunched. All words will be included in frequency and alphabetical listings although certain items (conjunctions, prepositions, etc.) will be omitted in the dictionary dealing with keywords and themes. They will be compiled in a special section of their own.

Type of computer: PDP-10. *Size of storage:* 64k. *Language:* FORTRAN, SNOBOL. *Disks or tapes:* 1 DEC RPoz disk; 1 mag., 4 DEC tapes. Program is available.

L469. Contextual Concordance to Valerius Maximus

Principal investigator: C.J. Carter, Lecturer, Dept. of Latin, U. of St. Andrews, St. Andrews, Fife, Scotland. *Associates:* A.E. Cairns, P.A. George.

Scope: A Concordance to Valerius Maximus on the lines of D.W. Packard's concordance to Livy (Harvard, 1968), but with consistent orthography and the additional listing of all questions, exclamations, speech/dialogue, parentheses, textual cruces, conjectural additions, deletions, and supplements. *Method:* Alphabetical sort of keyword to 16 places, by alphabetical sort of word following keyword to 8 places. Repetitions of keyword and word following are listed in order of occurrence. Simple word lists and word counts of each book are also produced.

Type of computer: IBM 360/44, *Size of storage:* 120k bytes. *Language:* FORTRAN IV (44 PS and OS). *Disks or tapes:* 3 2311 disks; 2 9-track tapes. *Special equipment:* 30" digital plotter, 60-character set to be replaced by 120 set with lower case. Program is available.

Status: Completed April 1973. Production of the complete concordance is now being followed by a number of stylistic studies of the text. P.A. George is also using this and other material for his program "Phonemic Analysis of Classical Latin" (see L470). Tape of machine-readable text available from A.Ph.A. data bank at Dartmouth, NH, USA, or per the compiler at St. Andrews University, Scotland.

L470. A Phonemic Analysis of Classical Latin

Principal investigator: P.A. George, Lecturer, Dept. of Latin, U. of St. Andrews, St. Andrews, Fife, Scotland. *Associates:* A.E. Cairns, A.S. Gratwick.

Objective: To establish 1) what phonetic features are functionally distinctive in Classical Latin, 2) what simultaneous combinations of distinctive features are functional (i.e., what the phonemes of Classical Latin are). *Method:* The program is designed to search for functional oppositions at the phonetic level both in texts and in morphemic inventories. It is hoped that a program for quantifying the functional load of any given opposition will be added in the near future.

Type of computer: IBM 360/44. *Size of storage:* 120k bytes. *Language:* FORTRAN IV (44 PS and OS). *Disks or tapes:* 3 2311 disks; 2 9-track tapes. *Special equipment:* 30" digital plotter, 60-character set to be replaced by 120 set with lower case. [*CHum,* November 1971]

L471. *The London Stage* Information Bank

Principal investigators: George Winchester Stone, Jr., Dean, New York U. Libraries, 10 Washington Square Place, New York, NY 10003; Ben Schneider, Jr., Professor, Dept. of English, Lawrence U., Appleton, WI 54911.

Scope: Edit the machine-readable corpus (*The London Stage 1660-1880)* and expand the data base to include bibliographies and plays of the period to provide an information service to scholars in the field. *Method:* Convert new material, refine retrieval system, develop editing system.

Type of computer: 1) IBM 360/44, 2) PDP-11/20. *Size of storage:* 128k. *Language:* 1) PL/I, 2) Basic. *Special equipment:* Intelligent CRT terminal.

References: "The London Stage Information Bank," *Restoration and 18th-Century Theatre Research* 9, 2 (November 1970), 56-58; "The London Stage Information Bank," *Computers and the Humanities* 4 (1971), 209-214; "Optical Scanning as a Method of Input," *The Computer and Literary Studies,* ed. A.J. Aitken, R.W. Bailey, and N. Hamilton-Smith (Edinburgh: The University Press, 1973); "Analysis of a Data Base for Information Retrieval," ibid. [*CHum,* November 1972]

L472. Morphemic Analysis of Russian

Principal investigators: Zdenik F. Oliverius, Professor, Dept. of Russian; Ron Addie, student programming assistant, Dept. of Mathematics, Monash U., Clayton, Victoria 3168, Australia.

Scope: The purpose of morphemic analysis, starting from the Steinfeldt frequency dictionary as its datum, is to establish an inventory of morphemes and to reveal the hierarchy and valency of morphemes. *Method:* The computer is used in only one part of the project, permitting a count of morpheme frequencies and their distribution.

Type of computer: CDC 3200. *Size of storage:* 32k. *Language:* FORTRAN 32. *Disks or tapes:* 2 disk packs; 4 tapes. Program is available in part. [*CHum,* November 1971]

L473. Computational Analysis of Romance Syllabic Structures

Principal investigator: Guntram A. Plangg, Visiting Assistant Professor, Dept. of Romance Languages, Ohio State U., Columbus, OH 43210. *Associates:* Thomas G. Whitney, Martha L. Knorre.

Scope: To analyze a base of syllabic structures for the typological affinity of Romance languages, particularly the quantitative relation of Western Romance dialects clearly defined in time and space. *Method:* Analysis of all the different phonemic (and prosodic) elements which constitute the syllable, especially the distribution and frequency. Units, groups, inventory, and sequences of minimal segments will be compared.

Type of computer: IBM 360/75. *Language:* SNOBOL4, PL/I. Program is available for Old French, Old Provençal, and Latin, from Thomas G. Whitney.

References: G.A. Plangg, "Zum Sprachtypus des Ladinischen und seiner Nachbarn," *Der Schlern* 43 (1969), 159-176. [*CHum,* November 1971]

L474. Punctuation Patterns in Racine

Principal investigator: Ivan Barko, Professor, Dept. of French, Monash U., Clayton, Victoria 3168, Australia. *Associates:* G. Millar, J. Rivers, H. McCulloch.

Objective: To analyze punctuation patterns in the manuscripts of Racine's correspondence. The ultimate purpose of the project is to lead to a new critical edition of all those works of the dramatist for which no manuscripts have come down to us. *Method:* Potential punctuation locations are coded according to syntactic and rhythmic criteria, and the presence or absence of each is then recorded. A concordance based on these is expected to reveal a pattern in Racine's use. It is hoped that further comparison between the punctation of the letters and that of the very small corpus of verse manuscripts will establish the general validity of our conclusions. Yet another comparison between Racine's punctuation habits and the punctuation patterns in the different editions published in the author's lifetime may validate one or several of these editions as reproducing Racine's own punctuation.

Type of computer: CDC 3200. *Size of storage:* 32k. *Language:* FORTRAN. *Disks or tapes:* 2 4-million-character disks (CDC); 4 tapes. *Special equipment:* 30" incremental plotter. Program is available.

References: G. Silva and C. Bellamy, *Some Procedures and Programs in Processing Language Data* (Clayton: Monash University, 1969).

L475. A Concordance of Flaubert's *Madame Bovary*

Principal investigators: Pierre H. Dubé, Assistant Professor, Dept. of Classics and Romance Languages, U. of Waterloo, Waterloo, ON, Canada; Charles Carlut, Professor, Dept. of Romance Languages, Ohio State U., Columbus, OH 43210.

Objective: A thematical study of *Madame Bovary*. *Method:* KWIC concordance. Variants of the text are included and flagged so that an evolution in Flaubert's creative thought processes may be readily analyzed. Flags will also be included to indicate whether a given word belongs to a narrative or dialogue mode. By studying keywords, in relation to the words surrounding them, we hope to discover new themes and/or amplify old ones.

Type of computer: IBM 360/75. *Size of storage:* 125k. *Language:* PL/I. *Disks or tapes:* 3 9-track tapes. *Special equipment:* Modified TN print train or ALA. Program is available.

L476. Investigation in Talmudic Lexicography

Principal investigator: Ben Zion Wacholder, Professor, Dept. of Talmud and Rabbinics, Hebrew Union College, Jewish Institute of Religion, 3101 Clifton Avenue, Cincinnati, OH 45220. *Associates:* E.A. Goldman, H.D. Uriel Smith, Roy D. Tanenbaum.

Objective: The generation of an English-Talmudic (Hebrew, Aramaic, et al.) dictionary, from which it will also be possible to develop a Hebrew and Aramaic thesaurus, sub-dictionaries of foreign words and special derivations, and other investigative resources. *Method:* The text, *Dictionary of the Targumin, the Talmus Babli and Yerushalmi, and the Midrashic Literature* by Marcus Jastrow, is pre-edited with grammatical and derivational notations, keypunched using the Computer-Compatible Semitic Alphabet developed at Hebrew Union College, and stored on magnetic tape. A manual editing follows, from which an expanded version of the original dictionary is produced. Through algorithms designed to eliminate most homographs, the basic data will be mechanically inverted: the significant English words in the Jastrow definitions now alphabetically displayed as headers, the Semitic entry words and their associated descriptors now constituting the definitions.

Type of computer: IBM 360/65. *Size of storage:* 120k. *Language:* PL/I F. *Disks or tapes:* 2 9-track tapes. Jastrow tape and pre-editing programs are available.

References: E.A. Goldman, H.D. Uriel Smith, R.D. Tanenbaum, "Transliteration and a 'Computer-Compatible' Semitic Alphabet," *Hebrew Union College Annual* (1971). [*CHum*, November 1971]

L477. Automatic Processing of the *Corpus of Latin Inscriptions*

Principal investigators: Eugène Chouraqui, Michel Janon, Jacques Virbel, Centre d'Analyse Documentaire pour l'Archéologie, Centre National de la Recherche Scientifique, Chemin Joseph Aiguier, Marseille 9, France; Paul Corbier, Paul-Albert Février, Faculté des Lettres, Institut d'Archéologie Méditerranée, Chemin des Fenouillères, Aix-en-Provence, France.

Scope: Automatic edition of Latin inscriptions (*C.I.L.*, and its updating), production of indexes and tables (including a vocabulary index); elaboration of a documentary system. Experiment on a corpus of 800 inscriptions from and about veterans in Roman Africa (*C.I.L.*, vol.8). *Method:* Elaboration of codes for the description of the texts (reading, interpretation), the language (morphology, linguistic particulars), the context (location, material, size, support, ornamentation, conservation, etc.), the content (semantic categorizations).

Type of computer: UNIVAC 1108. *Size of storage:* 64k words. *Language:* FORTRAN V, SLEUTH, SNOBOL. *Disks or tapes:* 2-3 tapes. *Special equipment:* 5 drums FH432 on extended central memory. Program is available.

Status: Completed 1972.

References: Michel Janon and Jacques Virbel, "Problèmes posés par la constitution d'un fichier automatisé des inscriptions latines," Colloque National sur les Banques de Données Archéologiques, C.N.R.S., Marseille, June 12-14, 1972; Jacques Virbel, "A propos de l'édition et de la consultation automatique du Corpus des Inscriptions Latines: Problèmes méthodologiques posées par l'analyse formelle des Textes" (à paraître dans *Archéologie et Calculateurs*, 1973); "Un système d'exploitation automatique du Corpus des Inscriptions Latines (SYCIL): Mode d'Emploi" (Rapport CADA-IAM, 1972).

L478. The Frequency of Syntactical Patterns in Present-Day Written Australian English

Principal investigator: Ralph D. Beebe, Dept. of Linguistics, Monash U., Clayton, Victoria 3168, Australia.

Scope: The project will embrace several genres of present-day written Australian English drawn from such sources as newspaper editorials, novels, and textbooks. *Method:* The syntactic structure of a sentence is first hand-generated in the form of a dependency-grammar tree. Using principles analogous to those employed by organic chemists in converting two-dimensional diagrams of molecular structure into linearizations, the tree structure is then converted to a linear string of symbols. Many such strings, fed to the computer in punched-card form, are then analyzed for their syntactic patterns, the frequencies of which are reported.

Type of computer: Burroughs B5500. *Size of storage:* 32k (time-shared). *Language:* ALGOL. *Disks or tapes:* 1 fixed head 19.2-million-character (1/2 mod only = 4.8M char.) disk; 6 7-track tapes. Program is available.

References: Linguistic Communications, Vol. 3, Monash University, Clayton, Victoria 3168, Australia (forthcoming). See also G57. [*CHum,* November 1971]

L479. MISTIC (*M*achine-*I*mplemented *S*yn*T*hesis of *I*maginative *C*ommunication)

Principal investigator: Mark Underwood, U. of Arizona, 2708 E. Mabel Street, Tucson, AZ 85716. *Associate:* Henry Tucker.

Scope: We are working on a semantic component to the Friedman transformational grammar program via some feasibility studies while looking for additional funding. Correspondence should be addressed to Interdisciplinary Research and Information Systems, Inc., 1112 N. Jones, Tucson, AZ 85716.

Type of computer: CDC 6400. *Size of storage:* 55k. *Language:* FORTRAN COMPASS (CDC Assembly). *Disks or tapes:* 2 disks; 5 7-track tapes. Program is available for MISTA, the crude "frame" poetry.

L480. Computer Archive of Modern English Texts

Principal investigator: Geoffrey N. Leech, Reader, Dept. of English, U. of Lancaster, Bailrigg, Lancaster, U.K. *Associates:* N.L. Fairclough, Bryan Higman, Rosemary Leonard, Helen Goodluck.

Objective: To compile an archive of Modern English on magnetic tape consisting of 1) the 1,000,000-word corpus of present-day American English compiled at Brown University, 2) a matching 1,000,000-word corpus of British English, 3) the spoken and manuscript materials collected and transcribed in the Survey of English Usage, University College, London.

Type of computer: ICL 1909. *Size of storage:* 32k, 24-bit 2-microsecond core. *Language:* ALGOL, PLAN. *Disks or tapes:* 3 disks, 3 tapes. [*CHum,* May 1972]

L481. Semantic Foundations for Question-Answering Systems

Principal investigator: John Olney, Researcher, Center for Computer-Based Behavioral Studies, UCLA, Consultant, System Development Corporation, 2500 Colorado Avenue, Santa Monica, CA 90406. *Associates:* Scott Munson, Donald Ramsey, Sylvia Tidwell, Judge Schonfeld, Rollin Weeks.

Objective: To obtain, from complete machine-readable transcripts of *Webster's Seventh New Collegiate Dictionary* (W7) and its pocket abridgment (MPD), data which can be used to improve natural-language processors, especially as applied to medical information problems. The focus is on deriving data from W7 and MPD definitions which can throw light on the semantic structure of the English vocabulary as a whole. *Method:* 1) Parse the transcribed dictionary entries into pronunciations, etymologies, definitions, etc., on the basis of typography; 2) extract from the parsed transcripts data sets useful for stylistic analysis of the dictionaries; 3) obtain semantic data sets, starting with entry parts which are well structured (e.g., by recording the spelling differences between main-entry forms and respective run-ons, sorting on these differences, and merging in W7 definitions of MPD run-ons, extensive semantic [and morphographemic] information about currently productive English suffixes can be obtained); 4) via manual semantic disambiguation of the words functioning as syntactic heads of their respective definitions, extract explicit generic hierarchies from W7 and MPD. Steps 1 and 2 have been completed, 3 is nearing completion, and 4 is well under way. Our suffixal data is now being used for detection of synonymitive relations among phrases recurring in medical titles.

Type of computer: 1) IBM 370/145, 2) PDP-10. *Size of storage:* 1) and 2) 90-350k. *Language:* 1) BAL, FORTRAN, PL/I; 2) PL/I. *Disks or tapes:* 1) and 2) 120 9-track tapes; 1) 3 2314 disks. Many programs, about 50 data sets, available for a nominal copying charge from John Olney.

References: R. Reichert, J.C. Olney, and J.Paris, *Two Dictionary Transcripts and Programs for Processing Them,* Vol. 1, System Development Corporation Document TM-3978/001/00, June 1969. (Available from PDLI; see *Finite String,* September 1970, 2-4.) J.C. Olney and D. Ramsey, "From Machine-Readable Dictionaries to a Lexicon Tester: Progress, Plans, and an Offer," *Computer Studies in the Humanities and Verbal Behavior.* [*CHum,* May 1972]

L482. A Stylistic Analysis of the *Chanson de Roland*

Principal investigator: John R. Allen, Associate Professor, Dept. of Romance Languages, U. of Manitoba, Winnipeg, MB R3T 2N2, Canada.

Objective: To establish whether any significant stylistic traits can be observed in the oldest sections of the *Chanson de Roland* which would suggest the structure of or partial content of earlier versions of the legend. To shed more light on the controversy over whether the Oxford manuscript is the work of a single poet. To gather evidence which would help resolve questions over the authenticity of the Baligant episode. *Method:* Establishment of two versions of the text, one based on the manuscript before the revisor's corrections, the other with his changes. Each word of both texts is then replaced by its lemma, with homographs distinguished and contractions broken down into component parts. The different laisses (strophes based on assonance) of both versions will be then separated, insofar as possible, according to age, based on assonance. Those sections that are oldest will receive a separate analysis. Stylistic features to be considered in the analyses are: word length, sentence length, entropy, transition frequencies, and semantic classification of nouns. Use of selected forms may also be a criterion for determining homogeneity of various sections.

Type of computer: IBM 360/65. *Size of storage:* 768k. *Language:* BASIC 6 and PL/I. *Disks or tapes:* 27 2314 disk packs; 3 9-track tape handlers, 1 7-track tape handler. Program is available.

L483. Linguistic/Literary Stylistics Analyses

Principal investigator: Edward A. Kline, Associate Professor, Dept. of English, U. of Notre Dame, Notre Dame, IN 46556.

Scope: Analysis of literary texts for assonance, alliteration, end rhyme, word classes, affixes, collocations, syntax patterns, punctuation, character and word recognition capacities. *Method:* Time-sharing mode is used. Console used to input texts. Analysis programs are canned. Outputs of frequency lists and/or maps of the text are provided on the console.

Type of computer: IBM 360/50. *Size of storage:* 32k. *Language:* APL. *Disks or tapes:* 2 tapes. *Special equipment:* IBM 2750 Console. Program is available.

L484. Middle English Dialectology Via Orthography

Principal investigator: Edward A. Kline, Associate Professor, Dept. of English, U. of Notre Dame, Notre Dame, IN 46556.

Objective: To determine the dialect and date of literary manuscripts from the twelfth through fifteenth centuries. For testing purposes, the texts, as found in F. Moss's *Handbook of Middle English,* serve as the data base. *Method:* Have the computer make up frequency lists of those items of an orthographical nature that can be used for determining dialects of Middle English.

Type of computer: IBM 370/155. *Size of storage:* 128k. *Language:* PL/I. *Disks or tapes:* 4 minitapes. Program is available.

L485. A Theory of Poetry and Its Computational Verification

Principal investigator: James Joyce, Computer Archives Project, 47 Potomac Street, San Francisco, CA 94117.

Scope: 1) Derivation of theory of poetry as organized pause. 2) Implementation of computer programs, to score juncture types in edited texts, both prose and poetry, and identify periodicity of pause for each piece. 3) Statistical interference to show a) prose pauses have longer periodicity than pauses in poetry; b) the first reliable periodicity of pause in poetry is the line; c) the first prominent multiple of periodicity for line is stanza. *Method:* Ross' syllable counter routine (modified for speed) used to give syllable number after which juncture occurs. Junctures are scored in prose and poetry using Trager and Smith levels of juncture.

Type of computer: IBM 370/155. *Size of storage:* 2000k. *Language:* PL/I F. *Disks or tapes:* 1 disk. *Special equipment:* Teleprocessing with text-editor.

L486. Un système de recouvrement des données pour l'analyse des attitudes des personnages de roman

Principal investigator: Florence E. DeHart, Associate Professor, School of Library and Information, U. of Western Ontario, London 72, ON, Canada. *Associate:* William Henry Hudgins.

Scope: Préparer et mettre à l'essai un système de recouvrement des données à partir d'un ordinateur, pour l'analyse des personnages de roman, qui peut être appliqué d'une façon constante et méthodique dans la vérification d'un grand nombre d'hypotheses en critique littéraire française. Les préoccupations philosophiques, psychologiques, et morales des

personnages dans six des romans de Flaubert, Gide, Sartre, Bernanos, Mauriac, et Camus seront analysées par le système en ce qui concerne les déterminants de la formation, de l'éveil et du changement d'attitudes par rapport à la sorte de fonction servie par l'attitude. Les particularités du style de la sémantique et de la syntaxe employées dans l'expression et la description des préoçcupations seront identifiées. *Method:* L'analyse du contenu sera appliquée aux six romans de deux façons: 1) Un thésaurus ou idioglossaire des symboles qui se trouvent dans les attitudes des personnages, les fonctions associées avec l'attitude, et les déterminants de l'attitude sera compilé afin d'aider à l'identification des images et des thèmes. 2) Un système non-traditionnel de recouvrement des données basé sur les concepts sera développé pour codifier les discours pertinents et les commentaires de l'auteur centrés autour des concepts qui sont exprimés dans un vocabulaire défiant toute présentation sous forme de thésaurus. *Madame Bovary* de Flaubert sera marqué et codifié sur bande magnétique dans le but d'essayer les programmes de l'ordinateur qui seront développés pour mettre en opération le système de recouvrement des données.

Type of computer: PDP-10. [*CHum*, May 1972]

L487. Thematic Analysis of Four Twentieth-Century French Novels

Principal investigators: Paul A. Fortier, Assistant Professor, Dept. of Romance Languages and Literatures; J. Colin McConnell, Dept. of Computer Science, U. of Manitoba, Winnipeg 19, MB, Canada. *Associates:* Carolyn Harynuk, Dolores Ferraton, Nancy Latoki, Michel Duclos, Lynne Friesen.

Scope: A pilot project is to set up a system for thematic analysis of novels. Robbe-Grillet's *La Jalousie,* Gide's *L'Immoraliste,* Sartre's *La Nausée,* and Bernanos' *Monsieur Ouine* are being studied at this stage. It is hoped that the resulting system will be usable with any prose fiction text in French, perhaps in other languages as well. *Method:* The complete system comprises four subsystems each of which is modular, consisting of several programs: 1) concord texts, 2) tag machine-readable concordance with part of speech and basic form; convert it to a search file of variable length records, permitting rapid searches; 3) gather lists from the eleven French synonym dictionaries, collage, alphabetize, remove duplicates and extraneous words, thus creating theme lists; 4) compare theme lists to search files or updated concordance to produce graphs, synoptic tables, and specialized concordances. Using the output from 4, the scholar analyzes the text.

Type of computer: IBM 360/65. *Size of storage:* 180k. *Language:* PL/I F. *Disks or tapes:* 1 disk, 2 tapes. *Special equipment:* French print chain, containing upper/lower case, accents, all punctuation. Concordance program is available.

References: "A System for Computer-Aided Thematic Analysis of French Texts" (mimeographed project description available from principal investigators); "Computer-aided Thematic Analysis of French Prose Fiction," *The Computer and Literary Studies,* ed. A.J. Aitken, R.W. Bailey, and N. Hamilton-Smith (Edinburgh: Edinburgh University Press, 1973).

L488. A Text-Oriented Study of the Cree Language

Principal investigator: H. Christoph Wolfart, Associate Professor of Linguistics, Dept. of Anthropology, U. of Manitoba, Winnipeg 19, MB, Canada. *Associate:* Francis Pardo.

Scope: As part of a linguistic study of Cree: 1) Automation of certain aspects of morphological analysis and related topics, 2) automation of work on a Cree-English dictionary. *Method:* Segmentation program reads texts, affix-lists, and combination rules; attempts to break text words into affixes and stems. Putative stems are checked against the dictionary, both automatically and by linguist intervention. Erroneous segmentations generate changes to segmenter program logic or input files, correct ones generate either new dictionary entries or further attestations for existing entries.

Type of computer: IBM 360/65 (MVT). *Size of storage:* 200k. *Language:* PL/I F. *Disks or tapes:* 2 tapes. Program is available.

L489. Comparative Reconstruction of Algonkian Proto-Forms

Principal investigator: John Hewson, Professor, Dept. of Linguistics, Memorial U. of Newfoundland, St. John's, NF, Canada. *Associate:* Harry Benson.

Scope: The comparison of raw lexical data from Fox, Cree, Menomini, and Ojibwa; the finding of cognates and the reconstruction of Algonkian proto-forms. *Method:* The consonant structure of each word is extracted and all proto-form possibilities generated on the basis of the known consonant correspondences. The generated forms are then sorted across all languages so that potential cognates appear as sequences of similar forms. These sequences are then printed along with the original data from which they were generated. Manual editing eliminates discrepancies and adds the vowels to the final proto-form.

Type of computer: IBM 370/155. *Size of storage:* 100k. *Language:* FORTRAN IV. *Disks or tapes:* 1 2314 disk; 1 9-track tape. [*CHum,* May 1972]

L490. Natural Language Analysis: *Portrait of the Artist*

Principal investigator: John B. Smith, Assistant Professor, Dept. of English, Research Consultant, Computation Center, Pennsylvania State U., University Park, PA 16502.

Scope: A study of imagery in James Joyce's *A Portrait of the Artist as a Young Man* that attempts to show 1) that major moments in the development of Stephen's personality are accompanied by large concentrations of important images in the text and 2) that the development of his personality can be traced by noting the changing patterns of association that take place at these moments. *Method:* The full text was scanned and a subset of images chosen as the basis for the study. A precise model for "richness" of imagery was developed that confirms the first hypothesis. A variety of programs were used in the second phase including principal component analysis and the construction of an image concordance. Work continues in developing a precise model for association among images.

Type of computer: IBM 370/168. *Size of storage:* 3000k. *Language:* PL/I. *Disks or tapes:* 10 tapes. Program is available.

References: "A Computational Analysis of Imagery in Joyce's *A Portrait of the Artist as a Young Man," Proceedings of IFIP '71*; "Visual Analogues of Mental Structures," International Communication Association Conference, Montreal, Canada, Spring 1973.

L491. Frequency Analysis of Dutch Language

Principal investigators: J.J.M. Bakker, Project leader, U. of Technology, Insulindelaan 2, Eindhoven, Holland; J. de Rooÿ, Chief member/commission "Spoken Language," Institute of Dialectology, Amsterdam, Keizersgracht 569/71, Amsterdam, Holland; P.C. Uit den Boogaart, full-time assistant, U. of Technology, Eindhoven. *Associates:* M.L. Alinei, B. van den Berg, H. Brandt Corstius, J. Daan, J.P.M. Eggermont, A. Feitsma, B. Tervoort, A. Weÿnen, A. van Wijngaarden, D.Zwama, A.M. Zwaneveld.

Objective: Setting up a standard corpus of both written and spoken Dutch language (1,000,000 words running text), analysis of frequency of words and of syntactical patterns in the corpus. *Method:* Taking samples from written and spoken Dutch in ten different fields of language usage. The whole corpus will be coded by a lexical coding system defining grammatical qualities of each word. Parts of the corpus will be coded by a syntactical coding system, which enables the user to get syntactical information on sentences and patterns.

Type of computer: IBM 360/30. *Language:* PL/I. *Disks or tapes:* 50 Singer 3M tapes. *Special equipment:* Magnetic Recorder Friden 4301. Program is available.

L492. An English Lexicon

Principal investigator: Raoul N. Smith, Assistant Professor, Dept. of Linguistics, Northwestern U., Evanston, IL 60201; Edward R. Maxwell, Instructor, Dept. of Linguistics, Northeastern Illinois U., 5500 North St. Louis, Chicago, IL 60625.

Objective: Construction of lexical data base for computational linguistics. Implementation of insights from transformational grammar, lexicology, lexicography, and CAI. *Method:* Dictionary parsing, concordances, interactive lexicon updating, networks representation, reference making.

Type of computer: 1) CDC 6400, 2) IBM 360/50. Program is available for concordance and interactive lexicon updating.

References: R. Smith, "Interactive Lexicon Updating," *Computers and the Humanities* 6,3 (January 1972), 137-146; E.R. Maxwell, "Semantic Structures," (Unpub.Ph.D. dissertation, Northwestern University, 1972).

L493. Analysis of the Voynich Manuscript

Principal investigator: Jeffrey P. Krischer, Division of Engineering and Applied Physics, Harvard U., Cambridge, MA; 5826 Stevens Forest Road, Columbia, MD 21045. *Associate:* John H. Tiltman.

Scope: The Voynich Manuscript is a codex of unknown origin and unknown content. The language of the MS. has resisted the efforts of investigators since the manuscript's discovery in 1912. The scope of this project is to perform an analysis of the language in an effort to determine the grammar and usage for translation. *Method:* Since the manuscript is divided into sections by drawings (botanical, astrological, biological, and pharmaceutical) along with accompanying text, the investigation is in two parts: 1) a study of the drawings to identify their content and similarity with drawings found in other manuscripts of the same period; 2) a statistical examination of the properties of the text using letter and word frequencies, digraphic counts, entropy, and hidden Markov models.

Type of computer: IBM 360/65, PDP-1, CDC 6400. *Language:* FORTRAN, MIDAS. The PDP-1 using MIDAS and a system for graphical input/output of nonstandard characters was used for transliteration. Statistical analysis was accomplished on the IBM 360 and 370 series and the CDC 6600.

References: "The Voynich Manuscript," *Computational Linguistics* (1969), Harvard University, Division of Engineering and Applied Physics; John H. Tiltman, "The Voynich Manuscript," presented to the Baltimore Bibliophiles on March 4, 1967. Both papers available from principal investigator.

L494. Basil of Caesarea's *Adversus Eunomium* I-V

Principal investigators: B. Barmann, Walter M. Hayes, Centre for Medieval Studies, U. of Toronto, Toronto 5, ON, Canada.

Objective: To produce a word index and frequency count, preliminary to critical text edition and authenticity studies of books IV-V, *Adversus Eunomium,* ed. J.P. Migne (PG. 29, 497-773). *Method:* Alphabetical word index, each word presented centered on printout in a context of 160 characters. The Greek text was punched in Roman letters. This has proved quite satisfactory for a readable printout.

Type of computer: IBM 360/65. *Language:* PL/I. [*CHum,* May 1972]

L495. Legal Texts of Alfonso X, el Sabio

Principal investigator: Robert A. MacDonald, Professor, Dept. of Modern Foreign Languages, U. of Richmond, Box 278, Richmond, VA 23173.

Objective: 1) Paleographic edition of the text *Espéculo,* with variants noted; linguistic analysis; historical study. 2) Edition of the *Fuero real,* with variants noted; linguistic analysis; historical study. *Method:* Statistical analysis of quantified factors makes up the linguistic analysis.

Type of computer: IBM 1620.

References: Computers and Medieval Data Processing 1, 2 (October 1971), University of Montreal; ibid., 2, 2 (October 1972); ibid., 3, 2 (October 1973); to appear in *Speculum* (January 1974).

L496. John Duns Scotus

Principal investigator: Theodore Stallknecht, Dept. of Information Science, Victoria U. of Wellington, P.O.Box 196, Wellington, New Zealand.

Objective: To achieve a machine-readable text and concordance of Scotus' works and to apply the text to various tests for authorship, etc.

Type of computer: 1) Burroughs B6700, 2) PDP-15, 2) IBM 1130. *Language:* SNOBOL, (Burroughs Extended) ALGOL, and LATIN (a local development). *Disks or tapes:* 9-track tapes. [*CHum,* May 1972]

L497. Alliteration in Barbour's *Bruce*

Principal investigator: Bernice Widgoff Kliman, 70 Glen Cove Drive, Glen Head, NY 11545.

Objective: To determine whether alliteration in *The Bruce* is used for particular stylistic effect. *Method:* SNAP programming is used to isolate lines which contain alliteration. The stylistic analysis is then done without the computer.

Type of computer: 1) XDS Sigma 7, 2) IBM 360/65. *Language:* SNAP. Three programs are available now; others will be available as they are developed. [*CHum,* May 1972]

L498. Older Scottish Textual Archive

Principal investigators: A.J. Aitken, Paul Bratley, Neil Hamilton-Smith, Edinburgh Regional Computing Centre, King's Building, Edinburgh, EH93JZ U.K.

Objective: To provide a generally accessible source of further lexicographical and statistical information on the Older Scots texts: *Barbour's Brus* (text of Edinburgh MS., transcribed by J.A.C. Stevenson); the *Chepman and Myllar Prints* (facsimile edition, 1950); the *Asloan* MS. (S.T.S., 2 vols.); the *Bannatyne* MS. (S.T.S., 4 vols.); the *Maitland folio* MS. (S.T.S., 2 vols.); *Gilbert Hay's Prose Works* (S.T.S., 2 vols.); the *Complaynt of Scotland* (text transcribed by A.M. Stewart); Rauff Coilyear (facsimile ed., 1966); *Henryson's Poems* (text prepared by D. Fox). Uses so far have included the provision of citations for single words, combinations of words, and blocks of words for the *Dictionary of the Older Scottish Tongue,* and complete and selective concordances to single texts and groupings of texts. *Method:* The program CONCORD allows the user to select works from the archive by a variety of criteria such as title, author, date of writing,

etc. CONCORD then uses these selected works to generate a word count or a concordance. The keywords of a concordance are restricted either to those lying in a chosen alphabetic range or to those in a given list. The keywords are indexed in either dictionary or reverse alphabetic order and then the citations are ordered by either the context beside the keyword or by the reference to source text.

Type of computer: IBM 370/155.

L499. Layamon's *Brut*

Principal investigator: Sidney Berger, Dept. of English, U. of California, Davis, CA 95616. *Associates:* John McGalliard, George Rompot.

Objective: To produce an accurate, useful, legible concordance of Layamon's *Brut,* edited by Sir Frederick Madden (London: Society of Antiquaries, 1847), the only complete edition of the *Brut* ever printed. *Method:* Various programs to: 1) proofread for non-allowable characters and for correct sequence (line numbering); 2) alphabetize a deck of headword and cross-reference cards; 3) generate alphabetized word-list and frequency count; 4) generate initial concordance–line numbers only given; 5) generate final concordance in upper/lower case chain printing; 6) generate a huge deck containing complete concordance in card-form (headword and cross-reference cards to be inserted manually into this deck). *Language:* COBOL.

L500. Metrical Stress in Chaucer

Principal investigators: Merle Fifield, Director; James J. Modena, Programmer; Dept. of English and Office of Research, Ball State U., Muncie, IN 47306.

Objective: Using F.N. Robinson, ed., *The Works of Geoffrey Chaucer,* 2nd ed., Boston, 1957, to study the possible number of rhyming iambic pentameter lines, each having a different sequence of the four allowable levels of metrical stress. Approximately 1000 lines of Chaucer's poetry, selected from a variety of periods and works, are in the process of being analyzed. The analysis will be used to define Chaucer's prosody in terms of his variety of metrical stress.

Type of computer: IBM 360/40. *Language:* FORTRAN IV.

L501. A Concordance to Four Moral Plays

Principal investigators: Michael J. Preston, Dept. of English, Temple Buell College, Denver CO 80220; Louis B. Hall, Professor, Dept. of English, U. of Colorado, Denver, CO 80202. *Associates:* J.D. Shuchter, Samuel S. Coleman.

Objective: To produce a concordance, rhyme index, ranking frequency list, and reverse index to Mark Eccles' E.E.T.S. *Macro Plays* and A.C. Cawley's Manchester *Everyman.*

Type of computer: CDC 6400. *Size of storage:* 64k. *Language:* FORTRAN IV, SNOBOL4, COMPASS. *Disks or tapes:* 1 6603 disk drive; 4 607 7-track tapes. Some programs are available; contact Sam Coleman, Center for Computer Research in the Humanities, University of Colorado, Boulder, CO 80302.

Status: Concording complete. Appendices in active preparation.

L502. A Complete Glossary to the Digby Plays

Principal investigators: Louis B. Hall, Professor, Dept. of English, U. of Colorado, Denver, CO 80202; Michael J. Preston, Coordinator, Center for Computer Research in the Humanities, U. of Colorado, Boulder, CO 80302. *Associates:* Donald C. Baker, Samuel S. Coleman.

Objective: To produce a concordance and indices from which to compile a glossary for Baker's forthcoming E.E.T.S. edition of the *Digby Plays.*

Type of computer: CDC 6400. *Size of storage:* 64k. *Language:* FORTRAN IV, SNOBOL4, COMPASS. *Disks or tapes:* 1 6603 disk drive; 4 7-track 607 tapes. Some programs are available.

Status: Computer work completed.

L503. A Complete Glossary to Christ's Burial and Resurrection

Principal investigators: Louis B. Hall, Professor, Dept. of English, U. of Colorado, Denver, CO 80202; Michael J. Preston, Coordinator, Center for Computer Research in the Humanities, U. of Colorado, Boulder, CO 80302. *Associates:* Donald C. Baker, Samuel S.Coleman.

Objective: To produce a concordance and other indices to aid compilation of glossary for Baker's forthcoming E.E.T.S. edition of the plays.

Type of computer: CDC 6400. *Size of storage:* 64k. *Language:* FORTRAN IV, SNOBOL4, COMPASS. *Disks or tapes:* 1 6603 disk drive; 4 7-track 607 tapes. Some programs are available.

Status: Completed.

L504. A Complete Concordance to the Works of Sir Thomas Malory

Principal investigator: Tomomi Kato, Assistant Professor, Dept. of English, U. of Electro-Communications, 1-5 Chofugaoka 1-chome, Chofu-shi, Tokyo, Japan.

Scope: All the words in *The Works of Sir Thomas Malory,* ed. E. Vinaver, 3 vols. (Oxford, 1967), are in KWIC with information about the scribes and whether the keyword appears in the description or the colloquial part. The frequency list for each keyword is given.

Type of computer: UNIVAC 1108. *Size of storage:* 194k. *Language:* COBOL II.

L505. Grammatical and Rhetorical Analysis of Donne's *Songs and Sonnets*

Principal investigator: William Cherubini, Professor, Dept. of English, Cleveland State U., Cleveland, OH 44115.

Scope: Study of Donne's *Songs and Sonnets* according to the same routine developed for the study of Sidney's *Astrophel and Stella.* (See L443) *Songs and Sonnets* (entire) is being analyzed for distribution of word-classes (Fries-Milic code, 24 classes) and various patterns. The sample is being searched for the occurrence of about 40 rhetorical figures; all tropes are coded according to the area of experience involved ("imagery") in terms of the scheme developed by C. Spurgeon in *Shakespeare's Imagery* (Cambridge, England, 1936).

Type of computer: IBM 360/50. *Size of storage:* 384k. *Language:* SNOBOL4. Program is available. [*CHum,* May 1972]

L506. Music for Shelley's Poetry

Principal investigator: Burton R. Pollin, Professor, Dept. of English, Bronx Community College, Bronx, NY 10542. *Associate:* Kenneth Weissman.

Scope: Musical compositions, 1250 in number and based on the text of the poems and verse dramas of Percy Bysshe Shelley, have been collected throughout the United States and Europe, for the purpose of producing an annotated listing with several indices, the whole to be printed and formatted for regular publication. *Method:* The musical compositions are briefly described as to title, composer, nature of the piece, publisher, place, date, number of pages, score, location of the original piece, provenance for a few rare or unfound items, and language of the original text. Indices, to be extracted from the text, keypunched in separate fields, are planned for the dates of publication, the title of the original poem or work by Shelley, the language of the original, etc. The whole is to be formatted (with upper/lower-case text) with a machine-justified introduction, guide letters, etc.

Type of computer: IBM 360/30. *Size of storage:* 32k. *Language:* BAL. *Disks or tapes:* 3 2311 disks; 2 2415 tapes. Program is available.

Status: Publication is expected in 1973 or 1974. [*CHum,* May 1972]

L507. An Organic Criticism of Laurence Sterne's *A Sentimental Journey*

Principal investigator: Betty Pasta, 1400 30th Street N.W., Washington, DC 20007.

Objective: To conduct content analysis of *Journey* by discovering depth and location of major images and uncovering less obvious themes; to determine location and degree of unusual terms as possible key to irony. *Method:* Dictionary of complete text with frequency count; KWIC index; word grouping with manual thesaural approach for keywords.

Type of computer: IBM 360/50. *Size of storage:* 256k. *Language:* PL/I. *Disks or tapes:* 2314 disk.

Status: Completed 1973.

L508. Edition and Word Index of Heinrich Kaufringer's Poems

Principal investigator: Paul Sappler, Deutsches Seminar der U. Tübingen, D-74 Tübingen, Wilhelmstrasse 36, West Germany. *Associate:* Wilhelm Ott.

Scope: Edition of the work of Heinrich Kaufringer, German poet c.1400, and a parsed word index. *Method:* Transcription of the text on paper tape from a preliminary version of the edition, compilation of a preliminary index, and revision of the text and the apparatus by means of the index, half-automatic parsing of the index, photocomposition of text and index.

Type of computer: 1) CDC 3300, 2) IBM 1130. *Size of storage:* 1) 30k, 2) 8k. *Language:* 1)FORTRAN, COMPASS, 2) FORTRAN IV. *Disks or tapes:* 1) 1 854 disk and 3 604 tapes, 2) 1-million-byte disk.*Special equipment:* Plotter for testing programs for photocomposition, paper tape reader. Program is available.

References: Heinrich Kaufringer, *Werke;* Herausgegeben von Paul Sappler, Bd. 1: Text (Tübingen: Max Niemeyer Verlag, 1972).

L509. The Manuscript Tradition of Freidanks Bescheidenheit

Principal investigator: Paul Sappler, Deutsches Seminar der U. Tübingen, D-75 Tübingen, Wilhelmstrasse 36, West Germany. *Associate:* Wilhelm Ott.

Scope: Analysis of connections and derivation of the manuscripts and prints by comparing the order of the sentences. *Method:* Examination of how the single sentences are arranged in the several copies and which groups of sentences occur in a number of manuscripts and prints.

Type of computer: CDC 3300. *Size of storage:* 30k. *Language:* FORTRAN, COMPASS. *Disks or tapes:* 1 854 disk, 3 604 tapes. *Special equipment:* Paper tape reader. [*CHum,* May 1972]

L510. A Concordance to the Poetry of Alexander Blok

Principal investigator: Demetrius J. Koubourlis, Assistant Professor, Dept. of Foreign Languages and Literatures, U. of Idaho, Moscow, ID 83843.

Scope: The present computerized concordance is the second in a series of concordances to the Russian poets, intended to provide a modern and uniform reference to their word use. This is thought to be especially needed as the Russians have not yet appreciated the value of the concordance as a literary tool. The edition used is the four-volume *Collected Works of Alexander Blok* (Moscow-Leningrad: Gosudarstvennoe Izdatel'stvo Xudozestvennoj Literatury, 1960). *Method:* As in the Mandel'shtam concordance (see L511), each line is transliterated from the Russian Cyrillic with page, poem, and line numbers. Next, the data are keypunched and processed for automated proofreading. After corrections have been made, titles are automatically extracted and abbreviated. Next, poem numbers are replaced by title abbreviations and, after selective decapitalization and conversion back to Cyrillic, the concordance is ready to go to the printer. The final format consists of every word form found in the source text with its frequency and all lines in chronological order, in which that word form occurs. Each line includes identificatory information, such as page number, abbreviated poem title, line number, and whether the line is a variant or not. A table of headword forms in descending order of frequency appears at the end of the concordance.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* PL/1 D. *Disks or tapes:* 6 IBM disks, 6 9-track tapes. *Special equipment:* Computer printing in Cyrillic done by the CIA, Washington, DC. Program is available.

Status: Completed September 1972.

L511. A Concordance to the Poetry of Osip Mandel'shtam

Principal investigator: Demetrius J. Koubourlis, Assistant Professor, Dept. of Foreign Languages and Literatures, U. of Idaho, Moscow, ID 83843.

Scope: A concordance to the *Collected Works of Osip Mandel'shtam,* ed. G.P. Struve and B.A. Filipoff (2nd. ed., Washington, DC: Inter-Language Literary Associates, 1967). *Method:* See L510.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* PL/1 D. *Disks or tapes:* 6 IBM disks, 6 9-track tapes. *Special equipment:* Computer printing in Cyrillic done by the CIA, Washington, DC Program is available. *Status:* Completed August 1972.

L512. Joseph Smith's Prose Style

Principal investigator: Herbert Guerry, Box 80, Dept. of Philosophy and Religion, Pembroke State U., Pembroke, NC 28372.

Objective: To determine who wrote *The Book of Mormon* and other early documents of the Latter Day Saints. *Method:* Prose passages are coded into machine-readable form and analyzed according to grammatical patterns and kinds of words used.

Type of computer: IBM 1130. *Size of storage:* 8k. *Language:* FORTRAN IV. *Disks or tapes:* 1-million-byte disk.

L513. A Concordance of Sean O'Casey's Plays

Principal investigator: Brother Darby T. Ruane, Professor, Dept. of English, Iona College, 715 North Avenue, New Rochelle, NY 10801. *Associate:* Jerry G. Simotas.

Objective: A literary analysis of Sean O'Casey's plays based on a computer-generated concordance of his plays. *Method:* Currently, we are putting the plays on cards, while continuing the development and debugging of a concordance building program that has been developed and designed to operate on a limited amount of data. O'Casey's *Shadow of a Gunman, Juno and the Paycock,* and *The Plough and the Stars* have been put on computer cards and tapes.

Type of computer: IBM 1130. *Size of storage:* 8k. *Language:* FORTRAN IV. *Disks or tapes:* 2 2315 cartridges.

L514. Twelve-Year-Old Child Language

Principal investigator: N. Boot, Nederlandse Taal en Lett. Prof., Einthovenweg 13, de Bilt, Holland.

Objective: Determining regularities and a proper performance measure in Dutch for native speakers. *Method:* The investigation consists of 3 phases: 1) statistical analysis of quantified factors (word-length, sentence-length, sentence-types by punctuation marks, sorting of lexicon, mean lexicon); 2) syntax and text analysis, error analysis (sentence dictionary, discovery of error types, discovery of text types, encoding errors and text types); 3) listing of dictionary with grammatical (lemma) information for further investigation.

Type of computer: 1) IBM 360/60, 2) IBM 360/65. *Size of storage:* 1) 256k, 2) 192k. *Language:* PL/I. The first-phase program is available. [*CHum,* May 1972]

L515. Computer-Based, Curricular-Alternative Program for Critical Analysis of French Literature

Principal investigator: Louis J. Legendre, Graduate Candidate, College of Education, Pennsylvania State U., University Park, PA 16802. *Associate:* Edward R. Fagan.

Objective: The proposed study would serve as a prototype for teaching literary critical analysis in any language. The teaching technique employed is such that a) each student can learn at his own pace; b) all students can "master" the program; c) teachers are freed from "traditional" lecture-recitation chores and can serve as learning guides or tutors for their students; c) evaluations of students' performances are recorded and stored for future reference as they move through the program, and e) total literary context (as opposed to excerpts from that context) is the basis for students' critical analyses. Arbitrarily, *Les Fables* of La Fontaine were chosen as the literary work for this study because the project director's native language is French and because there are several literary critical treatises extant on La Fontaine's work. *Method:* A computer-readable form of the entire literary text of *Les Fables* was produced, including a hierarchical information to the volumes, fables, sentences, and lines; for each word, two words were created, one for the letters of the word, the other for the accents and capitalization. From this basic form three files were created: a) dictionary of the words with their frequency; b) and c) an alphabetical and a textual word order of the words containing only the hierarchical information. These three files are interrelated by a double set of pointers. — A series of programs can be used to manipulate the computer-stored text of *Les Fables* in several ways, depending upon the interests and perceptions of the program user. A program user, for example, may want to explore how the concept, love, is treated by La Fontaine in *Les Fables.* The user assumes that words such as *heart, God, mother, sex, woman,* and *women* might also be connected with *love* as a concept. With such words introduced in the program, the computer will print out every sentence within which the given word occurs and sentences which use word corollaries such as *heart, God, mother.* Or the user can control how much of the literary context he wants to scan by varying the word combinations through the insertion of special computer-meaningful words (control words) such as *or, and,not,* which initiate logical (outpt control) functions just as those same words do in Boolean algebra.

Type of computer: IBM 360/67. *Size of storage:* 280k. *Language:* PL/I and Assembler. *Disks or tapes:* 9 9-track tapes. Program is available.

L516. Structural Stylistics and *La Princesse de Clèves* by Madame de Lafayette

Principal investigator: Susan W. Tiefenbrun, Assistant Professor, Dept. of French and Romance Philology, Columbia U., New York, NY 10027.

Objective: To identify, describe, and classify stylistic and semantic patterns woven intricately within the descriptive fabric of a seventeenth-century literary work, entitled *La Princesse de Clèves,* a French novel by Madame de Lafayette. A cross-correlation of linguistic parameters will produce more accurate evidence for the interrelationship of form and content in the novel and provide quantitative information as to how the intended messages of the novel are conveyed. *Method:* Each sentence of the novel has been decoded for stylistic and semantic structures perceptible to the reader. The program will produce a computation of the frequency and distribution of linguistic structures; averages with standard

deviation for the number of words in each sentence and comparison with random samples; cross-correlation of parameters. Thus we should be able to determine which stylistic and semantic tools are present in various contexts, the extent to which these are used (tabulations and summaries), and an extensive structural profile of the novel.

Type of computer: RCA Spectra 70/35. *Size of storage:* 65k. *Language:* COBOL. Program is available.

References: "Stylistic Potentials of *La Princesse de Clèves:* A Computer-Assisted Study in the French Novel," paper delivered at the Computer Research section of the Midwest Modern Language Association in St. Louis, Missouri, October 1972.

L517. A Stylistic Analysis of Ammianus Marcellinus' *History*

Principal investigator: Geoffrey J.D. Archbold, Associate Professor, Dept. of Classics, U. of Victoria, P.O. Box 1700, Victoria, BC, Canada.

Scope: As a first step, a concordance is being produced. The text is based on Seyfarth's 1969 edition for books 14 to 21 and Clark's 1910-1915 edition, reprinted in 1963, for books 22 to 31. The stylistic analysis will involve study of the vocabulary, investigation of the length of sentences, collection of data regarding the syntactical units within the sentences, and examination of the clausulae in the complete *History. Method:* A working version of the text is being prepared by the editing of the material on magnetic tape provided by L.J. Bolchazy of the State University of New York at Albany and originally prepared for William Bayless at Brown University. The editing is being accomplished through the use of a PL/I program. The concordance and stylistic analysis are to be prepared by the use of the program CONCORD supplied by the Edinburgh Regional Computing Centre.

Type of computer: IBM 370/145. *Size of storage:* 512k. *Language:* PL/I F. *Disks or tapes:* 6 2319 disks, 4 9-track tapes. [*CHum,* May 1972]

L518. Römische Hexameterpoesie

Principal investigator: Klaus Thraede, Professor, Manfred Wacht, Akad. Rat., FB Sprache und Literatur, U. Regensburg, 8400 Regensburg, U. Postfach 397, West Germany. *Associate:* Konrad Kögler.

Scope: 1) Gesamtkonkordanz der epischen Kunstsprache, 2) Versplatzindices für das ganze römische Epos bis Frühmittelalter, 3) Leitwort- und Motivuntersuchungen, 4) Statistik metrischer Besonderheiten *in Verbindung mit* Wortwahl und Syntax, 5) Überprüfung bisheriger Deutungen metrischer Eigentümlichkeiten (Möglichkeiten der Interpretation), 6) Textkritische Untersuchungen mithilfe der Versplatzindices, 7) Metrisch-sprachliche Abhängigkeiten (Sonderverse, Imitation), 8) Lexikalische Autorkonkordanzen (mit Frequenzwörterbuch). *Method:* 1) Präedition und Ablochung aller römischen Hexameter (bisher 90.000 Verse), 2) Metrische Analyse (Skandierung, Wortgrenzen, Synalöphen usw.), 3) Sortierung und Tabellierung metrischer Merkmale (abstrakt-metrisch), 4) Metrische Analysen speziell bezogen auf Abschnitts- und Satzstruktur (Satzanfang und -schluss; Enjambement; Gleichnisse, Reden usw.).

Type of computer: Siemens 4004/45-F. *Size of storage:* 100k. (See also S171)

L519. Project DOC

Principal investigators: William S-Y. Wang, Professor, Dept. of Linguistics, U. of California, Berkeley, CA 94720; Chin-Chuan Cheng, Assistant Professor, Dept. of Linguistics, U. of Illinois, Urbana, IL 61801. *Associates:* Matthew Chen, Hsin-I Hsieh, Teresa Cheng, Mary Streeter, Vanna Condax, Steve Baron.

Scope: The study of phonological change as a function of lexical diffusion and phonetic force. The primary data base so far has been various historical stages and dialects of Chinese. *Method:* The reconstruction of historical processes on the basis of a computerized pool of linguistic information.

Type of computer: 1) CDC 6400, 2) IBM 360/75. *Language:* SNOBOL4.

References: William S-Y. Wang, "Project DOC, Its Methodological Basis," *Journal of American Oriental Society* 90, 1 (1970), 57-66; William S-Y. Wang and Chin-Chuan Cheng, "Implementation of Phonological Change," *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society* (1970), 552-559; Matthew Chen and Hsin-I Hsieh, "The Time Variable in Phonological Change," *Journal of Linguistics* (1970); Mary L. Streeter, "DOC 1971: A Chinese Dialect Dictionary on Computer," *Computers and the Humanities* 6,5 (May 1972), 259-270. [*CHum,* November 1972]

L520. Project Hippo

Principal investigator: Gilles Maloney, Dept. of Literature, Laval U., Ste-Foy, Quebec 10, Canada. *Associates:* Jacques Désautels, Jacques Bouchard, Conrad Bourdon, Michel Keable.

Scope: The *Corpus hippocraticum:* test of authorship. *Method:* Storage of Greek texts of the *Corpus* on computer; each word is followed by a figure corresponding to a morphological category. Words are written in capital or small letters, with accents and diacritics, on APL system. The purpose is to get output describing what a writer actually does and apply statistical tests to chosen features.

Type of computer: 1) IBM 360/50, 2) IBM 370/155, *Size of storage:* 1) 512k, 2) 1000k. *Language:* 1) APL, 2) FORTRAN, ASM, PL/I(TSO). *Disks or tapes:* 1) 1 9-track tape, 1 2314 disk; 2) 1 9-track tape, 1 3330 disk. *Special equipment:* APL Console 2741. Program is available.

References: "Hippocrate dans la machine," paper delivered at the annual meeting of ACFAS, November 1971 (mimeo); "Hippocrate à la question," paper delivered at the annual meeting of the Classical Association of Canada, June 1972 (mimeo). [*CHum,* November 1972]

L521. *De mortibus persecutorum* Attributed to L.C. Firmianus Lactantius: A Study on Authenticity

Principal investigator: Daniel De Decker, Assistant, Library, U. de l'Etat à Mons, Avenue Paul Deschanel 154, B1030 Bruxelles, Belgium. *Associates:* J.-H. Michel, Fr. Masai.

Scope: Study on the author's complete vocabulary. Realization of an index and concordance with the help of the computer (including an inventory of the quotations, clausulae if possible). *Method:* Probably statistical.

Type of computer: CDC 6400. *Size of storage:* 64k. *Disks or tapes:* 8 841 disks, 2 tapes. Program is available. [*CHum,* November 1972]

L522. Dating of *Skillingsviser*

Principal investigators: Erik Meiling, Datalogisk Institut, U. of Copenhagen, Kollegiegarden v/412, Tagensvej 52, 2200 Kobenhavn, Denmark; Hans Hjordt Hansen, Datalogisk Institut, U. of Copenhagen, Herninggade 23 5tv, 2100 Kobenhavn, Denmark.

Objective: The program is particularly developed for this problem, but the common purpose is to compare two text files and find accordance, after which a dating will be made. *Method:* Word comparison and statistical analysis of two text files.

Type of computer: IBM 360/75. *Size of storage:* 100k. *Language:* PL/I. *Disks or tapes:* 1 2311 disk. Program is available.

References: "Dating of *Skillingsviser,*" report and output, both in Danish (mimeo). [*CHum,* November 1972]

L523. Eadwine's Canterbury Psalter: An Edition with Notes and Glossary

Principal investigator: Frank-G. Berghaus, Assistant, Dept. of Medieval Language and Literature (English), U. of Göttingen, Nikolausberger Weg 7b, Germany. *Associate:* Jürgen Mau.

Scope: Manuscript relationships of the 15 wholly or partly glossed Old English Psalters. Glossary to Eadwine's Canterbury Psalter (in progress). *Method:* Investigation of the c. 15,000 lexicographical divergencies in the different Old English texts with special reference to the Latin versions (Romanum and Gallicanum).

Type of computer: UNIVAC 1108. *Size of storage:* 128k 36-bit words. *Language:* FORTRAN V. *Disks or tapes:* 3 9-track tapes. *Special equipment:* 5-level teletyper. Program is available. [*CHum,* November 1972]

L524. A Concordance of *Heptaméron des Nouvelles* by Marguerite de Navarre.

Principal investigator: Suzanne Hanon, Research Fellow, mag. art., Dept. of Romance Philology, U. of Aarhus, Universitetsparken, 8000 Aarhus C, Denmark.

Scope: Listing of every word used with context of one line for Marguerite de Navarre's *Heptaméron.* Other linguistic analyses of statistical character, especially investigation of the word order of subject-verb groups. *Method:* KWIC concordance type; material punched on Dura Tempowriter, 8 channels with diacritics, capitals, subscripts, and critical corrections by Le Hir (editor). [*CHum,* November 1972]

L525. A Graphemic Analysis of Modern French

Principal investigator: Suzanne Hanon, Research Fellow, mag. art., Dept. of Romance Philology, U. of Aarhus, Universitetsparken, 8000 Aarhus C, Denmark. *Associates:* Jens Vendelbo, John Tantholdt.

Objective: To make a graphemic analysis of modern French, especially consonant clusters and diphthongs. To make a comparative study between several graphemic systems, so as to isolate foreign elements in languages like French, English, etc. Program is available. [*CHum,* November 1972]

L526. Concordances of Old Provençal and Old French Lyric Poets

Principal investigator: F. Ronald P. Akehurst, Associate Professor, Dept. of French and Italian, U. of Minnesota, Minneapolis, MN 55455. *Associate:* Richard L. Hotchkiss.

Scope: Concordances of William IX, Bernard de Ventadour, Marcabru, Cercamon, Arnaut Daniel, Comtesse de Die, etc; Grace Brulé, Gautier de Dargies, Guiot de Dijon, Blondel de Nesle, etc. *Method:* File of words in context is built and then sorted.

Type of computer: CDC 6600. *Size of storage:* 65k. *Language:* FORTRAN IV and COMPASS. *Disks or tapes:* 2 6603 and 4 801 disks; 4 7-track tapes. Program is available. [*CHum,* November 1972]

L527. A Concordance of Pascal's "Lettres provinciales"

Principal investigators: Hugh M. Davidson, Professor, Dept. of Romance Languages, Ohio State U., Columbus, OH 43210; Pierre H. Dubé, Assistant Professor, Dept. of Romance Languages, U. of Waterloo, ON, Canada.

Scope: A thematic investigation of the "Lettres provinciales." This shall be done in conjunction with our study of the *Pensées concordance.* The concordance shall be KWIC, with special tags for words in italics, words in variants, and dialogue. All diacritical marks shall be included. Words shall be printed in upper/lower case.

Type of computer: IBM 370/165. *Size of storage:* 125k. *Language:* PL/I. *Disks or tapes:* 3 2400 9-track tapes. Program is available in part. [*CHum,* November 1972]

L528. IMPAC (Indexing, Map Plotting, and Analysis by Computer)

Principal investigator: Gerald C. Keil, Fellow, Dept. of Computation, U. of York, Heslington, York, YO1-5DD, England.

Objective: Computer-aided analysis of the basic material contained in H. Orton et al., *Survey of English Dialects* (Leeds, 1962-), and selected additional material in the Survey Archives, University of Leeds. *Method:* 1) Preparation, on 8-track paper tape, of the corpus (narrow phonetic transcription will be preserved); 2) creation of a magnetic tape master file of the above, and provision for its correction and extension; 3) extraction onto disk or magnetic tape of subsets of the corpus (or of previously extracted subsets) for subsequent sorting and/or production of maps, tables, word-lists, concordances, and ultimately, a comprehensive index.

Type of computer: ICL 1906A. *Size of storage:* c. 32k (24-bit). *Language:* ALGOL 68/R. *Disks or tapes:* 3 160kcls tapes; 1 420 Mch FD disk. *Special equipment:* CalComp 663 incremental plotter. [*CHum,* November 1972]

L529. Stylometry

Principal investigator: Jürgen Mau, Professor, Dept. of Classics, U. of Göttingen, Nikolausberger Weg 13, Germany.

Scope: Let the linguistic form of a text be completely described by an ordered set of n values characterizing such features as, for example, mean sentence length, relative frequency of certain strings of characters or rhetorical figures, etc. Various texts showing no significant difference in the value of a subset of m characteristic parameters ($m < n$) may be classified together as belonging to the same literary species. Which are the features that can be characterized by measured values and have the above qualities? What information can be drawn from applying the method of multiple regression and correlation to a matrix of the described values?

Type of computer: 1) UNIVAC 1108, 2) IBM 1130. *Size of storage:* 1) 128k words, 2) 8k halfwords. *Language:* 1) FORTRAN V, 2) Basic FORTRAN IV, ASM. *Disks or tapes:* 3 9-track tapes. [*CHum,* November 1972]

L530. Affiliation

Principal investigator: Jürgen Mau, Professor, Dept. of Classics, U. of Göttingen, Nikolausberger Weg 13, Germany. *Associates:* D. Najock, H.-J. Ahnert.

Scope: Logical and set theoretical background of affiliation and evaluation of manuscripts. Development of programs therefor. *Method:* The manuscripts under consideration are attributed to the single bits of one or more machine words. Each word or string of words represents an error type. Comparison of these error types will lead to more or less contaminated stems.

Type of computer: 1) UNIVAC 1108, 2) IBM 1130. *Size of storage:* 1) 128k words, 2) 8k halfwords. *Language:* 1) FORTRAN V, 2) Basic FORTRAN IV, ASM. *Disks or tapes:* 3 9-track tapes. *Special equipment:* Teletype. Program is available. [*CHum,* November 1972]

L531. PROFILE, Part One

Principal investigator: James M. Watkins, Professor, Dept. of Foreign Languages, Five Colleges, Inc., Amherst, MA 01002. *Associate:* Robert H.Gonter.

Objective: To develop and apply a CAI language for use by all schools having access to the University of Massachusetts computer. *Method:* Sentence judging by groups. A student's sentence is compared with that supplied by the teacher. For the comparison, the sentence has been divided into structural and morphological groups. Within the number of tries determined by the teacher, the student will match all groups by 1) guessing, once shown the error, 2) requesting help through messages directed to each group, 3) asking for the correct answer and rewriting it.

Type of computer: 1) CDC 3800, 2) CDC 3600. *Size of storage:* 1) 32k, 2) 32k. *Language:* 1) and 2) FORTRAN IV (UMASS). *Disks or tapes:* 1) and 2) CDC 841 disks. *Special equipment:* IBM 2741, Datel, Teletype. Program is available.

References: PROFILE, Part One. Author's Manual. [*CHum,* November 1972]

L532. Natural Language Analysis: Four Fundamentalist Sermons

Principal investigator: John B. Smith, Assistant Professor, Research Consultant, Dept. of English/Computation Center, Pennsylvania State U., University Park, PA 16502. *Associate:* Bruce Rosenberg.

Scope: A study of the formulaic structure of four fundamentalist sermons. We are attempting to show that these sermons follow patterns similar to those found in the oral folktale or poem. The notion of formula is extended to a number of structural patterns that vary in scope from the phrase to the entire performance. *Method:* Transcriptions of four performances were encoded in their entirety as well as various measures indicating rapidity of delivery, audience response, timbre of voice, etc. A variety of techniques including Fourier analysis used to establish structural patterns.

Type of computer: IBM 360/67. *Size of storage:* 1024k. *Language:* PL/I. *Disks or tapes:* 10 tapes. Program is available.

References: Bruce Rosenberg, *The Art of the American Folk Preacher* (New York: Oxford University Press, 1971). [*CHum,* November 1972]

L533. EYEBALL: A Computer Program for Description of Style

Principal investigators: Donald Ross, Assistant Professor, Dept. of English, U. of Minnesota, Minneapolis, MN 55455; Robert Rasche, Associate Professor, Dept. of Economics, Michigan State U., East Lansing, MI 48823.

Objective: To indicate many features of the language of literary texts, and to describe features of style from the marked text. Language features which are marked are word, clause, and sentence lengths, a surface-structure grammatical code, and vocabulary distribution. Style features which are described in the present implementation also include word-class distributions, and complete word indexes for each text. *Method:* Natural-language input text is partially analyzed with two intermediate stages where human intervention allows for reliable parsing. An "augmented" or marked text is the main output.

Type of computer: 1) CDC 6600, 2) IBM 360/370. *Size of storage:* 1) 50k words (octal), 2) 80k bytes (decimal). *Language:* 1) FORTRAN, 2) FORTRAN G or H. *Disks or tapes:* Access to disk pack (2 files or data sets). Program is available.

References: "Description and User's Instructions for EYEBALL" (Minneapolis: University of Minnesota English Department, 1972, mimeo); "EYEBALL: A Computer Program for Description of Style," *Computers and the Humanities* 6 (1972), 213-221; "Beyond the Concordance: Algorithms for the Description of English Clauses and Phrases," *Computers and Literary Studies,* ed. A.J. Aitken, R.W. Bailey, N. Hamilton-Smith (Edinburgh: Edinburgh University Press, 1973).

L534. A Concordance to the Poetry of F.I. Tjutcev

Principal investigator: Geir Kjetsaa, Professor, Slavonic Dept., U. of Oslo, Box 1028, Oslo 3, Norway. *Associates:* Steinar Gil, Erik Egeberg.

Scope: Concordance to the poetry of F.I. Tjutcev; analysis of his poetry, mainly his vocabulary, but also syntactic studies. *Method:* Statistical analysis of quantified factors. Besides, research is carried out on the poetry of E.A. Baratynskij, M. Ju. Lermontov, A. Fet, A. Achmatova, and I. Annenskij, whose poems have been converted to machine-readable form by the Slavonic department at the University of Oslo, and on random selections (approximately 2000 words each) from poetic texts of 21 Russian poets born 1792 to 1821.

Type of computer: CDC 3300. *Size of storage:* 32k. *Language:* FORTRAN. *Disks or tapes:* 1 854 disk; 1 659 tape. [*CHum,* November 1972]

L535. Gerard Manley Hopkins and His Associates: Editorial and Critical Analysis

Principal investigators: Todd K. Bender, Associate Professor, Dept. of English, U. of Wisconsin, Madison, WI 53706; Robert J. Dilligan, Assistant Professor, Dept. of English, U. of Southern California, Los Angeles, CA 90007. *Associates:* James W. Parins, Shirley K. Johnson.

Objective: To generate texts and concordances of Hopkins and his circle of correspondents: Bridges, Patmore, Dixon. To proceed with descriptive and analytic statements about the use of language on the pre-semantic and semantic level. For example, to convert the orthographic to a phonetic text and then scan for patterns on assonance, alliteration, rime, and other features associated with pre-semantic verse patterning. Also to consider larger structures on the semantic level such as collocation of verbal imagery. *Method:* 1) Where necessary generate a scholarly text through computer collation. 2) Generate a verbal concordance to that text. 3) For pre-semantic analysis convert orthographic to phonetic text. 4) Various programs to search for patterns on the pre-semantic level. 5) Various programs to investigate collocation, tense-sequences, etc.

Type of computer: UNIVAC 1108. *Size of storage:* 256k. *Language:* FORTRAN V.

References: Robert J. Dilligan and Todd K. Bender, *Concordance to Gerard Manley Hopkins* (Madison, WI: University of Wisconsin Press, 1970). [*CHum,* November 1972]

L536. Computer-Assisted Investigation of the Prosody of John Keats and Gerard Manley Hopkins

Principal investigator: Robert Dilligan, Assistant Professor, Dept. of English, U. of Southern California, Los Angeles, CA 90007.

Scope: To produce a detailed computer index to the patterns of phonemic recurrence in the poetry of Keats and Hopkins with an eye towards understanding the evolution of their prosodic styles. *Method:* To scan verse using a revised version of Halle-Keyser's metrical theory. The programs determine the occurrence of stress maximum, elision, assonance, alliteration, and consonance for each line, poem, and corpus. The scansions of each line are cross-referenced by means of multiple field sorting with IBM's SORT/MERGE package. Statistical clustering techniques are applied to the results of each poem to see if they can be correlated with external criteria such as genre, date of composition, theme.

Type of computer: IBM 370/155. *Size of storage:* 400k. *Language:* PL/I, FORTRAN IV, BAL. *Disks or tapes:* 4 2314 disks, 3 9-track tapes. Program is available.

References: "The Lapses of Time: A Computer-assisted Investigation of English Prosody," *The Computer and Literary Studies,* ed. A.J. Aitken, R.W. Bailey, N. Hamilton-Smith (Edinburgh: Edinburgh University Press, 1973); "Ibant Obscvri: Robert Bridges' Experiment in English Quantitative Verse," *Style* (forthcoming). [*CHum,* November 1972]

L537. Concordance to the Work of Joseph Conrad

Principal investigators: Robert Dilligan, Assistant Professor, Dept. of English, U. of Southern California, Los Angeles, CA 90007; Todd Bender, Associate Professor, Dept. of English, U. of Wisconsin, Madison, WI 53706. *Associates:* Sybil Jacobson, James Parins.

Objective: To produce index concordances to the complete corpus of Conrad's prose and to develop statistical profiles of his lexicon for each work and for the corpus as a whole; to demonstrate the usefulness of microfiche for concordances whose bulk would argue against book format publication; to reduce the cost of concordances for users; to prepare texts for a computer-collated critical edition of Conrad. *Method:* Concordances to each work will be produced and issued as they are ready, along with basic information about word frequency. The individual concordances will also be saved on tape to be merged into a final concordance to all of Conrad. This final concordance will include lists of words unique to each work and other statistical information about the distribution of Conrad's lexicon.

Type of computer: IBM 370/155. *Size of storage:* 360k. *Language:* FORTRAN IV, BAL. *Disks or tapes:* 1 9-track tape; 2 2400 disks.

References: A Concordance to Joseph Conrad's "Heart of Darkness" (Carbondale: Southern Illinois University Press, 1973). [*CHum,* November 1972]

L538. A Concordance to the Poetry of Theodore Roethke

Principal investigators: Robert Dilligan, Assistant Professor, and Patrick Morrow, Assistant Professor, Dept. of English, U. of Southern California, Los Angeles, CA 90007.

Objective: To develop a successful context algorithm for lyric poetry based upon stanzaic units; to study the development of Roethke's diction through vocabulary-matching techniques and statistical measures; to reduce the cost of concordance production through efficient use of COM (Computer Output on Microfilm) and photocomposition devices. *Method:* Production of KWIC concordance to entire corpus and KWOC concordances for each poem and grouping of the collected works; isolation of lexical sets unique to each grouping; application of statistical clustering techniques to study the continuity and dispersion of stylistic traits.

Type of computer: 1) IBM 370/155, 2) RCA 70/830 VIDEOCOMP, 3) III COMP 80. *Size of storage:* 1) 360k, 2) 65k, 3) 16k. *Language:* 1) PL/I, FORTRAN IV, BAL; 2) Assembler; 3) Assembler. *Disks or tapes:* 3 2314 disks, 2 2400 tapes. The PL/I programs for context are available. [*CHum,* November 1972]

L539. A Concordance to the Poetry of Alexander Pope

Principal investigators: Robert Dilligan, Assistant Professor, Dept. of English, U. of Southern California, Los Angeles, CA 90007; Emmett Bedford, Assistant Professor, Humanistic Studies Division, U. of Wisconsin-Parkside, Kenosha, WI 53140.

Objective: To integrate the latest current technology in optical scanning directly from printed texts with videocomposition of graphic arts quality to produce a fully automated concordance to the poetry of Pope. The concordance will preserve all accidentals of the printed text and is designed to set a new standard for attractiveness as well as the utility of computer-generated material. *Method:* Direct Scanning of the Twickenham Edition of Pope has been completed by DISSLY Systems of Eaton, Pennsylvania. Preliminary sorting and post-sorting editing was carried out on the UNIVAC 1108 at the University of Wisconsin. Final sorting and editing, including formatting for videocomposition, is now in progress at the University of Southern California.

Type of computer: 1) UNIVAC 1108, 2) IBM 370/155, 3) RCA 70/830 VIDEOCOMP. *Size of storage:* 1) 65k (6 byte words), 2) 400k, 3) 65k. *Language:* 1) FORTRAN V; 2) PL/I, BAL, SORT/MERGE; 3) Assembler. *Disks or tapes:* 1) 2 7-track tapes, 2) 2 9-track tapes; 6 2314 disks. *Special equipment:* Fastrand Drum used in 1108 sorting. Program is available.

References: A Concordance to the Poetry of Alexander Pope (New York: Holt Information Systems, Holt, Rinehart & Winston, 1972). [*CHum,* November 1972]

L540. Statistical Analysis of the Vocabulary and Sentence Length of Elizabethan Drama

Principal investigators: Robert Dilligan, Assistant Professor, Dept. of English, U. of Southern California, Los Angeles, CA 90007; Louis Ule, 27 Mustang Road, Rolling Hills Estates, CA.

Objective: To investigate the vocabulary of early Elizabethan blank verse with an eye towards resolving certain questions of disputed authorship attribution. Twenty-five texts, including the entire corpus of Christopher Marlowe, have been keypunched, and are to serve as a basic data base for this enquiry. *Method:* KWOC concordances for each play have been produced through the use of CONSTAT, a concordance statistical program for natural language developed by Louis Ule. The output from this program is stored in separate data sets on magnetic tape and provides the basic data from which we hope to determine the stability and confidence levels for different stylistic traits such as vocabulary distribution, sentence- and word-length, character count, Yule's characteristic. A bevy of matching, clustering, and merging programs have been developed as part of the CONSTAT program package.

Type of computer: IBM 370/155. *Size of storage:* 400k. *Language:* BAL, FORTRAN IV H. *Disks or tapes:* 2 2314 disks, 2 9-track tapes.

References: "The Mathematics of Style," *Times Literary Supplement,* October 21, 1971. [*CHum,* November 1972]

L541. Statistical Style Analysis of the Book of Isaiah: Problem of Authorship

Principal investigators: Larry L. Adams, Research Analyst, Office of Institutional Research; Alvin C. Rencher, Associate Professor, Dept. of Statistics, Brigham Young U., Provo, UT 84601; Wayne Larsen, Statistician, Bell Telephone Laboratories. *Associates:* Ellis T. Rasmussen, William J. Adams, Robert C. Patch.

Scope: Statistical analysis of style of the Hebrew text from the Old Testament. *Method:* Statistical comparisons, based on several hundred stylistic variables consisting of approximately seventy types of literary elements, were made between the book of Isaiah and eleven other books from the Old Testament. Inter-text author variation was compared with intra-text variation using distribution-free methods of statistical analysis to avoid mistakes commonly made by style researchers who have used statistical procedures.

Type of computer: IBM 360/50. *Language:* FORTRAN VI, COBOL, SNOBOL, PL/I. *Disks or tapes:* 2 3200 fci tapes. Program is available. [*CHum,* November 1972]

L542. Use of Computers in Full Range of English Curricula

Principal investigator: Detmar W. Straub, Jr., Assistant Professor, Dept. of English, Gannon College, Erie, PA 16501. *Associate:* Philip Kelly.

Scope: In connection with the NSF-sponsored project PRISE, I have written a variety of programs for courses in the following subject areas: freshman composition, creative writing, medieval literature (applicable to all upper-level courses), science fiction (useful also in Computer and Society courses), and a course in computer projects in English. Inasmuch as the primary function of all of these programs is to add new excitement to traditional literary analyses, I am surveying the classes for their responses to this innovation.

Type of computer: PDP-10. *Language:* BASIC. Program is available. [*CHum,* November 1972]

L543. Iran-Afghanistan Bibliography Project

Principal investigators: K.A. Luther, Associate Professor, Dept. of Near Eastern Languages and Literature; Allan R. Emery, Assistant Director, Computing Center, U. of Michigan, Ann Arbor, MI 48104. *Associates:* Stanley T. Mendenhall, Margaret S. Fearey, Frances L. Trix, Orick L.Peterson.

Objective: To provide readily available bibliographic information, freeing scholars for research. *Method:* Each bibliographic entry is searched for selected keyword parameters. Each entry contains a line of abstracted keywords which is searched in addition to the title.

Type of computer: IBM 360/67. *Size of storage:* 16k. *Language:* Assembler-G. *Disks or tapes:* 1 Telex disk, 1 tape. *Special equipment:* Special Michigan Terminal System subroutines.

References: Jerome W. Clinton et al., "On the Feasibility of an Automated Bibliography of Iranian Studies," *Iranian Studies* 2, 4 (Autumn 1969), 151-169 (with sample search profile and printout). [*CHum,* November 1972]

L544. A Concordance to the Poetry of Nikolai Gumilev

Principal investigator: Demetrius J. Koubourlis, Assistant Professor, Dept. of Foreign Languages and Literatures, U. of Idaho, Moscow, ID 83843.

Scope: The present computerized concordance is the third in a series of concordances to the Russian poets which intends to provide a modern and uniform reference to their word use. The edition used is the 4-volume *Collected Works of N. Gumilëv,* ed. G.P.Struve and B.A. Fillipov (Washington, DC: Victor Kamkin, 1964). *Method:* See L510.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* PL/I D. *Disks or tapes:* 6 disks, 6 9-track tapes. *Special equipment:* Computer printing in Cyrillic done by the CIA, Washington, DC. Program is available. [*CHum,* November 1972]

L545. Automatic Morphemic Segmentation of Modern Russian

Principal investigator: Demetrius J. Koubourlis, Assistant Professor, Dept. of Foreign Languages and Literatures, U. of Idaho, Moscow, ID 83843.

Scope: Zellig Harris was able to show that a given phoneme sequence of English could be segmented into morphemes without appeal to meaning. This method was based on the interdependence of linguistic units, in this case, phonemes. A computer right-to-left and left-to-right scanning of a sample of English text in phonemic representation produced word and morpheme boundaries. The implications are of tremendous significance, not only on the theoretical plane but on the practical as well (machine translation, deciphering of texts written in unknown languages). Preliminary tests have been conducted by the investigator on modern Russian, modern Bulgarian, literary Polish, and modern Greek. The results of these tests indicate that Harris' findings may lead to the discovery of a language universal. The preliminary tests further established that the numerical differences between predecessors and successors tend to approach zero, i.e., the number of unique phonemes preceding a given phoneme is equal or nearly equal to the number of unique phonemes following that phoneme. *Method:* The corpus consisted of a phonemically transcribed version of *Ballada o soldate,* ed. L.C. Thompson, W. Konick, V. Gross (New York: Harcourt, Brace & World, 1966). The unit for computer search was the accentual unit, a word form with clitics.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* PL/I. *Disks or tapes:* 6 disks, 5 tapes.

References: Zellig Harris, "From Phoneme to Morpheme," *Language* 31, 2 (1955), 190-222; Zellig Harris, *Mathematical Structures of Language* (New York: Interscience Publishers, 1968), pp.11-16, 20-28. [*CHum,* November 1972]

L546. The Automatic Disambiguation of Biaspectuals in Modern Russian

Principal investigator: Demetrius J. Koubourlis, Assistant Professor, Dept. of Foreign Languages and Literatures, U. of Idaho, Moscow, ID 83843.

Scope: This project will attempt to automatically disambiguate biaspectual verb forms using specially designed computer programs. Available evidence indicates that verbal aspect in Russian is context-dependent and, in a given context, a verb form is either perfective or imperfective. This evidence leads to the null hypothesis that determination of this aspect-context relationship in uniaspectual verb forms can be instrumental in positing aspectual characteristics for biaspectual verb forms. The present project will attempt to measure statistically the aspect-context relationship in aspectually unambiguous instances occurring in contemporary Russian prose and test the null hypothesis. Whether the null hypothesis is proved or disproved, this study will increase knowledge about the contextual characteristics of aspect choice which, in turn, will be of use to all those interested in Russian verbal aspect. If the null hypothesis is proved, as preliminary results indicate, I hope to develop a technique for automatically disambiguating biaspectual verb forms in Russian, which will be of theoretical and practical interest to automatic language translation specialists. *Method:* After a suitable corpus is created in computer-processable form, searching for and tabulation of varied combinations of aspectual features takes place. Next, statistical measurements of the contingency of co-occurring features are obtained. And, finally, processing for the statistical determination of aspect choice in biaspectual cases is performed.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* PL/I D. [*CHum,* November 1972]

L547. Computer-Rationed Stress in Russian Language Materials

Principal investigator: Demetrius J. Koubourlis, Assistant Professor, Dept. of Foreign Languages and Literatures, U. of Idaho, Moscow, ID 83843.

Scope: Stress in advanced Russian language readers (texts) has not been treated uniformly. Some authors stress all but monosyllabic words, others stress even monosyllabic words when they are used non-clitically, and still others may either alternate stressed chapters with unstressed ones or may not indicate stress at all. Because stress is generally marked manually, it is the rare reader (textbook) which enjoys a high degree of consistency as well as freedom from error. Both of these shortcomings can be remedied by relegating stress marking to a computer. The concept of computer-rationed stress reaches far beyond any of the manual stress-marking systems in use. As the goal in learning how to read Modern Russian correctly is reading in stressless form, the student is guided, via computer-rationed stress, to rely on stress-free form. *Method:* The screenplay of Josef Heifetz's *Dama s Sabachkoj,* a film based on Chekhov's story, is the basis for an advanced Russian language reader, featuring computer-rationed stress. Optionally offered with a reader for classroom use will be a copy of the film. A computer program determines the frequency of word forms in a finite text and then decides how many of the first appearances of each word form will appear stressed. A different stress mark appears on word forms in regulated intervals as check and reinforcement. The final computer output consists of the original text (optionally in camera-ready form) with rationed stresses. This method is readily applicable to any language with mobile stress.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* PL/I D. *Disks or tapes:* 2 disks, 3 tapes. *Special equipment:* Cyrillic print chain needed for preparation of camera-ready copy. [*CHum,* November 1972]

L548. Concordance of the Plays of Christopher Marlowe

Principal investigator: Molly Shelton Hussing, M.A. candidate, Dept. of English, U. of Texas at Arlington, Arlington, TX 76010. *Associate:* James Wyatt.

Objective: To produce a concordance as well as an adaptable program for further study of Marlowe's works. *Method:* The first phase will be to concord only the two *Tamburlaine* plays. They are currently being coded.

Type of computer: XDS Sigma 7. *Size of storage:* 64k. *Language:* SNOBOL4. *Disks or tapes:* 2 9-track tapes. [*CHum,* November 1972]

L549. Cervantes' Works: Old-Spelling Edition and Concordance

Principal investigator: R.M. Flores, Assistant Professor, Dept. of Hispanic and Italian Studies, U. of Victoria, BC, Canada.

Objective: To recover Cervantes' orthography and provide a KWIC old-spelling concordance to, and a machine-generated old-spelling edition of, his works. The initial concordance body to *La Galatea* and *Don Quixote* is complete. The project is now in its second phase. Multiple readings are being singled out and compositorial spellings are being replaced with Cervantes' own spellings.

Type of computer: IBM 370/145. *Size of storage:* 512k bytes. *Language:* PL/I F. [*CHum,* November 1972]

L550. An Index of the Machine-Detectable Figures of Speech in Lucan

Principal investigator: Theodore Crane, Assistant Professor, Dept. of Classics and Italian, U. of South Carolina, Columbia, SC 29208. *Associate:* Penelope Crockett Crane.

Objective: Our purpose is to list the location and frequency of such machine-detectable figures of speech as alliteration, anaphora, conduplicatio, homeoptoton, homeoteleuton, polysyndeton, and polyptoton in the 10 books of Lucan's first-century A.D. Latin epic poem *De Bello Civili.* Eventually the same investigations will be applied to Vergil's *Aeneid* and a comparison made. An index of proper names for both is also planned. *Method:* The program language allows us to type one line of poetry on each computer card. The programs and their application to the generated data will be developed and tested. The cards will be converted to magnetic tape and all the results will be made available through the American Philological Association Repository of Greek and Latin Texts, Department of Classics, Dartmouth College, Hanover, NH 03755.

Type of computer: IBM 360/65. *Size of storage:* 128k. *Language:* SNOBOL4. *Disks or tapes:* 2 9-track tapes. [*CHum,* November 1972]

L551. Recherches en vue de l'analyse automatique des productions écrites des élèves de l'école élémentaire

Principal investigators: Roberte Tomassone, Maître Assistante; Bernard Combettes, Maître Assistant, Linguistique Appliquée, Centre de Recherches et d'Applications Linguistiques (C.R.A.L.), Boulevard Albert I, 54-Nancy, France. *Associates:* M. Thoss, J.B. Jouin.

Scope: Mettre au point une méthode de description aussi complète que possible de la langue des élèves dans le but de pouvoir comparer les productions d'élèves suivant différentes méthodes, différents types de pédagogie (au niveau de la syntaxe, du vocabulaire). Corpus limité pour l'instant à 500 copies—un total de 40.000 mots environ. *Method:* 1) Codification; 2) établissement des index alphabétiques de fréquence; 3) établissement des listes des différents types de phrases, transformations utilisées, fréquence de chacune de ces catégories; 4) classement des phrases en fonction de leur degré de complexité: liste des différents schémas de phrases rencontrés; 5) analyse multidimensionnelle sur les caractéres linguistiques ci-dessus et des variables socio-culturelles.

Type of computer: CII 10070. *Size of storage:* 320k. *Language:* COBOL. *Disks or tapes:* 2 disques fixes, 8 bandes moyenne densité. *Special equipment:* Traceur de courbes. Program is available. [*CHum,* November 1972]

See also E1, E2, G5, G15, G18, G25, G53, G57, M43, P1, P2, P4, P5, P6, P9, P12, P16, P17, P18, S16, S25, S26, S39, S41, S42, S54, S77, S83, S93, S142, S143, S159, S161.

Music

M1. Computer Music

Principal investigators: M.V. Mathews, Member of Technical Staff; F.R. Moore, Associate Member of Technical Staff; Behavioral and Acoustical Research, Bell Telephone Laboratories, Murray Hill, NJ 07974.

Objective: Synthesis of music and sounds on a digital computer.

Type of computer: 1) Honeywell 635, 2) DDP-224. *Size of storage:* 1) 40k, 2) 32k. *Language:* 1) and 2) FORTRAN. *Disks or tapes:* 1) 1 tape, 2) 1 disk. *Special equipment:* Digital-to-Analog converters. Program is available.

References: M.V. Mathews, *Technology of Computer Music* (Cambridge, MA: MIT Press, 1969).

M2. Computer Sound Generation

Principal investigator: David Cohen, Associate Professor, Dept. of Music, Box 2886, University, AL 35486. *Associates:* Luther Nale, Robert F. Tinker.

Objective: To develop flexible and widely usable computer program(s) for sound generation employing FORTRAN IV and to write a nontechnical manual which will enable composers with no previous computer experience to use the program(s). *Scope and method:* The programs are similar to the Bell Laboratory Music IV Program, but are structured differently. The computer-produced BDC tape must be run through a digital-to-analog converter to produce the sound.

Size of storage: 32k. *Disks or tapes:* 1 tape. *Language:* FORTRAN IV. May also require machine language. *Special features:* Digital-to-analog conversion. Listing of simplified program and preliminary description available. Nontechnical manual and sample audio tapes available shortly. Description and listing of the full-scale program (PAN) will be made available, when completed.

Status: Technical difficulties with our digital-to-analog converter have delayed completion of descriptive manual and thorough testing of program. A larger, more general, program is still being written. *Future plans:* Full-scale program to be completed by Fall 1967. Simplified program (PERFORM) to be thoroughly tested and fully described. [*CHum*, May 1967]

M3. University of Chicago Chanson Project

Principal investigator: Lawrence F. Bernstein, Associate Professor, Dept. of Music, U. of Pennsylvania, Philadelphia, PA 19104.

Objective: Automated bibliographic and style-analytical study of the sixteenth-century chanson. Generation of a "Thematic Concordance" of the repertory; machine extraction of quantitative stylistic data. *Scope:* The entire repertory of sixteenth-century French secular polyphonic chansons.

Type of computer: IBM 7094. *Language:* CLML Programming Language, FAP, and machine language. Program will be available when debugged.

Status: Pilot project run on slightly less than 5 percent of total repertory. Results satisfactory. *Future plans:* Scope of project to be expanded to include entire repertory.

References: Lawrence F. Bernstein and Joseph P. Olive, "Computers and the 16th-Century Chanson: A Pilot Project at the University of Chicago," *Computers and the Humanities* 3, 3 (January 1969), 153-160.

M4. Musical Composition

Principal investigator: Gerald Strang, Lecturer, Dept. of Music, U. of California, Los Angeles, CA 90024. *Associates:* Leon Knopoff, John Gardner.

Objective: To investigate the possibilities of computer sound synthesis and compositional techniques. *Method:* Adaptation of Bell Telephone Laboratories' MUSIC V.

Type of computer: IBM 360/91. D/A conversion on Hewlett-Packard. *Language:* FORTRAN G, machine language. Current UCLA version of Bell Labs MUSIC V is MUSIC VC.

Status: Fully operative. Since 1969 the following music has been produced by Dr. Strang: *Compusition —8, Compusition —9, Prelude and Fugue, Compusition —10—Atmosonus.*

References: "The Computer in Musical Composition," *Computers and Automation* 15, 8 (August 1966), 15; "The Problem of Imperfection in Computer Music," *Music by Computers,* ed. Heinz von Foerster and James Beauchamp (New York: John Wiley & Sons, 1969); "Ethics and Esthetics of Computer Composition," *The Computer and Music,* ed. Harry B. Lincoln (Ithaca: Cornell University Press, 1970).

M5. Thematic Indexing by Computer, Including Development of Necessary Hardware to Permit Printing of Music by Computer

Principal investigators: Harry B. Lincoln, Professor, Dept. of Music; Christian Granger, Manager of Academic Support, Computer Center; SUNY, Binghamton, NY 13901.

Status: In progress, indices of sixteenth-century secular (madrigals) and sacred music (motets, masses, hymns, etc.); about 5000 titles (with incipits of all voices) completed to date. Complete, the thematic indices of printed sources of the frottola and the Palestrina repertory (Casimiri edition). Being developed is the necessary hardware (print symbols for high-speed printer) to print music notation as part of output. A limited number of special music type characters have been developed by the IBM Glendale Laboratory; details can be obtained through IBM representatives.

References: "A Computer Application to Musicology; The Thematic Index" in papers of the International Federation of Information Processing Societies (Edinburgh, 1968); "The Thematic Index: A Computer Application to Musicology," *Computers and the Humanities* 2, 5 (May 1968), 215-220; "Toward a Computer Typography for Music Research: A Progress Report," Information Processing '71 (IFIP Congress, Ljubljana), 1971 (TA7, pp.32-35; "Uses of the Computer in Music Composition and Research," *Advances in the Computer* (New York: Academic Press, 1974), pp.73-114.

M6. A Topical Index of Musical Treatises

Principal investigator: Cecil D. Adkins, Associate Professor, School of Music, North Texas State U., Denton, TX 76203.

Objective: A multilingual (original language and English) topical index of theoretical musical treatises to the year 1800, providing source order and alphabetical listings of their contents. The index also contains locations of rare items together with translations and information on significant commentaries. *Method:* The index is programmed so that one may request printouts of individual items or complete treatises. Information retrieval has been programmed with a KWIC index. Printouts are also supplied chronologically.

Type of computer: IBM 1620 and 1440. *Language:* SPS (1620); modified COBOL. Program is available but is highly specialized and would have to be adapted to a specific usage.

Status: Index now contains about 25,000 entries; subsequent entries are being added at about the rate of 300 per week. *Future plans:* Completed index is expected to contain in excess of 100,000 entries (over a period of several years). [*CHum,* May 1967]

M7. Computer Music

Principal investigator: L. Knopoff, Professor, U. of California, Los Angeles, CA 90024. *Associates:* Gerald Strang, William Hutchinson, Mantle Hood.

Objective: Analysis, synthesis, and performance of music.

Type of computer: IBM 7094. *Size of storage:* 32k. *Language:* FORTRAN IV, MAP. *Disks or tapes:* 8 729 tapes. *Special feature:* Direct data connection (2 channels). Program is available.

Status: 1) Analysis of Javanese gender improvisation techniques plus Javanese Saron nuclear-themes (in progress). 2) Synthesis (in progress). 3) Performance (production). [*CHum,* May 1967]

M8. Integrated Composition of Music and Dynamic Visual Art

Principal investigator: Jack Citron, IBM Los Angeles Scientific Center, 10889 Wilshire Boulevard, Los Angeles, CA 90024.

Scope: "Learning" programs allow one to gain experience communicating with each part of the integrated program separately and also help specify the most desirable form for the final high-level language.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV E or H with GRAF. *Disks or tapes:* 1 2311 disk, 1 7-track and 2 9-track tapes. *Special features:* IBM 2250 display and controlled moving picture camera. Program is available.

Status: Static and dynamic visual art programs exist in CPS and are being debugged in GRAF. Four music programs are being amalgamated into a final form. *Future plans:* The art and music programs will be combined when audiovisual output hardware is ready. [*CHum,* May 1967]

M9. Investigation of Various Compositional Devices

Principal investigator: Edward G. Kobrin, graduate student, Dept. of Music, U. of Illinois, Urbana, IL 61801. *Associate:* T.H.A. Ashford.

Scope: Investigation of the use of various controls in the writing of music for large ensembles (symphony orchestras, symphonic bands) with special consideration to transposing instruments. *Method:* 1) Serialize given parameters with emphasis on strict synchronization of vertical and horizontal serialization. 2) Choose instruments for transposition. 3) Reprogram for correct pitch of these instruments.

Type of computer: CDC 6400. *Size of storage:* 40k. *Language:* ALGOL 60. Program is available.

References: "A Solution to the Problems of Vertical Serialization," *Perspectives of New Music* (Fall-Winter, 1968). [*CHum,* November 1968]

M10. MUSIGOL, An Electronic Music Preprocessor

Principal investigator: Donald MacInnis, Associate Professor, Dept. of Music, U. of Virginia, Charlottesville, VA 22903. *Associates:* William A. Wulf, Paul S. Davis, Donald Bowers, Donald Wright.

Objective: A versatile sound-generating program to study the effects of carefully controlled parameter changes on the perceived sound wave, and as a sound source for original musical compositions. *Scope and method:* MUSIGOL is a skeleton program which becomes complete when the researcher or composer supplies descriptions of "instruments" with all their parameters, and the "score." MUSIGOL then "plays" the music by creating a magnetic tape with amplitude sample numbers at 20,000/sec.

Type of computer: Burroughs 5500, and Ambilog 200 hybrid computer (as converter). *Size of storage:* 24k plus disk memory of 1.2 million words. *Language:* Extended ALGOL. *Disks or tapes:* 3 magnetic tapes, 10 disk files. *Special features:* Includes facilities for a CalComp plotted graph of the final complex wave. User's manual and/or line printout available.

Status: MUSIGOL is operational; improvements are constantly being added; the psychoacoustic research is in progress. Program is continuing. [*CHum,* May 1967]

M11. Analysis of Musical Relationships in Contemporary Music

Principal investigators: Theodore H.A. Ashford, Instructor; Edward G. Kobrin; School of Music, Northwestern U., Evanston, IL 60201.

Objective: To analyze contemporary scores (since 1960) and test the findings by writing compositions in similar styles using the computer. *Method:* Analysis of scores using the tools of information theory. Constructing programs to do the analyses and to write compositions. Scores selected include Berio, Feldman, Drew, Boulez.

Type of computer: CDC 3600. *Language:* ALGOL 60.

Status: Initial stages. [*CHum,* May 1967]

M12. MUSE: A Syntax-Directed System for Music Analysis

Principal investigators: George W. Logemann, Assistant Professor; Gary Berlind, graduate student; New York U., Institute for Computer Research and Humanities, New York, NY.*

Objective: To develop a convenient retrieval language and system for music analysis.

Type of computer: IBM 360/30. *Size of storage:* 65k. *Language:* PL/I, Assembly, BTAM. *Disks or tapes:* 2 2311 disks. *Special features:* Remote access devices.

References: Murray Gould and George W. Logemann, "ALMA: Alphameric Language for Music Analysis," *ICRH,* 1966; Gary Berlind and George W. Logemann, "TPOSE: A Program for Musical Transposition," *ICRH.* [*CHum,* May 1967]

M13. The Structure of Atonal Music

Principal investigator: Allen Forte, Professor of Theory of Music, Dept. of Music, Yale U., New Haven, CT 06520. *Associate:* Stephen Smoliar.

Scope: Study of structure in atonal compositions. More general applications are foreseen. *Method:* The complete score is represented using the input language developed by S. Bauer-Mengelberg. An analytical reading program operates on this representation to provide output subsequently processed at higher analytical levels.

Type of computer: IBM 370. *Language:* SNOBOL4, FORTRAN (high level).

References: A Syntax-Based Score Reading Program, Project MAC Technical Report TR-32, May 1967; *The Structure of Atonal Music* (New Haven: Yale University Press, 1973).

M14. A Syntax-Directed Compiler to Analyze Musical Compositions and Text

Principal investigator: Eors N. Ferentzy, Assistant Professor, Dept. of Computer Science, U. of Toronto, Toronto, Canada. *Associate:* Jim Gabura.

Objective: To gather statistics on the strings analyzed. *Scope and method:* A combination and extension of syntax, directed parsing, and Markovian analysis.

Type of computer: IBM 360/50, 75, IBM 7094/II. *Size of storage:* 20k-50k words. *Language:* Specifications written in formula ALGOL, implementation in PL/I. *Disks or tapes:* IBM disk unit. Program is available.

Status: Basic routines specified in formula ALGOL. *Future plans:* To use the system for the statistical characterization of musical and linguistic styles. [*CHum,* May 1967]

M15. Harmonic Structure in Variation Movements by Webern

Principal investigator: Mary E. Fiore, Professor, Dept. of Music, State University College at Buffalo, 1300 Elmwood Avenue, Buffalo, NY 14222.

Scope: All movements in variation form in their entirety. *Method:* 1) Encode music scores. 2) Scan for melodic patterns. 3) Isolate and reference chords. 4) Reference chords to melodic patterns and form.

Type of computer: CDC 6400. *Size of storage:* 64k. *Language:* SNOBOL4. *Disks or tapes:* 1 6603 disk and 4 604 tape drives. Program is available.

References: "The Formation and Structural Use of Vertical Constructs in Selected Serial Compositions," dissertation, Indiana University, 1962; "Harmonic Analysis as a Determinant of Piano Performance: Selected Twentieth-Century Compositions," paper read at Piano Forum, Buffalo, 1965; "Motive in Webern's Piano Variations," *Computer Uses in Music* (Ithaca, NY: Cornell University Press). [*CHum,* May 1968]

M16. Opera Codex

Principal investigator: John Nagosky, Associate Professor, U. of South Florida, Tampa, FL 33620.

Objective: To place into a single volume as much objective information on opera as would be most useful to a musician.

Type of computer: IBM 7040. *Size of storage:* 32k. *Language:* FORTRAN IV.

Status: Preliminary determinations and analysis of the operas. [*CHum,* May 1967]

M17. Music Information Retrieval (MIR)

Principal investigator: John Selleck, technical assistant, Dept. of Music, Princeton U., Princeton, NJ 08540. *Associates:* Arthur Mendel, Lewis Lockwood.

Scope: To devise an information retrieval system for music employing the Intermediate Musical Language (IML) devised here at Princeton also.

Type of computer: IBM 7094, IBM 360/50. *Size of storage:* 32k. *Language:* FORTRAN II, BEFAP. *Disks or tapes:* BEFAP takes 1 7-track tape. Program is available.

Status: Minor results in the area of simultaneity counting have been obtained with the 7094 version of the system, now obsolete. The new 360 version, all in FORTRAN, is nearly completed. Program is available.

References: Michael Kassler, "Toward Musical Information Retrieval," *Perspectives of New Music* (Spring-Summer 1966), p.59; Arthur Mendel, "Some Preliminary Attempts at Computer-Assisted Style Analysis in Music," *Computers and the Humanities* 4, 1 (September 1969), 41-52; Tobias Robison, "IML-MIR: A Data Processing System for the Analysis of Music," *Elektronische Datenverarbeitung in der Musikwissenschaft,* ed. Harald Heckmann, (Regensburg: Gustav Bosse Verlag, 1967). [*CHum,* November 1970]

M18. Investigation of Tonal Properties of the Gregorian Repertory

Principal investigator: Finn Egeland Hansen, Institute of Musicology, U. of Aarhus, Denmark.

Objective: Statistical analysis of frequencies of tones and tone combinations. Program will be available.

Status: Program and data tapes in preparation. *Future plans:* Construction of standard programs and codes for musicology. [*CHum,* May 1967]

M19. A Pilot Study on "Rhythmic Levels"

Principal investigator: John G. Suess, Assistant Professor, Dept. of Music, U. of Wisconsin-Milwaukee, Milwaukee, WI 53211.

Objective: To establish a testable language for determining levels of functional rhythm. *Scope:* Limitation to a study of Webern's Symphony Opus 21. Method not yet established.

Type of computer: IBM 1620 and 1401. *Size of storage:* 60k. *Disks or tapes:* 2 disks for 1620; 4 tapes and 2 disks for 1401.

Status: Initial phase of language determination. [*CHum,* May 1967]

M20. Investigation of Style of Keyboard Composition of Ruttini and Haydn

Principal investigator: Carlo Lombardi, graduate student, New York U.; Dept. of Music Education, Adelphi U., Garden City, NY 11530.

Type of computer: 360/30. *Size of storage:* 65k. *Language:* PL/I, Assembly, BTAM. *Disks or tapes:* 2 2311 disks. *Special features:* Remote access devices. [*CHum,* May 1967]

M21. A Computer Music-Typographist

Principal investigator: Jeff F. Raskin, Dept. of Visual Arts, U. of California at San Diego, La Jolla, CA 92037.

Objective: To utilize standard hardware and list processing graphics system of original design to produce, rapidly and inexpensively, engraved quality music reproductions, suitable for standard photo-publishing. *Method:* Input is direct from manuscript. Output is on a digital incremental plotter.

Type of computer: 1) IBM 7074, 2) IBM 1401, 3) IBM 360/50, 4) IBM 360/67. *Size of storage:* 1) 10k, 2) 4k, 3) 256k, 4) 1000k. *Language:* FORTRAN IV, DAFT, 1401 AUTOCODER, 360 Assembly language.

Status: Pilot project completed. [*CHum,* May 1967]

M22. Generation and Plotting of Musical Tone Systems by Binary Logarithm

Principal investigator: Andrew G. Pikler, Ret. USN Research Psychologist, 198 Laurel Hill Avenue, Norwich, CT 06360. *Associates:* David Kellish, Thomas Carty, Barry Stevens.

Objective: Comprehensive knowledge of the tonal inventories of all systems (Pythagorean, harmonic, tempered); retrievals of scales and chords; automatic plotting of tonal spirals and helices; computation of the defects of systems vs. musical contexts; optimization. *Scope and method:* Programming of infinite and finite sets from theoretical equations, linear and angular.

Type of computer: IBM 1620 with IBM 1627 plotter; 2) IBM 7040. *Size of storage:* 1) 40k, 2) 32k. *Language:* FORTRAN II D and FORTRAN IV. *Special features:* 1) indirect addressing, automatic divide, disk. Program is available, dependent on U. of Connecticut release.

Status: Last phase of the work covered the subdivision of the octave in terms of dualistic intonation. Temporarily discontinued because of the retirement of persons involved.

References: "Logarithmic Frequency Systems," *Journal of the Acoustical Society of America* 39, 6, 1106-1110; "History and Theory of the Tonal Spiral," *Journal of the Acoustical Society of America* 40, 5 (November 1966), 1253; "Generation and Plotting Musical Tone Systems with the Digital Computer," paper read at the Sixth International Congress on Acoustics in Tokyo, August 1968.

M23. Music Printing by Plotter and Cathode-Ray Tube

Principal investigator: Barry S. Brook, Executive Officer, Ph.D. Programs in Music, Graduate School, City University of New York, 33 west 42 Street, New York, NY 10036. *Associate:* Larry Heimrath.

Objective: Develop programs for representation of musical notation.

Type of computer: IBM 1620 and IBM 1627 incremental digital pen plotter. *Language:* Machine language.

Status: Investigating next stage: cathode ray tube to microfilm recorder. *Future plans:* Development of a complete printout method for musical symbols, scheduled to begin in 1974.

M24. RILM: International Repertory of Music Literature

Principal investigator: Barry S. Brook, Executive Officer, Ph.D. Programs in Music, Graduate School, City University of New York, 33 West 42 Street, New York, NY 10036. *Associate:* Larry Heimrath.

Objective: International computer-indexed abstracts of music literature covering articles, books, dissertations, congress reports, Festschriften, etc., from 1967 on. *Scope:* Bibliographical control of scholarly literature on music. Quarterly publication with in-depth cumulative computer-indexing.

Type of computer: IBM 360. Program is available, when ready.

Status: Vol. VI, no.1 scheduled for publication August 1973. *Cumulative Index,* vols. I-V in press. *Future plans:* Publication of a multilingual thesaurus now in preparation. Application of system to other journals (two already under way: *CHum, RILA*).

References: "RILM," *Computers and the Humanities* 1 (1967), 103-108; *Nysuc Music Library Association Notes* 23 (1967), 462-467; *College Music Symposium* (1969).

M25. Theories of Tonal Bass Harmonization

Principal investigator: John Rothgeb, Dept. of Music, Yale U., New Haven, CT 06520.

Scope: Historical and contemporary theories (interpreted as procedures) for harmonizing unfigured basses are evaluated with respect to effectiveness, adequacy, and simplicity. A highly adequate theory is designed. *Method:* Pitch-class and interval-class grammars of the tonal bass are developed. Grammars specified by relational networks (or equivalent list structures) are used analytically to assign structural descriptions to tonal basses.

Type of computer: IBM 7040. *Size of storage:* 32k. *Language:* SNOBOL3. [*CHum,* May 1968]

M26. Music Simulation on the 1130 Computer

Principal investigator: John E.H. Hancock, Associate Professor and Director, Computer Center, Reed College, Portland, OR 97202. *Associate:* Mark Roberts.

Scope: Traditional classical music, pentatonic scales, electronic music. *Method:* Simulation of tempered scale notes by using various cycled instructions.

Type of computer: IBM 1130. *Size of storage:* 8k. *Language:* 1130 Assembler and FORTRAN. *Disks or tapes:* 1 disk. [*CHum,* November 1967]

M27. Root Progression and Composer Identification

Principal investigator: Joseph E. Youngblood, Associate Professor, Dept. of Music Literature, U. of Miami, Coral Gables, FL 33124.

Scope: First movement each of Hindemith, Clarinet Quintet, Op.30; Bartók, String Quartet No.4; Schönberg, String Quartet No.3; Hindemith, String Quartet No.6. *Method:* 1) Code (in a modified DARMS) and keypunch. 2) Extract roots (modified Hindemith) and calculate root progressions. 3) Calculate redundancy through three successive roots. 4) Calculate *chi-square* for various groupings.

Type of computer: 1) IBM 7040, 2) IBM 1401. *Size of storage:* 1) 32k. *Language:* FORTRAN IV. *Disks or tapes:* 10 729 tapes (6 mod 2, 4 mod 4). Program is available for steps 2 and 3.

Status: Completed.

References: "Root Progression and Composer Identification," *The Computer and Music,* ed. Harry B. Lincoln (Ithaca, NY: Cornell University Press, 1970), pp. 172-178.

M28 . FORTRAN Music Programs Involving Numerically Related Tones [Deleted]

[*CHum,* November 1967]

M29. Computer Applications to Béla Bartók's Serbo-Croatian Material

Principal investigator: Benjamin Suchoff, Curator, Bartók Archives, 333 East 79 Street, New York, NY 10021.

Scope: Continuation of pilot project commenced in 1966 at SUNY, Binghamton: incipit interval sequence extraction program to yield lexicographical index of melodies and other data. *Method:* 1) Transmutation of incipits into DARMS. 2) Keypunch. 3) Transfer to data bank (disk). 4) Extract incipit interval sequence as plus or minus integers. 5) Sort-merge: plus precedes minus and lesser precedes greater interval. 6) Identical strings are juxtaposed in single-space vertical format. 7) Print lexicographical index in tabular form.

Type of computer: 1) IBM 360/40, 2) IBM 360/30. *Language:* 1) FORTRAN IV, 2) PL/I.

References: "Computer Applications to Bartók's Serbo-Croatian Material," *Tempo* No.80 (London, Spring 1967); Béla Bartók, *Rumanian Carols and Christmas Songs* (Vol.4 of *Rumanian Folk Music*), ed. Benjamin Suchoff (The Hague; Martinus Nijhoff); cf. Appendix: "Lexicographical Index [computer-derived] of Colinda Melodies"; Idem, *Tabulation of Serbo-Croatian Material,* ed. Benjamin Suchoff (The New York Bartók Archives); cf. Appendix: "Lexicographical Index [computer-derived] of Parry Melodies." [*CHum,* November 1967]

M30. Improvisation and Music Theory of the Javanese Gamelan

Principal investigator: Mantle Hood, Professor, Institute of Ethnomusicology, U. of California, Los Angeles, CA 90024. *Associate:* Fredric Lieberman.

Scope: 1) Selected sample of 82 nuclear themes illustrating the various modes and forms, transcribed from traditional Javanese notations. 2) Selected sample of 38 improvisations on the gender, transcribed from performance tapes. *Method:* 1) Transcribe. 2) Keypunch. 3) Transfer to tapes. 4) Statistical analysis. 5) Information-theory analysis. 6) Pattern-search.

Type of computer: IBM 360/75. *Size of storage:* 512k. *Language:* PL/I F. *Special equipment:* IBM 2260 display stations. [*CHum,* November 1967]

M31. Faktorenanalytische Typenbestimmung an einstimmigen Volksmelodien

Principal investigator: Reiner Kluge, wissenschaftlicher Mitarbeiter, Rechenzentrum beim II. Mathematischen Institut, Humboldt-Universität zu Berlin, 108 Berlin, Unter den Linden 6, DDR.

Scope: 1000 einstimmige Volksmelodien aus der Landschaft Altmark, in transkribierter Form (Sammlungen Parisius, Stockmann u.a.). *Method:* 1) Codieren. 2) Lochen auf ZRA-1-Lochkarten. 3) Auditive Kontrolle. 4) Melodie-Analyse (nach 150 Merkmalen). 5) Quantitative Ahnlichkeitsbestimmung. 6) Faktorenanalyse. 7) Typenbeschreibung. 8) Druck der Typen-Tabelle.

Type of computer: ZRA 1 (Zeiss). *Size of storage:* 4k. *Language:* ZRA-1-Code. *Special equipment:* Elektronische Orgel für auditive Kontrolle. Program is available.

References: "Volksliedanalyse und-systematisierung mit Hilfe eines Rechenautomaten," *Bericht über den Kongress der Gesellschaft für Musikforschung,* Leipzig, 1966 (in Druck); "Zur automatischen quantitativen Bestimmung musikalischer Ahnlichkeit," *Bericht über den X. Kongress der I.M.S.,* Ljubljana 1967 (in Druck). [*CHum,* November 1967]

M32. A Music Compiler for Musicologists and Ethnomusicologists

Principal investigator: Jerome Wenker, 11 Carlton Club Drive, Piscataway, NJ 08854.

Scope: Design and implement a compiler for use in the analysis of European "Art" music and ethnomusicological transcriptions from other cultures.

Type of computer: 1) IBM 360 or 370 (OS), 2) CDC 6600, 3) UNIVAC 1106, 1107, 1108. *Size of storage:* 1) 225-280k bytes, 2) 120k octal, 3) 45k. *Language:* 1), 2), and 3) FORTRAN and ASM. *Disks or tapes:* 1), 2), 3) 1 standard tape, 1) 1 2314 disk, 2) and 3) 1 standard disk. Program is available.

Status: The mnemonic music notation MUSTRAN containing both ethnomusicological symbols and common music notation and a basic set of analysis programs are now available for any one of the following machines and operating systems: IBM 360 or 370 (OS/MVT or OS/MFT); CDC 6600 (SCOPE); and UNIVAC 1106, 1107, or 1108 (Exec 2 and Exec 8). *Method:* The MUSTRAN notation allows the preparation of music in machine-readable format and the existing analysis programs evaluate selective aspects of monophonic music. Once additional analysis programs have been produced to determine the required features of a music compiler, that compiler will be designed and implemented.

References: "A Computer Oriented Music Notation Including Ethnomusicological Symbols," *Musicology and the Computer,* ed. Barry S. Brook (New York: The City University of New York Press, 1970), pp.91-129. A revised version of this article may be obtained from the author. [*CHum,* November 1972]

M33. Statistical Comparison of Musical Styles

Principal investigator: Frederick Crane, Associate Professor, School of Music, U. of Iowa, Iowa City, IA 52240.

Scope: Polyphonic chansons of the early fifteenth century, many of which are anonymous, are the subject of a project to determine whether a set of statistics resulting from the analysis of a musical work can be used to identify the composer of the work. *Method:* Compositions encoded in DARMS are analyzed, partly by hand, for many aspects of style, each represented by a numerical value. The resulting data are subjected to cluster analysis, with the object of separating the works into distinct clusters, each of which consists of works by only one composer.

Type of computer: IBM 360/65. *Size of storage:* 200k. *Language:* SNOBOL4, FORTRAN IV G. Program is available.

References: Frederick Crane and Judith Fiehler, "Numerical Methods of Comparing Musical Styles," *The Computer and Music,* ed. Harry B. Lincoln (Ithaca, NY: Cornell University Press, 1970), pp. 209-222.

M34. Canon One

Principal investigator: Mangham D. Lehr, Instructor, Dept. of Creative Arts, Purdue U., Stanley Coulter Hall, Lafayette, IN 47907.*

Scope: A program to determine the points of stretto in a given tonal melody in strict canon. *Method:* A melody is compared against a copy of itself through twelve transpositions to determine the acceptable starting points.

Type of computer: 1) IBM 7094 Mod.1, 2) CDC 6500. *Size of storage:* 1) 32k, 2) 65k. *Language:* 1) SNOBOL3, 2) SNOBOL4. *Disks or tapes:* 1) 8 tapes, 1301 mod 2 disks, 2) 4 tapes, disks. *Special equipment:* 4 on-line teletype, 252 graphic display system. Program is available. [*CHum,* May 1968]

M35. Computer Synthesis of Musical Sounds and Psychoacoustical Stimuli

Principal investigator: Donald MacInnis, Associate Professor, Dept. of Music, U. of Virginia, Charlottesville, VA 22903. *Associates:* William A. Wulf, Donald Bowers, Donald Wright.

Method: Program calculates 20,000 amplitude samples per second of any desired complex wave form; these amplitudes, in the form of numbers between 0 and 4096, are then converted from digital to analog. Program is similar to Bell Lab's MUSIC 4, but is written entirely in Burroughs 5500 Extended ALGOL. An Ambilog 200 computer is used as D-to-A converter. User can define original instruments to suit his imagination, then "call" these by specifying parameters on individual "note" cards.

Type of computer: Burroughs 5500 and Ambilog 200. *Size of storage:* 24k. *Language:* B5500 Extended ALGOL. *Disks or tapes:* 4 7-channel 62,500 char/sec tape units. *Special equipment:* CalComp plotter. Program is available.

References: "Sound Synthesis by Computer: MUSIGOL, a Program Written Entirely in Extended ALGOL," *Perspectives of New Music* 6 (Winter 1967-68). [*CHum,* May 1968]

M36. Statistics in Style Research

Principal investigator: Stefan M. Kostka, Dept. of Music, U. of Texas, Austin, TX 78712.

Scope: Various aspects, primarily vertical, of four of Paul Hindemith's string quartets. *Method:* 1) The music was encoded melodically, including only pitch and rhythm. 2) The melodic deck was used to produce an annotated harmonic deck. 3) Both decks were then used as sources of data for analysis programs.

Type of computer: CDC 3600. *Language:* FORTRAN IV, Assembler. *Disks or tapes:* 2 tapes.

References: "A Computer-Assisted Study of Some Aspects of Style Development in the String Quartets of Paul Hindemith" (Unpub. Ph.D. dissertation, U. of Wisconsin, 1969).

M37. Music Synthesis by Computer

Principal investigator: John Clough, Professor, School of Music, U. of Michigan, Ann Arbor, MI 48108. *Associates:* R. Teitel, J. Leonard, E. Sosman.

Scope: Research in programming systems for music generation by means of digital computers and digital-to-analog converters; design of a special-purpose programming language for composers; educational applications of computer-generated sound.

Type of computer: IBM 360/44. *Size of storage:* 128k. *Language:* FORTRAN G and Assembler. Program is available. [*CHum,* May 1970]

M38. Studies Relating to Automation of Bartók Archives

Principal investigators: Benjamin Suchoff, District Director of Music, Woodmere-Hewlett Public Schools, Hewlett, NY 11557; Jack Heller, Professor, Dept. of Computer Science, SUNY, Stony Brook, NY 11790. *Associate:* Jane Steiner.

Scope: To produce permuted indices, including a lexicographical index of musical themes from data in "Transylvanian Hungarian Folk Songs" (1921) by Béla Bartók and Zoltán Kodály. *Method:* 1) Data derivation by means of annotators assigned to specific categories assembled from Bartókian ethnomusicological methods, and procedures; 2) transmutation of music notation into machine-readable characters by means of the DARMS music representation; 3) data bank construction by means of the GRIPHOS (*G*eneral *R*etrieval and *I*nformation *P*rocessor for *H*umanities-*O*riented *S*tudies) system (see G13); 4) tabulations obtained in the form of a variety of permutated indices.

Type of computer: IBM 360/67. *Size of storage:* 240k. *Language:* PL/I(5). *Disks or tapes:* 3 2314 disks, 3 2400 tapes. *Special equipment:* T.N. chain.

References: Jack Heller and Gary Berlind, "The ICRH Index Generator, General Considerations and Input Specifications," *ICRH Newsletter* (New York University) 2, Nos. 6-7, 1967; Benjamin Suchoff, "Computer Applications to Bartók's Serbo-Croatian Material," *Tempo* 80 (London, 1967); "Computerized Folksong Research and the Problem of Variants," *Computers and the Humanities* 2 (1968), 155-158; B. Suchoff, "The Computer and Bartók Research in America," *Hungarian Musico-Historical Studies* 3 (Budapest, 1971); also in *Journal of Research in Music Education* 19, 1 (Spring, 1971); B. Suchoff, "Computer-Oriented Comparative Ethnomusicology," *Music and the Computer,* ed. Harry B. Lincoln (Ithaca: Cornell University Press, 1970); "Some Problems in Computer-Oriented Bartókian Ethnomusicology," *Ethnomusicology* 13, 3 (September 1969) and *Revista de Etnographie si Folclor* 14, 5 (Bucharest, 1969); *Building a Collection Data Bank at Museum Computer Network* (New York: IBM Application Brief GK 20-0370-0).

M39. Correlational Analysis to Establish Authenticity of Eighteenth-Century Symphonies

Principal investigator: Jan LaRue, Professor, Dept. of Music, Graduate School of Arts and Science, New York U., New York, NY 10003. *Associate:* Murray Gould.

Objective: To test the power of musical style analysis as a means of establishing attributions of eighteenth-century symphonies, many of which are incorrect. The specific test is a comparative study of an early Mozart symphony which in one source is attributed to Dittersdorf. *Method:* Quantification of factors of sound, harmony, melody, and rhythm as they interact during the course of phrase formation, followed by correlation studies by computer.

Type of computer: IBM 360/30. *Size of storage:* 20k. *Language:* FORTRAN, PL/I. [*CHum,* May 1969]

M40. Notation of Computer-Generated Music

Principal investigator: Edward G. Kobrin, graduate student, Dept. of Music, U. of Illinois, School of Music, Urbana, IL 61801.

Objective: To obtain a suitable standard of notation for music conceived on the computer. *Method:* 1) Separate subroutines are assigned and called for each note of the chromatic scale. 2) The octave and correct transposition is calculated within the subroutine.

Type of computer: 1) CDC 6400, 2) IBM 360. *Size of storage:* 1) 43k, 2) 65k. *Language:* FORTRAN IV. *Disks or tapes:* 1 CDC 6603 disk, 1 IBM disk and 1 tape for each machine. *Special equipment:* 565 CalComp plotter. Program is available. [*CHum,* May 1969]

M42. A Thematic Locator for Mozart's Works as Listed in Köchel's *Chronologischthematisches Verzeichnis*

Principal investigators: George R. Hill, Apt. 13-B, 3800 Parkview Lane, Irvine, CA 92664. *Associates:* Murray Gould, David Noon, Hans-Rudolf Dürrenmatt, Jan LaRue.

Objective: To produce a thematic index arranged by pitch contours to the works of Mozart. *Method:* Music incipits from the Köchel catalog encoded onto specially designed IBM mark-sense cards arranged to simulate a musical staff; marks converted to punches by IBM card punch; incipits ordered and printed in alphabetic notation by computer.

Type of computer: IBM 360/30. *Size of storage:* 32k. *Language:* PL/I F. *Disks or tapes:* 3 2311 + data cell (2321) disks; 2 tapes. *Special equipment:* Card punch equipped with column binary punching capability. Program is available. [*CHum,* May 1969]

M43. Word-Concordance of J.S. Bach's Vocal Texts

Principal investigator: Arthur Mendel, Professor, Dept. of Music, Princeton U., Woolworth Center of Musical Studies, Princeton, NJ 08540. *Associates:* Paul W. Gunzelmann, Tobias Robison, John Selleck.

Scope: To make it possible to locate every use of every word in the texts set by Bach or probably set by him, for all the purposes which word-concordances normally serve.

Type of computer: IBM 360/65.

Status: Completed 1970.

References: The finished concordance is available ($40).

M44. Linguaggio Musicale della Sonata a Quattro Italiana del Secolo XVI

Principal investigator: Mario Baroni, Assistente, Istituto di Storia della Musica, Università di Bologna, via Trombetti 4, 40126 Bologna, Italy.

Scope: Dati caratteristici della "Sonata a quattro" e rilevanze stilistiche dei singoli autori. *Method:* Analisi quantitativa e distributiva dei parametri del suono. Raccolta dei dati riguardanti la Sonata come genere e differenze fra questi dati e quelli dei singoli autori.

Type of computer: IBM 7090. [*CHum,* May 1969]

M45. An Analytical Approach to Folk Music

Principal investigator: Jerome Wenker, 11 Carlton Club Drive, Piscataway, NJ 08854.

Scope: A series of analysis programs applied to 277 Anglo-American ballad tunes from England, Scotland, Ireland, New England, and the Southern Appalachians. These programs include interval counts and percentages, scale analysis, identification of existing 3- and 4-note interval sequences. evaluation of common and selected interval sequences, and some rhythmic analysis. Regional differences were found to exist and, following further evaluation of the results, a paper will be prepared for publication. *Method:* Music data is prepared using the MUSTRAN notation (see M32) and converted into an internal format using the MUSic TRANslator. The interval format is saved on disk or tape and is then available for the individual analysis programs. Some analysis programs count occurrences of selected factors (intervals, interval patterns, scale runs, etc.) by song, group of songs, and corpus. These results are then available for statistical analysis. Other analysis programs are strictly concerned with song analysis (number and length of phrases, table of pitch versus note duration, etc.) and do not produce data appropriate for statistical analysis.

Type of computer: 1) IBM 360 or 370 OS, 2) CDC 6600, 3) UNIVAC 1106, 1107, 1108. *Size of storage:* 1) 225-280k (bytes), 2) 120k (oct), 3) 45k. *Language:* For all machines, FORTRAN and ASM. *Disks or tapes:* For all machines, 1 (standard) tape; 1) 1 2314 disk, 2) and 3) 1 (standard) disk. Program is available.

References: "A Computer-Oriented Music Notation Including Ethnomusicological Symbols," *Musicology and the Computer,* ed. Barry S. Brook (New York: The City University of New York Press, 1970), pp.91-129. A revised version of this article is obtainable from the author.

M46. A Syntax-Directed, Changeable System for Processing Data for Musical Analysis

Principal investigator: Raymond F. Erickson, Assistant Professor, Dept. of Music, Queens College, Flushing, NY 13657.

Objective: To provide a general-purpose tool for syntax-error detection of (musical) data in string form and for the conversion of such data strings into tables or other output format desired by the user for subsequent musical analysis. *Method:* The system is a PL/I implementation of a modified form of the syntax-directed compiler designed by T.E. Cheatham and Kirk Sattley, Jr. The completed system consists of a series of subprograms, some of which may be altered or provided by the user to accommodate any encoding method.

Type of computer: IBM 360/50. *Size of storage:* 1000+k bytes. *Language:* PL/I. Program is available.

References: "A General-Purpose System for Computer-Aided Musical Studies," *Journal of Music Theory* 13, 2 (1970). [*CHum,* May 1970]

M47. An Analytical and Indexing System for Incipits of Renaissance Polyphonic Compositions

Principal investigator: Lynn Mason Trowbridge, Instructor, Dept. of Music, Alma College, Alma, MI 48801.

Objective: To develop a system for the cataloging of musical incipits ordered according to their musical relatedness (this first part has been completed). To develop a system for the study and comparison of musical style as found in the chansons of four generations of composers, c. 1400-1550. *Method:* The system for the cataloging of incipits was completed using the method previously described. Style comparisons will be carried out by a series of programs designed to interlock with those already completed. These programs will print an analytical synopsis of each composition as well as a statistical summary of all the compositions in each of the repertories.

Type of computer: IBM 360/75. *Size of storage:* 160k. *Language:* ANS COBOL, FORTRAN IV H, Assembly G. *Disks or tapes:* 3 9-track tapes. Disks used only as temporary storage. Program is available.

References: C. Hamm and H. Kellman, "The Musicological Archive for Renaissance Manuscript Studies," *Fontes Artis Musicae,* no.3 (1969), pp.148-149; L. Trowbridge, "A Computer Programming System for Renaissance Music," summary of project delivered at the Renaissance Study Session, Annual Meeting of the A.M.S., Toronto, November 1970; L. Trowbridge, "The Burgundian Chanson: An Index and Analysis of Incipits by Application of Electronic Data Processing." (Unpub. Master's thesis, University of Illinois, 1971.)

M48. Computer Pattern Recognition of Standard Engraved Music Notation

Principal investigator: David S. Prerau, Dept. of Electrical Engineering, U.S. Department of Transportation, Transportation System Center, Kendall Square, Cambridge, MA 02142.

Objective: To recognize visually the music information specified by a sample of standard engraved music notation. The input to the programs is an array of black and white points which corresponds to a few measures of a page of music. The programs attempt to recognize the music information specified by the input. If the programs are successful, the output will be the DARMS notational strings which correspond to the input sample. Though there are several notational items which are not considered, this is the first investigation which attempts to recognize a significant set of music symbols. The programs are written in such a way as to facilitate future expansion, though this expansion will not be trivial. *Method:* The programs are divided into three major groups: input, isolation, and recognition. The input section used a flying-spot scanner and obtains, after some conversions, a packed array which corresponds to a part of a page of music. (Present sample size is about three measures of a duet.) The isolation programs solve the difficult problem of isolating the symbols and staff-lines, almost all of which are graphically connected. The recognition programs do the final recognition, using symbol properties such as normalized size, notational syntax, and graphical features.

Type of computer: MIT CTSS TimeSharing System (IBM 7094). *Size of storage:* 32k. *Language:* MAD and SLIP. *Special equipment:* A flying-spot scanner. [*CHum,* May 1970]

M49. A Bibliography of Music Research Done in the Southern Division, MENC, 1912-1968

Principal investigator: Robert W. John, Professor, Dept. of Music, U. of Georgia, Athens, GA 30601. *Associates:* Karen Rice, William Kethley.

Scope: An index of all masters theses and doctoral dissertations written in Southern colleges and universities in the field of music from 1913 through 1968. Also includes a KWIC index of titles. *Method:* Survey method.

Type of computer: IBM 1401. *Size of storage:* 16k. *Language:* AUTOCODER. *Disks or tapes:* 12 7730 tapes. *Special equipment:* Also used IBM 7094 for sorts 12-729. Program is available, IBM KWIC Modified.

References: Study is available from University Microfilms through the Music Educators National Conference, 1401 16th Street, NW, Washington, DC. [*CHum,* May 1970]

M50. Cumulative Index of *Journal of Research in Music Education, Vols. 1-16, 1954-1969*

Principal investigator: Robert W. John, Professor, Dept. of Music, U. of Georgia, Athens, GA 30601.

Scope: Cumulative index,including KWIC index of titles of this journal.

Type of computer: IBM 1401. *Size of storage:* 16k. *Language:* AUTOCODER. *Disks or tapes:* 7 7330 tapes. *Special equipment:* IBM KWIC Index Modified. Program is available. [*CHum,* May 1970]

M51. A Pilot Computer Catalog of Early American Shape-Note Hymn Tunes

Principal investigator: Robert W. John, Professor, Dept. of Music, U. of Georgia, Athens, GA 30601.

Scope: To index and catalog hymn tunes found in (selected) easy American hymn books printed in "shape-notes." *Method:* Hymn books were selected from personal library, and the first 50 hymns of each of ten books were used. Hymns were indexed and classified under the following: Title, melody (Solfeggio), meter, words, key, time signature, number of parts, book and page, position on page—upper or lower.

Type of computer: IBM 1401. *Size of storage:* 16k. *Language:* AUTOCODER. *Disks or tapes:* 4 7330 tapes. [*CHum,* May 1970]

M52. A Pilot Computer Catalog of Early American Sheet Music

Principal investigator: Robert W. John, Professor, Dept. of Music, U. of Georgia, Athens, GA 30601.

Scope: Much of this music is preserved in volumes bound by the original purchaser. Because each volume contains 25-50 pieces of music differing from any other collection, libraries have had serious catalog problems with this type of material. This project was to assist in solving the problem. *Method:* Indexing was done under 14 controls: title, publisher, city of publication, date of publication, composer, lyrics, arranger, original price, plate number of mark, medium, number and description of pages, condition of this copy, shelf list number (volume index) and KWIC index of titles. Twelve bound volumes containing around 500 works were used.

Type of computer: IBM 1401. *Size of storage:* 16k. *Language:* AUTOCODER. *Disks or tapes:* 4 7730 tapes. [*CHum,* May 1970]

M53. Experimental Music Information System

Principal investigator: Robert P. Goldberg, staff member, Group 28, M.I.T. Lincoln Laboratory, Lexington, MA 02173.*

Scope: An experimental system has been developed that permits a user to explore problems involved in the storage and retrieval of musical information. The MUSIC system was written in LISP and designed to run under the CP/CMS timesharing system for the IBM 360/67. This system provides a terminal language for inputting musical themes, a translator to map them into an internal representation that describes them, a retrieval system to identify these themes and an output system to display music on the IBM 2250 scope and print it on the SC-4020 microfilm recorder. *Method:* Experiments have been performed in which the system is used in conjunction with a data base consisting of themes from Mozart's piano concertos.

Type of computer: IBM 360/67, running CP/CMS timesharing system. *Size of storage:* 512k. *Language:* LISP 1.5. *Special equipment:* IBM 2250 scope, Stromberg Carlson SC4020 microfilm recorder. [*CHum,* May 1970]

M54. Computer Music

Principal investigators: Leland Smith, Professor, Artificial Intelligence Project, Dept. of Computer Science; John M. Chowning, Assistant Professor, Dept. of Music; Stanford U., Stanford, CA 94305. *Associate:* David Pool.

Scope: Computer-generated sound; computer-assisted composition; and musical notation for computer input. Music printing by computer; music teaching.

Type of computer: 1) PDP-10, 2) IBM 360/67. *Language:* FORTRAN IV and machine language. *Disks or tapes:* 1) 2 7-track tapes, 2) 8 7-track and 9-track tapes; 1) 4 IBM 3330 disks. *Special equipment:* D-A converter on PDP-10. Complete sound generating program is available for use on PDP-10 with at least 32k available.

M55. Computer-Assisted Musical Composition and Analysis

Principal investigator: Gary Nelson, Assistant Professor, Dept. of Creative Arts, Purdue U., West Lafayette, IN 47907.

Objective: Development of a generalized and transportable computer program for musical composition and analysis. The program (COMPOSE) facilitates operations on musical data by making available a variety of data structures (linked lists, trees, networks, decision tables, description lists, and hierarchies) which are compatible with compositional and analytical thought and with the structural organization of musical scores. *Method:* Functions are provided for the manipulation of COMPOSE data structures so that the programmer is concerned only with the conceptual representation of a structure and not with its internal storage scheme. Input and output functions for each data structure are supplied. Input and output in the data formats of the MUSIC 4 and MUSIC 5 sound synthesis programs and in the DARMS music encoding system are also acceptable to COMPOSE. The source deck of the program is organized in blocks to facilitate insertion or deletion of program modules. COMPOSE is therefore not a fixed program but may be tailored to the needs of individual researchers.

Type of computer: 1) CDC 6500. *Size of storage:* 75k (octal). *Language:* SNOBOL4 Version 3. Program is available.

References: "COMPOSE: An Expansible Computer Program for Structuring Composer-Described Musical Events" (Unpub. Ph.D. dissertation, Washington University, Saint Louis, MO, 1973).

M56. Computerized Analysis and Classification of Asian Musical Style

Principal investigators: Alan L. Kagan, Assistant Professor; Charles T. Brown, Research Assistant; Dept. of Music, U. of Minnesota, Minneapolis, MN 55455.

Scope: A tool for research and teaching in the area of Asian studies, the program will test various computerized forms of analysis. It will provide a classification and locator catalog for oral traditions of Asian musical and spoken literature. Both music and texts will be encoded and analyzed. *Method:* Private and public holdings of recordings, film tracks, and texts will be surveyed and elicited. These will be classified, transcribed, encoded, and analyzed. Printout lists, correlations, distribution maps, and statistical data will be produced.

Type of computer: CDC 6600. *Size of storage:* 65k. *Language:* FORTRAN IV, SNOBOL.

References: "Interim Manual I," *CHINOPERL News,* No.1 (August 1969), pp.30-37. (*CHINOPERL News* is a publication of the China Program at Cornell.) [*CHum,* May 1970]

M57. Chords and Chord Progressions in J.S. Bach's 371 Chorales

Principal investigator: Merton Shatzkin, Professor, Conservatory, U. of Missouri-Kansas City, 4420 Warwick Blvd., Kansas City, MO 64111.

Scope: Encode all pitches, with durations, to discover the frequency of each vertical structure occurring and the frequency of each (second-order) progression thereof, irrespective of durations. Two runs are planned--one that shows voice-crossing and the other that does not (here the "bass" is always the lowest notated pitch, the "tenor" the next-lowest, etc.). Results are not expressed in terms of traditional "chord analysis," but in terms of structures and voice-leading.

Type of computer: IBM 360/30. *Size of storage:* 65k. *Language:* FORTRAN F. *Disks or tapes:* 3 2311 disks; 2 2415 tapes. Program is available. [*CHum,* May 1970]

M58. A Computer-Assisted Study of Ecuadorian Urban Music

Principal investigators: Johannes Riedel, Professor; Charles G. Boody, Research Fellow; Dept. of Music, U. of Minnesota, Minneapolis, MN 55455.

Objective: To examine encoded melodies for 1) scale (or mode) used, 2) melodic movement, 3) basic melodic shape, 4) rhythmic patterns, and 5) phrase structure. This data will then be used to develop a means of differentiating musical subtypes within the repertory. As a side product, an index of the melodies is being generated. Melodies used in the study are drawn from 10 volumes of transcriptions made under the direction of Riedel during his 10 years as a musician in Ecuador.

Type of computer: CDC 6600. *Size of storage:* 131k. *Language:* FORTRAN II and IV, SNOBOL4. Program is available. [*CHum,* November 1970]

M59. A Complete Edition of the Parisian Two-Part Organa

Principal investigator: Hans Tischler, Professor, Dept. of Musicology, Indiana U., Bloomington, IN 47401.

Objective: Chiefly the purpose is to locate all parallel passages in this vast repertory to establish accuracy of reading with regard to rhythm and accidentals and to further clarify the centonization involved in the compositional process. *Method:* Collating and comparing.

Type of computer: CDC 3400/3600.*Size of storage:* 32k. *Language:* FORTRAN 63.*Disks or tapes:* 1 disk, 6 tapes. Program is available from Dr. Theodore Karp, University of California, Davis, CA 95616.

Status: Completed 1972.

References: "On the Need of New Editions of Early Polyphonic Music," *Orbis Musicae* 1 (Tel Aviv, 1971), 67-71; Review of *Jürg Stenzl: Die 40 Clausulae der Hs. Paris 15139 (StV), Journal of American Musicological Society* 25 (1972), 107f; Review of Wulf Arlt: *ein Festoffizium des Mittelalters aus Beauvais, Speculum* 47 (1972), 742f.

M60. Music Notation by Computer

Principal investigator: Donald Byrd, Applications programmer, Wrubel Computing Center, Indiana U., Bloomington, IN 47401.

Objective: To allow computer composing, analysis, etc., programs to produce output in standard music notation. Any symbols used in Western music since the Renaissance, including twentieth-century music employing basically standard notation, will eventually be possible. Ethnomusicological symbols will also be available. *Method:* A program, called SMUT (System for MUsic Transcription), is being written to accept alphanumeric data following certain conventions and draw the music on a digital plotter, or with fairly small modifications, on an electrostatic printer, CRT or any other graphic device. Version 1.1 of SMUT is now available, along with a version of Iannis Xenakis' STOCHOS composing program that produces data for it. In addition, an interface called SMIRK to allow SMUT to accept data in MUSTRAN form (see references) is in progress, and a much more comprehensive version of SMUT is planned. SMUT 1.1 is mostly in ANSI FORTRAN and could easily be converted for most machines.

Type of computer: 1) CDC 6000 or 7000, 2) IBM 360 or 370. *Size of storage:* 1) 28k decimal words, 2) 35k decimal words. *Language:* 1) FORTRAN (mostly ANSI), 2) FORTRAN G or H. *Special equipment:* Digital plotter (CalComp 565). Program is available.

References: "Transcription by Plotter," *Random Bits* 5, 9 (May 1970). (See also M32, M76, and M79.)

M61. Analysis of Rhythmic Performance

Principal investigator: Ingmar Bengtsson, Professor, Institute of Musicology, Uppsala U., Dragarbrunnsgatan 63 III, 753 20 Uppsala, Sweden. *Associates:* Stig-Magnus Thorsén, Krister Malm, Alf Gabrielsson, Dan Malmström, Bengt Edlund.

Scope: Analysis of rhythmic performance (including rhythmic "dialects") particularly the existence and character of systematic variations at different levels. At the first stage the work is concentrated on the duration factor, i.e., systematic variations in this factor ("SYVAR-D"). *Method:* Using different equipment for getting analogous graphs (a melodiwriter and analyzer for multipart music) and for finding time-values for every sound event (attack-attack), the types and patterns of deviations from the fictive regular and mechanical scheme of simple time-value relationships are studied and compared.

Type of computer: CDC 3600. *Size of storage:* 32k. *Language:* FORTRAN, COMPASS. *Disks or tapes:* 2 7-track tapes. Program is available. Contact the Uppsala Datacentral, Uppsala.

References: I. Bengtsson, A. Gabrielsson, and S.-M. Thorsén, "Empirisk rytmforskning" (in Swedish, short summary in English), *Svensk tidskrift för musikforskning* 51 (1969), 49-118; I. Bengtsson, P.-A. Tove, S.-M. Thorsén, "Sound Analysis Equipment and Rhythm Research Ideas at the Institute of Musicology in Uppsala," *Studia Instrumentorum Musicae Popularis* 2 (1972), 53-76; an article about the computer program forthcoming in *Elektronische Datenverarbeitung in der Musikwissenschaft*, vol. 2.

M62. Retrieval of Historical Data from Swedish Music Archives

Principal investigator: Axel Helmer, Director, Swedish Music History Archives, Strandvägen 82, S-115 27 Stockholm, Sweden. *Associates:* Eva Helenius, Aare Mörner, Wolfram Uhlmann.

Objective: Analysis of concerts in Stockholm from 1801 onwards. Listing of composers, performers, types of concerts, works, instruments, keywords, etc. *Method:* Compilation of advertisements, program sheets, reviews; data entered on forms; punched and computed. Biographical, topographical, chronological, and other kinds of lists should be obtained; selective lists can be called for.

Type of computer: IBM 360/45 I. *Size of storage:* 512k bytes. *Language:* FORTRAN, Assembler. *Disks or tapes:* 2 2314 disk packs; 4 2401 tapes.

References: Axel Helmer and Wolfram Uhlmann, "Retrieval of Historical Data. Towards a Computerized Concert Index at the Swedish Archives of Music History," *Fontes Artis Musicae* 16, 1/2 (1969), 48-56. [*CHum,* May 1971]

M63. MUSIC7

Principal investigator: Hubert S. Howe, Jr., Assistant Professor, Dept. of Music, Queens College, CUNY, Flushing, NY 11367.

Scope: Synthesis of electronic music by digital computer and digital-to-analog converter. *Method:* MUSIC7 is a direct descendant of the author's previous MUSIC4BF program, which was in turn a FORTRAN adaptation of MUSIC4B, which was written by Godfrey Winham and the author. MUSIC4B was a revision and expansion of Max Mathews' MUSIC4. MUSIC7, which includes an independent programming language for defining and assembling instruments, is organized as a system procedure library in the Meta-Symbol language. (Separate versions of MUSIC4BF for the IBM System 360 (OS) and the CDC 6600 are also available.)

Type of computer: XDS Sigma 7. *Size of storage:* 65k. *Language:* FORTRAN IV/Meta-Symbol. *Disks or tapes:* 2 tapes. *Special equipment:* Digital-to-analog converter. Program is available.

References: MUSIC7 Reference Manual (New York: Queens College Press). [*CHum,* May 1971]

M64. Computer Investigation of Highland Bagpipe Music

Principal investigator: Michael Ian Shamos, Dept. of Computer Science, Yale U., Dunham Laboratory, 10 Hillhouse Avenue, New Haven, CT 06473.

Objective: To investigate rhythmic and melodic patterns in music for the Highland bagpipe. An eventual outcome of this work will be computer-assisted investigation of the authenticity of *Piobaireachd.* Some of the smaller tasks to be completed before so involved a project is attempted, however, are: 1) Development of a machine-readable notational scheme. The bagpipe is able to produce but nine notes; however, the grace note repertoire is extremely rich, consisting of over 150 distinct patterns. A notation has been developed whereby each note or grace note combination is represented by two characters from the PL/I 60-character set. This 2-character code contains all necessary information on pitch and duration, and is mnemonic in nature. A typical bagpipe tune is representable in its entirety by a 500-byte record. 2)

Creation of a file of *Piobaireachd* and *ceol beag*. This involved the manual transcription of bagpipe music into the machine-readable code. A light tune requires about ten minutes to transcribe, a *Piobaireachd* about half an hour. It is anticipated that over 400 *Piobaireachd* and over 2000 *ceol beag* will eventually be entered into the file. 3) The first study to exercise the system will be a scan of grace note frequencies by type of piece. 4) A set of programs is being developed to translate staff notation to *canntaireachd* (the vocal representation of *Piobaireachd*) and from *canntaireachd* to staff notation. 5) A complete thematic index to bagpipe music will be produced. 6) Statistical study will be made of rhythmic and melodic motives. 7) Selection of "sets": A problem confronting the performer is the selection of pieces that may be played in succession with increasing tempo while still sounding continuous. The selection is based on matching rhythmic and melodic patterns among marches, strathspeys, and reels. Application of the computer to this problem will be studied. A related, but much more difficult problem is the selection of competition medleys. 8) A major goal of the study is to be able to comment scientifically upon the authorship of *Piobaireachd,* much of which is in dispute. 9) Possible other projects will deal with computer composition of music for the bagpipe and the relation between a pipe tune and its accompanying drum setting. *Method:* Once a manageable file is created, most of the studies will be done by statistical methods using as data counts of various patterns found in the music. Bagpipe music is relatively simple, as no harmonic elements are present (other than that of the relation of chanter notes to the drones, which is largely subjective). 10) Publication of a complete concordance to the *Nether Lorn Canntaireachd.* 11) Use of a program to develop accurate settings of *Piobaireachd* in staff notation (suggested by Dr. Patrick Terry, University of Cambridge).

Type of computer: 1) IBM 360/67, 2) PDP-10. *Size of storage:* 1) 100k, 2) 32k. *Language:* 1) PL/I F, SNOBOL4, 2) SNOBOL4.

M65. Music Synthesis by Computer [Deleted] [*CHum,* May 1971]

M66. Rhythmic Problems and Melodic Structure in Organum Purum: A Computer-Assisted Study

Principal investigator: Raymond F. Erickson, Assistant Professor, Dept. of Music, Queens College, CUNY, Flushing, NY 11367. *Associate:* William G. Waite.

Objective: 1) To develop a comprehensive and rigorous methodological framework for a computer-assisted study of the medieval repertory of Notre Dame organa dupla as preserved in three major manuscript sources. 2) To perform computer-assisted stylistic analysis of the encoded music to settle questions regarding the use of modal rhythm in the duplum and the extent of melodic borrowing as a compositional technique. *Method:* 1) Some 30 Office organa were encoded in a DARMS-like representation to preserve all paleographic evidence, e.g., the original note forms, variations in size and shape of note heads, the vertical alignments between tenor and duplum as written by the scribe. 2) The encoded data were checked for both encoding and information errors and corrections were made. Three programs were developed for this purpose. 3) The corrected data strings were converted into tables for fast retrieval purposes. 4) A tenor realignment program adjusted the MS vertical alignments on the basis of medieval musical criteria. 5) A phrase-structure analysis program yielded output from which conclusions were drawn concerning the musicological purposes of the investigation.

Type of computer: 1) IBM 360/50, 2) IBM 7094/7040 Direct Coupled System. *Language:* 1) PL/I F, 2) SNOBOL3, FORTRAN IV. A generalized version of the table-driven compiler used for step 3 of Method is available.See reference 1 below.

References: 1) "A General-Purpose System for Computer-Aided Musical Studies," *Journal of Music Theory* 13, 2 (1969), 276-294, 2; (to appear in a volume of computer studies ed. by Harald Heckmann [Kassel]): "Methodological Problems and Some Solutions in a Computer-Assisted Study of Notre Dame *organum duplum.*" [*CHum,* May 1971]

M67. DUPACK (DARMS Utility Package)

Principal investigator: Raymond F. Erickson, Assistant Professor, Dept. of Music, Queens College, CUNY, Flushing, NY 11367. *Associates:* Stefan Bauer-Mengelberg, David Gomberg, Anthony Wolff.

Scope: To provide a program package for musicological research based on DARMS, an encoding system for music invented by Stefan Bauer-Mengelberg, the canonical form of which has been defined by David Gomberg. It will include three programs: 1) a syntax-checker to detect invalid characters; a (user-)DARMS to Canonical-DARMS translator (designed and implemented by Anthony Wolff; cf. M74); and 3) a string-to-tables conversion program to transform complicated Canonical DARMS data strings into easy-to-use tables that can be operated on by user-written musical analysis programs in any language. *Method:* No backup syntax-directed analysis techniques and PL/I optimization factors will be employed to make the syntax checker as efficient as possible.

Type of computer: IBM 360 and IBM 370. *Language:* PL/I F.

References: Raymond F. Erickson, "Rhythmic Problems and Melodic Structure in Organum Purum: A Computer-Assisted Study," chap.IV (Unpub. Ph.D. dissertation, Yale University, 1970); Raymond F. Erickson, "The Ford-Columbia Music Representation (DARMS)," an informal but complete description of the encoding system, available from the author.

M68. Katalog *Deutsches Tenorlied des 16. Jahrhunderts*

Principal investigators: Harald Heckmann, Director; Norbert Böker-Heil, Deutsches Rundfunkarchiv, D6 Frankfurt am Main, Bertramstrasse 8, West Germany. *Associate:* Ilse Kindermann.

Scope: An automatically compiled catalog of the *Deutsche Tenor-Lieder* that will contain indices of the following items; 1) all single-voice incipits according to their intervallic structure (a diastematic-lexicographical "DIALEX"-index involving the immediate succession of melodically concordant entries), 2) concordances and variants of the polyphonic settings, 3) sources, 4) text incipits, 5) names (composers, editors, etc.). The inclusion of a computer-generated index of polyphonic incipits written in the original, i.e., White Mensural Notation, is projected. *Method:* Musical incipits and alphameric information are transcribed on a special coding form and then keypunched. For the music a newly developed input language, the WMN-Code, is used (cf. M69). Together with the already working DIALEX program, several sorting programs will provide the printed output for the indices. A WMN-GRAPHIC system now in preparation is expected to reconvert the keypunched music to mensural notation (on microfilm).

Type of computer: 1) IBM 360/65, 2) UNIVAC 1108, 3) UNIVAC 9300. *Size of storage:* 1) 768k bytes, 2) 65k (36-bit words). *Language:* 1) FORTRAN IV, 2) FORTRAN V. *Disks or tapes:* 1) 2 9-track tapes, 2) 2 7-track tapes. *Special equipment:* Stromberg Datagraphics SD 4060 Microfilm Recorder.

Status: From 61 printed and manuscript sources up to now almost 6000 polyphonic incipits have been encoded, keypunched, and stored on magnetic tape. INKOMP, a program which provides the alphabetical list of composers, has been successfully tested. Each composer's name is followed, in alphabetical order, by the text incipits of his works. Various forms of a name produce their own entries, with references to the normalized spelling. The WMN-GRAPH program is ready to decode the keypunched material for both convenient proofreading of the data and final cataloging of the polyphonic incipits. Its output closely resembles sixteenth-century music calligraphy.

References: "Weisse Mensuralnotation als Computer-Input und -Output," *Acta Musicologica* 43, 1-2 (1971), 21-33.

M69. Programmierte Musikanalyse

Principal investigators: Norbert Böker-Heil; Harald Heckmann, Deutsches Rundfunkarchiv, D6 Frankfurt-am-Main, Bertramstrasse 8, West Germany. *Associate:* Ursula Böker-Heil.

Scope: Polyphonic mass settings of the fifteenth century are analyzed with respect to personal and/or regional traits of style. Principally the investigation aims at the relevance of musical microstructures and their interdependencies. At a final stage an attempt will be made to define strictly certain stylistic features in terms of information theory. *Method:* Data is prepared by transcribing the music to WMN-Code, i.e., a numerical representation of the graphic features of White Mensural Notation. A conversion program recalculates the pitches and evaluates the mensuration signs and rhythmic symbols in such a way that every contrapuntal, melodic, or rhythmic configuration can be referred to by an arithmetic or logical expression. The entropies of single and grouped patterns will be considered the measuring units for stylistic significance. An attempt will be made with a "learning" program, which on a stochastic basis defines and subsequently tests its own hypotheses on the informative value of musical microstructures.

Type of computer: 1) IBM 360/65, 2) UNIVAC 1108, and 3) UNIVAC 9300. *Size of storage:* 1) 768k bytes, 2) 65k (36-bit words). *Language:* 1) FORTRAN IV, 2) FORTRAN V, 3) UNIVAC 1108 Assembly language. *Disks or tapes:* 1) 2 9-track tapes, 2) 2 7-track tapes. *Special equipment:* Stromberg Datagraphics SD 4060 Microfilm Recorder. CalComp 763 plotter.

Status: A test set of some 60 polyphonic Kyrie settings from the Trent Codices has been prepared for computation. Analysis, on a medium level of abstraction, is now carried out by two programs which present their results in the form of plotted diagrams.

M70. Instrumental Composition: Project 2

Principal investigator: Gottfried Michael Koenig, Artistic Director, Institute of Sonology, Utrecht State U., Plompetorengracht 14-16, Netherlands. *Associate:* G. Vermeulen.

Scope: The composition of instrumental music according to data and combinatorial rules given by the composer on the basis of existing sub-programs. Compositional-theoretical investigation of 1) the programability of musical structures, 2) the adaptability of composers to mechanical realization of structures by the computer. *Method:* The composition of musical structures with the use of a complete program. Data input by means of a questionnaire to be filled in by the composer (investigator). Input data and results preserved for later musicological analysis.

Type of computer: Electrologica X8. *Size of storage:* 24k. *Language:* ALGOL. *Disks or tapes:* 1 disk. Program is available.

References: Electronic Music Reports 3 (1970), Institute of Sonology, Utrecht State University. [*CHum,* November 1971]

M71. Harmonic and Phraseologic Computer-Aided Music Analysis

Principal investigator: Claudio Stephan, Chairman, Grupo de Pesquisas Musicais (Musical Research Group), Rua 15 de Novembro, 607-ap.84, 04709 Sao Paulo, Brazil. *Associates:* Antonio C. Amancio, Antonio C.F. Campos, Arthur N. Campos, Bianca M. Mazza, Clarisse F. Stephan, Magaly M. Oliveira.

Objective: To analyze selected music scores, harmonically and phraseologically, and edit them revised and commented, rendering them more easily understood by students and performers. *Method:* Study of music theories through the best books used in the most important universities, aiming at the compilation of a new treatise on music theory, as comprehensive as possible, to supply the theoretical basis for the elaboration of a computer program for the desired analysis. [*CHum,* November 1971]

M72. Development of a Standard Orchestra for Computer-Generated Sound

Principal investigator: Reg Boudinot, doctoral candidate, Dept. of Music Theory-Composition, School of Music, U. of Miami, Coral Gables, FL 33124.

Objective: To define the standard non-percussion orchestral instruments for use with the MUSIC V sound-generating program.

Type of computer: IBM 370/155. *Size of storage:* 150k. *Language:* FORTRAN IV G (Release 20); Assembler F (Release 20). *Disks or tapes:* 2 2314 disks, 1 7-track tape. MUSIC V is available from Bell Labs, instrument definitions and subroutines from the investigator. [*CHum,* November 1971]

M73. Automatic Music Printing

Principal investigator: David A. Gomberg, Assistant Professor, Dept. of Computer Science, SUNY at Buffalo, 4226 Ridge Lea Road, Amherst, NY 14226.

Objective: To input any music manuscript encodable in the Ford-Columbia Music Representation (DARMS) and output directions to an appropriate printing device (currently the Photon printer). The programs will accomplish all layout work (page and line breaks, allowing for page turns, etc.), part extraction and printing, transposition, etc.

Type of computer: IBM 360. *Language:* PL/I F. *Special equipment:* Photon printer.

M74. Development of a DARMS-to-Canonical DARMS Translator

Principal investigator: Anthony B. Wolff, Research Associate, School of Advanced Technology, SUNY, Binghamton, NY 13901. *Associate:* Stefan Bauer-Mengelberg.

Objective: For a given piece of music, the DARMS (Ford-Columbia) representation offers a nearly infinite set of possible encodings, some much less compact than others. Selection of DARMS options in any instance is governed by the structure of the music being encoded. For example, certain options facilitate polyphony while others facilitate homophony. While linguistic complexity is advantageous to the encoder, it greatly complicates the task of processing the code. The purpose of the present program is to translate input DARMS into Canonical (or complete) DARMS, thus guaranteeing a unique representation for a piece of music. *Method:* The canonizer is being written in two phases: the first, a syntax analyzer constructed of a network of finite state automata whose state transitions drive the second phase, and the second, a group of subroutines which perform the actual transformations on the input stream, resulting in Canonical DARMS. Maximum machine independence is being built into the canonizer. It is intended that this program ultimately be associated with DUPACK (DARMS Utility Package).

Type of computer: IBM 370/155. *Language:* PL/I F. *Disks or tapes:* IBM 3330 disks, IBM 3420 tapes.

References: Stefan Bauer-Mengelberg, "The Ford-Columbia Input Language," *Musicology and the Computer,* ed. Barry Brook (New York: The City University of New York Press, 1970), pp. 48-52; Raymond F. Erickson, "Musical Analysis and the Computer," *Computers and the Humanities* 3, 2 (November 1968), 84-104; Harry B. Lincoln, "Some Criteria and Techniques for Developing Computerized Thematic Indices," *Elektronische Datenverarbeitung in der Musikwissenschaft,* ed. Harald Heckmann (Regensburg: Gustav Bosse Verlag, 1967), pp.57-62.

M75. Afro-American Religious Music, 1619-1861: An Analytical Study

Principal investigator: Portia Katrenia Maultsby, Lecturer, Dept. of Afro-American Studies, Indiana U., Bloomington, IN 47401.

Scope: A series of analysis programs has been applied to 100 Negro spirituals. These programs include interval counts and percentages, scale analysis, identification of existing 3- and 4-note interval sequences, evaluation of common and selected interval sequences, and some rhythmic analysis. This analysis program, designed by the investigator in conjunction with Jerome Wenker, was used in the writing of the investigator's Ph.D. dissertation. *Method:* Music data is prepared using the MUSTRAN notation (see M32) and converted into an internal format using the MUSic TRANslator. The interval format is saved on disk or tape and is then available for the individual analysis programs. Some analysis programs count occurrences of selected factors (intervals, interval patterns, scale runs, etc.) by song, group of songs, and corpus. These results are then available for statistical analysis. Other analysis programs are strictly concerned with song analysis (number and length of phrases, table of pitch versus note duration, etc.) and do not produce data appropriate for statistical analysis.

Type of computer: UNIVAC 1108. *Size of storage:* 45k. *Language:* FORTRAN, ASM. *Disks or tapes:* 1 disk, 1 tape. Program is available. Write to investigator or Jerome Wenker, 11 Carlton Club Drive, Piscataway, NJ 08854.

References: "Seeks Computer Derivation of Musical Roots," *MACC Newsletter* 6, 5, University of Wisconsin Academic Computing Center, 1972. [*CHum,* November 1972]

M76. MUSTRAN II—An Extended *MUS*ic *TRAN*slator

Principal investigator: Jerome Wenker, 11 Carlton Club Drive, Piscataway, NJ 08854.

Scope: An extension of the MUSTRAN translator (see M32 and M45) is in process of being developed. This notation includes all known symbols of common music notation, all known symbols used in ethnomusicology, all known graphemic notations used in various projects for computer printing of music (see, in particular, M60 and M67), additional graphemic notations which appear to be advantageous, and selected conventions advantageous for automated music performance. During translation, additional information is automatically supplied by the computer, thus simplifying procedures to analyze, print, or perform music. It must be emphasized that MUSTRAN II is fully upward compatible with MUSTRAN and that no data correctly prepared for MUSTRAN will have to be re-encoded for MUSTRAN II (unless MUSTRAN II data is to be added). Similarly, data encoded for MUSTRAN II that does not use any of the purely MUSTRAN II conventions can correctly be processed by the current music translator. *Method:* The set of symbols to be included is, at present, about 95 percent complete. For almost all of these the encoding notations have been designed, and the modifications to MUSTRAN have been flowcharted and coded. The final implementation awaits access to computer facilities. The MUSTRAN II notation conventions will not be formally released until the working program proves their validity. Until then the existing MUSTRAN programs, which are very stable, are available for IBM, CDC, and UNIVAC equipment (see M32 and M45). [*CHum,* November 1972]

M77. A Music Performance Program

Principal investigator: Jerome Wenker, 11 Carlton Club Drive, Piscataway, NJ 08854.

Scope: A series of programs is being produced which will accept as input the output of MUSTRAN (or MUSTRAN II) and output data which will be appropriate to Max Mathews' MUSIC4 program as modified and extended by John Gardner at UCLA (MUSIC 5B). The result is that music encoded in the mnemonic human-oriented notation of MUSTRAN can, without further human intervention, be processed to produce an audio tape containing the prescribed music. Thus, to produce a sound pattern consisting of a quarter note on G, a dotted eighth on B, and a sixteenth note on A, all in the treble clef, with a metronome count of a quarter note equals 80, the notation GS, M4=80, 4G, 8B., 16A, will supply all required information for the initial program. *Method:* The existing MUSTRAN and the forthcoming MUSTRAN II programs will be used to process the input data. The existing versions of MUSIC4 will be used to produce the final audio tape. The intermediate processing will be done by a separate program in process of development. [*CHum,* November 1971]

M78. An Analysis of Transcriptions

Principal investigator: Jerome Wenker, 11 Carlton Club Drive, Piscataway, NJ 08854. *Associates:* James Porter, Michael Moore.

Scope: The transcription of the music of a folksong is a descriptive notation of how the transcriber heard the tune. Through use of the computer and Dr. Seeger's melograph, an attempt is being made to minimize the subjectivity of this process. *Method:* A human-produced transcription of a folktune will be encoded in MUSTRAN notation and processed on the computer through an intermediate program and through MUSIC4 to produce an audio tape which precisely performs the descriptive notation of the transcriber. Both the original tune being transcribed and the generated audio tape will be processed by the Melograph at UCLA, which will produce a visual graph of each version. By comparing the two graphs it will be possible to modify the transcription to obtain a new transcription which when processed by the computer and the melograph will produce a graph that is closer to the graph of the original tune than the original transcription. This process will be repeated until the two graphs are sufficiently similar that no further modifications will be productive. At this point a detailed, objective transcription of the tune can be analyzed. With such detailed transcriptions, it is anticipated that it will then be possible to begin work on an objective description of the individual

"style" of a performer. *Special equipment:* Melograph at UCLA. For further details see M35, M76, and M77. [*CHum,* November 1972]

M79. Musical Graphics

Principal investigators: Donald Byrd, Programming Consultant, Research Computing Center, Indiana U., Bloomington, IN 47401; Jerome Wenker, 11 Carlton Club Drive, Piscataway, NJ 08854.

Objective: To use the computer to produce sheet music in photo-ready form for publication or performance. The goal is to have the computer automatically produce commercially satisfactory sheet music from data prepared using a mnemonic, human-oriented notation with a minimum of graphemic symbols. Of course, the facilities will be available so that the print format can be completely specified if that is desired. *Method:* The music to be processed is first encoded using an expanded form of the MUSTRAN notation (see references). The MUSTRAN translator then converts the data into a compressed binary form suitable for analysis, performance,and, as here, music printing. A second processor, called SMUT, will use the output of the MUSTRAN translator to format and print sheet music using the CalComp plotter as the output device. The actual printing routines will be separated from SMUT so that it should be fairly easy to replace them in order to produce output on film or on any other device.

Type of computer: CDC 6600. *Size of storage:* 120k (octal). *Language:* FORTRAN IV. *Disks or tapes:* 1 disk, 1 7-track tape. *Special equipment:* CalComp plotter. MUSTRAN is currently available for IBM 360 and IBM 370; UNIVAC 1106, 1107, and 1108; and CDC 6600. SMUT is available for CDC 6600.

References: (MUSTRAN) Jerome Wenker, "A Computer Oriented Music Notation Including Ethnomusicological Symbols," *Musicology and the Computer,* ed. Barry S. Brook (New York: The City University of New York Press, 1970)—see also M32 and M45; (SMUT) Donald Byrd, "Transcription by Plotter," *Random Bits* 5, 9 (May 1970). (See also M60.) [*CHum,* November 1972]

See also G18, G23, L429, L506, S54.

Philosophy

P1. Index Thomisticus

Principal investigator: Roberto Busa SJ, Professor, Dept. of Philosophy, Aloisianum College, Aloisianum 21013 Gallarate, Italy.

Scope: Indexes and concordances of the 118 works of St. Thomas Aquinas and 61 works of other authors (ninth-sixteenth centuries) connected with him; 1,700,000 text lines of Latin containing 10,600,000 words; at the level of forma and lemma, with frequency counts, etc. *Method:* Pre-editing; lemmatizing by Latin machine dictionary; automatic construction of context; photocomposition. Computer analysis enlarged to other Latin, Greek, Hebrew, Russian, German, and Italian texts: mainly Latin Bible, Forcellini Lexicon Totius Latinitatis, Qumran nonbiblical Scrolls, one work of Goethe, one of Kant, scientific Russian abstracts, including 150,000 Latin forms coded for computer use; 25,000 Latin lemmas coded for automatic inflection; 6000 Latin endings coded for computer use, etc.

Type of computer: 1) IBM 1401, 2) IBM 7090, 3) IBM 360. *Language:* AUTOCODER, Aiocs, Assembler, FORTRAN, PL/I. *Disks or tapes:* 9- and 7-track tapes; 3) 4 disks. Program is available.

References: Complete list available from investigator upon request.

P2. A Concordance to Plato

Principal investigator: Leonard Brandwood, Lecturer, Dept. of Classics. U. of Manchester, Manchester 13, England. *Associates:* R.B.E. Napper, Lorne H. Bouchard.

Scope: Inflected forms listed under their base lexical form. Classification of different usages according to grammatical and syntactic criteria. *Method:* 1) Pre-edit text to distinguish ambiguities (homographs, etc.). 2) Punch paper tape on teleprinter (Greek characters). 3) Code onto magnetic tape. 4) Output of vocabulary in alphabetical order on paper tape. 5) Alphabetical order revised in accordance with standard grammatical paradigms. 6) Output of concordance according to combined alphabetical and grammatical order. These stages have been completed, and the output is now being edited manually for publication.

Type of computer: Atlas I. *Size of storage:* 16k. *Language:* ALGOL, FORTRAN IV, Atlas Autocode. *Disks or tapes:* 8 tapes.

Status: An index to the Burnet text of Plato was expected to reach completion by the end of 1970 or early 1971, to include references by Stephanus, and page, section and line number to every occurrence of each word. Entries were to be arranged according to the paradigms of the traditional grammar, the more important variant readings included from the critical apparatus both of Burnet and of subsequent editions (e.g., Budé). Words in quotations noted as such.

References: "Analysing Plato's Style with a Computer," *Bulletin of the Institute of Classical Studies* (London, 1956). [*CHum,* November 1970]

P3. Kierkegaard's Pseudonyms

Principal investigator: Alastair T. McKinnon, Professor, Dept. of Philosophy, McGill U., Montreal 2, Canada. *Associates:* Norman Swartz, Roger Webster.

Scope: 16 samples (400,000 words) from Kierkegaard's pseudonymous and "acknowledged" writings. *Method:* Various statistical comparisons of the pseudonymous works with the established Kierkegaard norm.

Type of computer: IBM 7044. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 6 729 Model V, 3 729 Model G tapes. [*CHum,* May 1968]

P4. Concordance to Kierkegaard's *Samlede Vaerker*

Principal investigator: Alastair T. McKinnon, Professor, Dept. of Philosophy, McGill U., Montreal 2, Canada. *Associate:* Roger Webster.

Scope: Text of approximately 5000 pages (*Samlede Vaerker,* 3rd ed.). *Method:* 1) Punch text. 2) Run various proofreader programs. 3) Establish page and line correlations between the Danish third edition and the Danish second and the English, French, and German translations.

Type of computer: IBM 7044. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 6 729 Model V, 3 729 Model G tapes. [*CHum,* May 1968]

P5. Concordance to Wittgenstein's *Philosophische Untersuchungen*

Principal investigator: Alastair T. McKinnon, Professor, Dept. of Philosophy, McGill U., Montreal 2, Canada.

Scope: Text of approximately 230 pages (3rd ed.). *Method:* 1) Punch text. 2) Proofread.

Type of computer: IBM 7044. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 6 729 Model V, 3 729 Model G tapes. [*CHum,* May 1968]

P6. Theology on the Canadian Frontier

Principal investigator: Alastair T. McKinnon, Professor, Dept. of Philosophy, McGill U., Montreal 2, Canada. *Associates:* Arnold Smith, Roger Webster.

Scope: Data is 50 statistically typical hymns from the "Canadian frontier." *Method:* Writing of word counts and a line concordance to serve as a basis for an interpretative essay on the subject.

Type of computer: IBM 7044. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 6 729 Model V, 3 729 Model G tapes. [*CHum,* May 1968]

P7. Allgemeiner Kantindex

Principal investigators: G. Martin, Professor, Philosophisches Seminar A der U. Bonn, 53 Bonn, Adenauerallee 98a, West Germany; G. Ungeheuer, Professor, Institut für Phonetik und Kommunikationsforschung, 53 Bonn, Am Hof, West Germany. *Associates:* D. Krallmann, H.A. Martin.

Scope: Word indexes, name index, concordances and index of references for the complete works of Imanuel Kant based on the Academic Edition, Berlin. *Method:* 1) Keypunch. 2) Transfer to tapes. 3) Produce concordances. 4) Text statistics. 5) Grammatical analysis. 6) Sort by grammatical classes and patterns. 7) Produce word index and index of references. 8) Material computer-typeset.

Type of computer: IBM 370/165.

Status: The grammatically classified text location index to Academy Edition volume 1 has been completed and will be available in the near future. The text location indexes to Academy Edition volumes 2-9 are in preparation. Project is interrupted at present.

References: G. Martin (Hrsgb.), *Allgemeiner Kantindex,* zweite Abteilung, Bd.16, 17; *Wortindex* (bearb.: D. Krallmann u. H.A. Martin), Berlin 1967, 1968; G. Martin, "Der Allgemeine Kantindex und seine Philosophischen Probleme," *Kantstudien* 60 (1969), 198-201; H.A. Martin, "Der Allgemeine Kantindex und seine Elektronischen Probleme," *Kantstudien* 60 (1969), 201-203; D. Krallmann, "Linguistische Aspekte des Kantindex," *Kantstudien* 60 (1969), 203-209; Autorenkollektiv, "Untersuchungen zur Sprache Kants," *IPK—Forschungsbericht* 26 (1970), Verlag

Helmut Buske, Hamburg; W. Lenders, "Bedeutungsanalyse philosophischer Begriffe—Ein Problem des Kantindex," *Revue Internationale de Philosophie* (1973).

P9. The Actual Use of Selected "Philosophical" Words in Present-Day Edited American English Prose

Principal investigator: Herbert Guerry, Dept. of Philosophy and Religion, Box 80, Pembroke State U., Pembroke, NC 28372.

Objective: To determine the actual use of various words or concepts that are discussed frequently by philosophers. *Method:* Examples of use are being taken from the million-word Brown University *Standard Sample of Present-Day Edited American English.* These data are then used as a basis for generalizations about actual use. The use of a large sample of prose helps to avoid the idiosyncrasies that often creep into linguistic analyses.

Type of computer: IBM 360/20. *Size of storage:* 16k. *Language:* MOD 20 BAL. *Disks or tapes:* 2 EBCDIC, 9-track tapes; 2 IBM 2311 disks. Program is available from Department of Linguistics, Brown University.

Status: Completed 1971.

P10. Philosophy Documentation Center

Principal investigator: Richard H. Lineback, Professor, Dept. of Philosophy, Bowling Green State U., Bowling Green, OH 43403.

Objective: 1) To publish *The Philosopher's Index: An International Index to Philosophical Periodicals.* The index is processed by computer and covers approximately 250 journals. 2) To operate the "Philosopher's Information Retrieval System" (PIRS). This system, now in the experimental stage, is designed to provide bibliographies on any topic or author related to philosophy. 3) To publish the *Directory of American Philosophers,* a handbook on philosophy teachers in the U.S. and Canada with information on available fellowships and assistantships. Additional information on philosophical societies, journals, and publishers of philosophy works. 4) To publish the *International Directory of Philosophy and Philosophers,* a handbook on philosophy teachers throughout the world except for the U.S. and Canada. Additional information on philosophical institutes, philosophical societies, journals, and publishers of philosophical works. 5) To publish a series of bibliographies of famous philosophers.

Type of computer: IBM 360/75. *Size of storage:* 512k. *Language:* PL/1 F. *Disks or tapes:* 8 2314 disks, 4 2401 tapes. *Special equipment:* 6 2260 terminals, 10 2741 terminals. Program is available.

P11. Philosophy Information Retrieval System

Principal investigator: Richard H. Lineback, Professor, Dept. of Philosophy, Bowling Green State U., Bowling Green, OH 43403.

Objective: To write a program which will efficiently retrieve bibliographic data and abstracts in the area of philosophy. User may stipulate: 1) period of time he is interested in, for example, 1950-1960, 2) author(s), 3) languages he can read, 4) journal(s) that he is interested in searching, and 5) any logical combination of subject headings.

Type of computer: IBM 360/75. *Size of storage:* 512k. *Language:* PL/I.

P12. Index Descartes

Principal investigators: Pierre Costabel, Directeur d'études, Dept. d'Histoire des Sciences, E.P.H.E.VI; Jean-Robert Armogathe, Maître-assistant, Dept. d'Histoire du Catholicisme moderne, E.P.H.E. V; (for both) Maison des Sciences de l'Homme, 54 Boulevard Raspail, Paris CEDEX 06, France. *Associates:* Th. Bert, R. Bragues, J.P. Deschepper, J. Lebrun, J.L. Marion.

Scope: Index verborum and index of concepts of Descartes. Already available: *Discours de la Méthode* and *Essais, Méditations, Regulae ad Directionem Ingenii.* Projected: Latin text of the *Méditations. Method:* Search for frequencies and for lexical and functional coefficients: statistical calculations on co-occurrences. Utility: indexing and problems of interpretation (historical, dating, stylistic, philosophical).

Type of computer: 1) IBM 1410, 2) IBM 360/75. *Language:* 1) and 2) PL/I. Program is available. Write to Centre de Calculs Sciences Humaines, Maison des Sciences de l'Homme.

References: J.R. Armogathe, "Saisie et maniement de l'information dans le *Discours de la Méthode," Les applications de l'informatique aux textes philosophiques* (Paris: C.N.R.S., 1971), pp. 70-75; J.L. Marion, *Revue Internationale de Philosophie* 4 (1972).

P13. Computerized Storage and Dialog-Retrieval of Philosophical Information

Principal investigators: Alwin Diemer, Professor, and Norbert Henrichs, Dept. of Philosophy, Düsseldorf U., D-4 Düsseldorf, Horionplatz 10, West Germany.

Scope: User-oriented bibliographical Dialog-Retrieval (special version of Siemens' GOLEM), working on the basis of "Textword" abstracts and signified contextual relations.

Type of computer: Siemens 4004/45. *Size of storage:* 256k. *Language:* Assembler.

References: N. Henrichs, "Bericht über das Philosophy Information center an der Universität Düsseldorf," *Conceptus* 4 (1970), 133-144; N. Henrichs, "Philosophische Dokumentation, Literatur-Dokumentation ohne strukturierten Thesaurus," *Nachrichten für dokumentation* 21 (1970), 20-25; "Literatur-Dokumentation mit GOLEM," SIEMENS-Schriftenreihe "data praxis," D-8 München 70, Hofmannstr.51, Bestell-Nr. D14/4279.

P14. ALBUM: Retrieval System for Bibliographies, Indices, and Catalogs

Principal investigators: Alwin Diemer, Professor; Norbert Henrichs, Dept. of Philosophy, Düsseldorf U., D-4 Düsseldorf, Horionplatz 10, West Germany.

Scope: ALBUM (Anwendungsvariables, literaturauswertendes, bibliographiengenerierendes, unitermakkumulierendes Maschinenprogramm) is an "object-oriented" retrieval system for computer printing of bibliographies, journal indices, or library catalogs. Names and subjects can be listed in their contextual relations.

Type of computer: Siemens 4004/45. *Size of storage:* 64k. *Language:* Assembler.

References: N. Henrichs and H. Rabanus, "ALBUM—ein Verfahren für Literatur-Dokumentation," *Siemens Schriftenreihe Data-Praxis* 026-1. [*CHum,* May 1971]

P15. A Critical Edition of the Papers of the Metaphysical Society, 1869-1880

Principal investigator: J.S. North, Assistant Professor, Dept. of English, U. of Waterloo, Waterloo, ON, Canada. *Associate:* Alan Willard Brown.

Objective: To collect all 96 papers delivered to the Metaphysical Society during its lifetime and provide biographical and bibliographical information about each contributor, as well as annotations for each paper. *Method:* All the material will be stored using the Administrative Terminal System, which facilitates proofreading and manipulation of editorial material. The completed magnetic tape will then be sent out for photocomposition.

Type of computer: IBM 360/75. *Size of storage:* 3000k. *Language:* ATS. *Special equipment:* 2741 Communication Terminal for entering data. Program is available.

References: Alan Willard Brown, *The Metaphysical Society, Victorian Minds in Crisis, 1869-1880* (New York: Columbia University Press, 1947). [*CHum,* May 1971]

P16. Materials for the History of Ethical Terminology

Principal investigator: L.V. Berman, Associate Professor, Humanities Special Programs, Stanford U., Stanford, CA 94305. *Associate:* R. Finkel.

Objective: To determine 1) the development of ethical terminology within one language such as Hebrew, Arabic, Latin, and others; 2) whether ethical terms take on different meanings when translated from one language to another and, if they do, to find out in what ways they differ. *Method:* Various Hebrew medieval manuscript versions of Averroes' *Middle Commentary on the Nicomachean Ethics of Aristotle* are being put into machine-readable form, divided into the lines of the Bekker edition of Aristotle. An index of ethical terminology is then to be prepared.

Type of computer: IBM 360/67. *Language:* SNOBOL4. *Special equipment:* Hebrew typing element for 2741 terminal. Program is available.

References: "Preliminary Report on a Computer-Aided Critical Edition of the Hebrew Versions of Averroes' *Middle Commentary on the Nicomachean Ethics," Hebrew Computational Linguistics,* Bulletin No. 6 (Bar-Ilan University, Israel).

P17. Moralia of Gregory the Great: Index of Terms and Concepts

Principal investigators: Gerhart B. Ladner, Dept. of History, U. of California, Los Angeles, CA 90024; Michael Phelps, Nasson College, Springvale, ME 04083.

Objective: To study key terms and concepts, important for the history of ideas, and their association with other terms, in Gregory the Great's *Moralia,* and to compare their use with that in other authors.

Type of computer: FORTRAN IV. *Disks or tapes:* 9-track tapes.

References: "Gregory the Great and Gregory VII: A Comparison of Their Concepts of Renewal, with a Note on the Computer Program Used by David W. Packard," *Viator IV* (University of California Press, 1973).

P18. Automatic Collation of Buridan's Questions on the Metaphysics

Principal investigator: Penny Gilbert, Dept. of Romance Languages, U. of Manitoba, Winnipeg, Canada. *Associate:* J. Colin McConnell.

Objective: To prepare a critical edition of Buridan's long questions on the metaphysics by having the computer compare the manuscript texts and locate the variants. *Method:* Programs are modular. Texts are input on a typewriter terminal, one line per card image. The first module breaks the text into single word records. The second module works on two texts at a time and compares them to locate variants which are saved in a variant file. After all manuscripts have been compared, the variant files are merged and sorted into correct order. The next module prints hard copy. Working with this printout, the editor prepares data for the update module in which insignificant variants are deleted and citations added. A final module will print the edition in standard scholarly format.

Type of computer: IBM 360/65. *Size of storage:* 85k. *Language:* PL/1 F. *Disks or tapes:* 3 9-track tapes. *Special equipment:* EMRF print chain. [*CHum,* November 1972]

Social Sciences

S1. Statistical Analysis of the Socialist Underground in Nazi Germany, 1933-1939

Principal investigator: William S. Allen, Associate Professor, Dept. of History, U. of Missouri, Columbia, MO 65201.

Objective: To determine average characteristics of persons active in the Socialist underground. *Scope and method:* Cross referencing and evaluating 400 biographies from court cases.

Type of computer: IBM 7040. *Language:* FORTRAN IV.

References: "Eine Statistische Analyse der socialdemokratischen Widerstand in Nordrhein-Westfalen: Vorläufige Bericht," paper delivered for Friedrich-Ebert Stiftung Conference, Bad Godesborg, September 1966. Published in *Forschungsberichten,* vol. 2, 1967. [*CHum,* May 1967]

S2. An Index to the Evidence for Travel and Communication in Antiquity

Principal investigator: Eugene N. Borza, Associate Professor, Dept. of History, Pennsylvania State U., University Park, PA 16802.

Objective: To develop a work listing all ancient literary references to travel time and routes as an aid for scholars with chronological problems, the solutions of which depend on the speed and circumstances of travel. *Method:* The references will be noted according to a predetermined code, and the computer will be asked to classify these and index them in a manner which will provide ease of use.

Type of computer: IBM 360/67. *Language:* FORTRAN H. *Disks or tapes:* 2 2400 tapes and 2311 disks. Program is available.

Status: Pilot study, showing work completed on Herodotus, Thucydides, Xenophon, Livy, and Strabo is now complete.

References: Travel and Communications in Classical Times. A Guide to the Evidence, A Pilot Study (University Park, PA, 1969).

S3. Congressional Voting Patterns and the Kansas-Nebraska Act

Principal investigator: Gerald Wolff, Associate Professor, Dept. of History, U. of South Dakota, Vermillion, SD 57069.

Objective: To test sectionalism and party affiliation on this issue. *Method:* Guttman Scalogram Analysis.

Type of computer: IBM 7040. Program is available at University of Iowa.

References: "The Slavocracy and the Homestead Problem of 1854," *Agricultural History* (April 1966); "A Scalogram Analysis of the Kansas-Nebraska Bill of 1854 and Related Roll-Calls in the House of Representatives," *Computers and the Humanities* 8, 2 (March 1974). [*Chum,* May 1967]

S4. Investments in England's Overseas Expansion, 1575-1630

Principal investigator: Theodore K. Rabb, Professor, Dept. of History, Princeton U., Princeton, NJ 08540.

Objective: Analysis of investors in some 40 trading companies according to their social classes, membership in Parliament, date of admission to the various companies, offices (e.g., directorships) attained, etc. *Method:* All the data were subjected to exhaustive correlations, reduction of the information to chronological series, and tests of the differences in the activity of the different classes.

Type of computer: IBM 1401 and 7094. *Size of storage:* 32k (7094). *Language:* FORTRAN II and IV. *Disks or tapes:* 2 tapes. Program is available from the Widener Library, Harvard University; SSUS, Princeton University Computer Center; and ICPR, University of Michigan.

References: "Investment in English Overseas Enterprise, 1575-1630," *Economic History Review* 19 (1966), 70-81; *Enterprise and Empire* (Cambridge: Harvard University Press, 1967); "On Nominalism and Idealism, Historical and Statistical," *The Historical Journal* 15, 4 (1972).

S5. Tendencies in the Careers of American Scientists and Engineers, 1846-1876

Principal investigator: Robert V. Bruce, Professor, Dept. of History, Boston U., Boston, MA 02215. *Associate:* John B. Cusack.

Objective: To determine correlations between careers and social, political, economic, and other external factors, as part of research for a study of the development of American science, 1846-1876. *Scope and method:* Descriptive statistics, applied to 1100 American scientists and engineers listed in the *Dictionary of American Biography* as having been active at any time between 1846 and 1876, inclusive.

Type of computer: IBM 1620 Model 2. *Size of storage:* 60k. *Language:* All phases of FORTRAN. *Disks or tapes:* 2 1311 disks. *Special features:* Binary feature, floating point hardware, 1443 printer. Program is available.

Status: Completed. *Future plans:* Significant results will be embodied in a forthcoming book. Some have been published.

References: Science, Democracy, and Union, 1846-1876, a volume in preparation for the series "The Impact of the Civil War on American Life" (Alfred A. Knopf); "A Statistical Profile of American Scientists, 1846-1876," *Nineteenth-Century American Science: A Reappraisal,* ed. George Daniels (Evanston, IL: Northwestern University Press, 1972).

S6. Evolution of Social Organization in Prehistoric Arizona

Principal investigator: Paul S. Martin, Chief Curator Emeritus, Dept. of Anthropology; Director, Field Museum-National Science Foundation Undergraduate Research Participation Program; Lecturer, Dept. of Anthropology, U. of Chicago, Chicago, IL 60637. *Associates:* John M. Fritz, James N. Hill, William A. Longacre.

Objective: To document cultural processes and to formulate laws of cultural regularity.

Type of computer: IBM 7094.

References: Paul S. Martin et al., "Chapters in the Prehistory of Eastern Arizona, II," *Fieldiana: Anthropology* 55 (1964); James N. Hill, *Broken K.: A Prehistoric Society in Eastern Arizona,* in press. [*CHum,* May 1967]

S7. The Political Process in American Communities, 1870-1900

Principal investigator: J.R. Hollingsworth, Associate Professor, Dept. of History, U. of Wisconsin, Madison, WI 53706.

Objective: To analyze comparatively for a number of American middle-sized communities the following questions: What people were involved in the political process, for what purposes, in what ways, and for what consequences? [*CHum,* May 1967]

S8. A Study of Voting Alignments in the United States Senate during Three Transitional Periods of the Twentieth Century, 1909-1915, 1917-1923, and 1929-1935

Principal investigator: Howard W. Allen, Assistant Professor, Dept. of History, Southern Illinois U., Carbondale, IL 62901. *Associates:* Jerome M. Clubb, Aage R. Clausen.

Objective: To assess the influence of political parties, ideology, and constituency interests and pressures upon the legislative process. *Method:* Guttman scale analysis, modifications of the Rice Institute measures of group cohesion and conflict, and various correlation statistics.

Type of computer: IBM 1401 and 7090.

Status: Most of our basic analysis is completed; we are now preparing special reports about some of our more important findings. *Future plans:* To complete the remainder of the basic analysis and produce a series of articles and perhaps a more lengthy manuscript reporting the results of our research.

References: Howard W. Allen and Jerome M. Clubb, "The Party Loyalty in the Taft and Wilson Years: The Senate," *The Journal of Politics* (August 1967). [*CHum,* May 1967]

S9. The Use of Statistical Models in the Interpretation of Anthropological Data

Principal investigator: Robert A. Benfer, Teaching Assistant, U. of Texas, Austin, TX 78712.

Objective: To test the usefulness of various models developed elsewhere for anthropological use; the modification of these models where necessary; and the generation of new models where feasible. *Method:* To utilize previously existing programs for various models, and to test the usefulness of these models with anthropological data.

Type of computer: CDC 6600. Program is available through the program library.

Status: Most of the work involving the use of the computer is complete. *Future plan:* More emphasis on development of new models for anthropological use.

References: "A Design for the Study of Archeological Characteristics by Populational, Genetical and Psychological Models," *American Anthropologist* (accepted for publication). [*CHum,* May 1967]

S10. Computer Simulation of the Applications of Historical Technique

Principal investigator: George G. S. Murphy, Associate Professor, U. of California, Los Angeles, CA 90024. *Associate:* M.G. Mueller. *Method:* We consider techniques used by historians (stages, types, genetic method), and then write programs which will execute such techniques.

Type of computer: IBM 7094 and IBM 360. *Language:* FORTRAN IV. Program is available.

References: "On Making Historical Techniques More Specific: 'Real Types' Constructed with a Computer," *History and Theory* 6 (1967), 14-32.

S11 (See S86).

S12. Backgrounds of the Highest Elected Elite in Japan, 1890-1960

Principal investigator: George Akita, Professor, Dept. of History, U. of Hawaii, Honolulu, HI 96822.

Scope: By inquiring into the political, economic, social, and educational backgrounds of all those who were elected to the House of Representatives (Die of Japan) to answer some basic questions on the political history of modern Japan. *Method:* Collect data from standard biographies, special biographies, year books. This stage has been completed. I am now concerned with the problems of translating, cleaning, and standardizing the data. [*CHum,* May 1969]

S13. The Charitable Impulse in Old and New England, 1660-1760: An Enquiry into the Origins of National Character

Principal investigator: Kenneth A. Lockridge, Assistant Professor, Dept. of History, U. of Illinois at Chicago Circle, Chicago, IL 60680.

Objective: A study in the sources of national character, based largely on data from thousands of wills. Chief subject is the relative frequency and nature of the charitable impulse at death. *Method:* Cross-tabulation of data abstracted from documentary sources. Project 50 percent complete.

Type of computer: In process of changing computers and advancing method.

References: Article in progress, "The Charitable Impulse in New England, 1660-1670." [*CHum,* May 1968]

S14. The Rise of Native Capital in Texas, 1850-1860: A Test of Incipient Transition of the Rostow Thesis

Principal investigator: G.R. Woolfolk, Professor, Dept. of History, Prairie View A & M College, Prairie View, TX 77445. *Associate:* Anasuya S. Rao.

Objective: 1) To discover the points of emergence of the industrial dimension in a frontier and/or "traditional" economy and the basic characteristics of that emergence. This will test Rostow's concept of transition. 2) To discover if the plantation economy serves as an effective bar to entrepreneurial pursuits. *Method:* Scope was confined to Texas in the decade 1850-1860. Theoretical approaches to both entrepreneurial evolution and Southern economic development from relevant scholarship provided a frame of reference for the critical probe of primary data at state and federal levels. Data was coded by individual businesses and by county designations of wealth-producing activities for 1850 and 1860. Code sheets were turned over to the data processing center of A & M University of Texas for processing and programming.

Status: Major programming and ancillary activities are not entirely completed, but major trends of the data processed indicate the final result.

S15. The Analysis of Folk Narratives

Principal investigator: Benjamin N. Colby, Professor, School of Social Sciences, U. of California, Irvine, CA 92664. *Associate:* Roger Knaus.

Objective: To isolate eidochronic elements and study their distribution in a search for distinctive cultural patterns in folk narrative. *Method:* A combination of standard techniques for counting and grouping words, retrieving sentences, etc., with a variety of statistical techniques.

Type of computer: PDP-10. *Size of storage:* 112k. *Language:* LISP.

References: "The Analysis of Culture Content and the Patterning of Narrative Concern in Texts," *American Anthropologist* 68, No.2, Part 1 (1966), 374-388; "Culture Patterns in Narrative," *Science* 151, No. 3712, 793-798.

S16. Analysis of the Minoan Hieroglyphic Script

Principal investigators: Richard S. Morgan and John J. Reich, Dept. of Classics, U. of Manitoba, Winnipeg 19, MB, Canada.

Scope: All Minoan hieroglyphic inscriptions to be examined (seal stones, sealings, and graffiti); also the Zakro sealings and the Aghia Triadha sealings. *Method:* 1) Details of inscriptions punched onto cards. 2) Printout of corpus and transfer to magnetic tape. 3) Sort and list symbols, proveniences, materials, colors. 4) Formulae involving two or more symbols listed and correlated.

Type of computer: IBM 360/75. *Size of storage:* 256k. *Language:* Assembler (DOS converting to OS–level F). *Disks or tapes:* 2 2400 tapes.

Status: Corpus of Scripta Minoa I, Kenna's Cretan Seals, Giamalakis and Chapouthier (Mallia) now assembled.

References: Mention of this project in *Calculi* 5 (March 1967), and in *Nestor* 457 (April 1967), "The Role of the Naturalistic Signs in the Minoan Hieroglyphic Script," *Atti e Memorie del Primo Congresso Internazionale di Micenologia* (Roma 1967). [*CHum,* May 1968]

S17. Proportion Analysis of Hallstatt and La Tène Spearheads

Principal investigator: Ralph M. Rowlett, Instructor, Dept. of Anthropology, U. of Missouri, Columbia, MO 65201. *Associate:* Dee F. Green.

Objective: Determine which parameters of such spearheads are useful for chronological determination and which are useful for cultural group attribution. *Scope:* All spearheads of iron with archaeological associations and possibility of dating from Hallstatt I through La Tène IIb from Champagne, France. *Method:* Generalized distance, numerical taxonomy.
Language: FORTRAN. Program is available.
Status: Initial stages only; program still under development. [*CHum,* May 1967]

S18. Computerization of the Ethnographic Atlas

Principal investigator: Lewis H. Larson, Jr., Associate Professor, Dept. of Anthropology, Georgia State U., 33 Gilner SE., Atlanta, GA 30303. *Associate:* Dayle R. Doyle.

Objective: To place the Ethnographic Atlas data (*Ethnology,* vol.1-date) on computer cards and program the Atlas categories and data for a printout. This was done to facilitate cross-cultural research within the limitations of the Atlas data. Currently the project is limited to data recovery.

Type of computer: IBM 7040. *Size of storage:* 32k. *Language:* COBOL. *Disks or tapes:* Minimum IBSYS configuration. Program is available.

Status: The project is continuing as new sections of the Atlas are published. [*CHum,* May 1967]

S19. Modernity and Traditionalism in Tlacolula, Oaxaca, Mexico

Principal investigator: Martin Diskin, Instructor, M.I.T., 14N-310, Cambridge, MA 02139.

Objective: Construction of a scale to measure modern/traditional attitudes, life style, etc. *Method:* Sample survey with questionnaire.

Type of computer: IBM 7094. *Language:* Crosstabs (designed at Project MAC, M.I.T.). Program is available.

Status: Computer stage effectively finished. Report being written. *Future plans:* Further field study, possibly with precoded material. [*CHum,* May 1967]

S20. Chronological Ordering

Principal investigator: Marcia Ascher, Professor, Dept. of Mathematics, Ithaca College, Ithaca, NY 14850; Robert Ascher, Professor, Dept. of Anthropology, Cornell U., Ithaca, NY 14850.

Objective: To order in time archaeological collections.

Status: Mathematical analysis of problem, beyond that described in reference below.

References: "Chronological Ordering by Computer," *American Anthropologist* 65 (1963), 1045-1052; Marcia Ascher, "A Mathematical Rationale for Graphical Seriation," *American Antiquity* 25 (1959), 212-214.

S21. DREASYS: Divostin Data Retrieval and Analysis System

Principal investigator: Alan McPherron, Associate Professor, Dept. of Anthropology, U. of Pittsburgh, Pittsburgh, PA 15260. *Associate:* Joel Gunn.

Objective: Development of interactive time-sharing program package for studying spatial distributions and performing multivariate analysis of archeological data.

Type of computer: PDP-10. *Language:* FORTRAN. Program is partly available.

S22. The British House of Commons, 1841-1847

Principal investigator: William O. Aydelotte, Professor, Dept. of History, U. of Iowa, Iowa City, IA 52240.

Objective: Analysis of divisions (roll calls) in British House of Commons, 1841-1847, using Guttman scaling technique.

Type of computer: IBM 360/65. *Size of storage:* 1024k bytes. *Language:* FORTRAN IV G. *Disks or tapes:* 2 9-track tapes. *Special equipment:* Data Computer Systems CP-4 Remote I/P Unit. Program is available.

Status: Information on votes in Parliament has been summarized in cumulative scales of the Guttman type; the scales have been compared with each other by various means. Information on personal and political background has also been tabulated and has been correlated with party affiliation and scale position.

References: "Voting Patterns in the British House of Commons in the 1840s," *Comparative Studies in Society and History* 5, 2 (January 1963), 134-163; "Parties and Issues in Early Victorian England," *Journal of British Studies* 5, 2 (May 1966), 95-114; "The Conservative and Radical Interpretations of Early Victorian Social Legislation," *Victorian Studies* 11, 2 (December 1967), 225-236; "The Disintegration of the Conservative Party in the Eighteen-Forties: A Study of Political Attitudes," delivered at conference sponsored by Mathematical Social Science Board, University of Chicago, June 1969, and at the annual meeting of the American Political Science Association, September 1969 (was to be published in an anthology edited by Robert W. Fogel, Allan G. Bogue, and myself). [*CHum,* November 1970]

S23. "PENNINK"

Principal investigator: G. Miller, Computer Center, U. of Pennsylvania, Philadelphia, PA 19104.

Objective: University of Pennsylvania version and extension of Philip J. Stone's General Inquirer.

Type of computer: IBM 7040. *Size of storage:* 32k. *Language:* FORTRAN IV and MAP.

Status: Completed. *Future plans:* Additional extension. [*CHum,* May 1967]

S24. Research on Decision-making in Crisis

Principal investigator: Ole R. Holsti, Professor, Dept. of Political Science, U. of British Columbia, Vancouver, BC.

Objective: To develop and test hypotheses about the effects of crisis-induced stress on decision-making processes. Content analysis of political documents has been a primary source of data. *Method:* Content analysis by computer; aggregate data analysis; other relevant techniques.

Type of computer: IBM 7090. *Disks or tapes:* 32k. *Language:* BALGOL. *Disks or tapes:* 1 disk, 14 tapes. Program is available from Cambridge Computer Associates, Cambridge, MA.

Status: A continuing project. *Future plans:* Development of better techniques of computer content analysis; continuation of empirical case studies.

References: "Computer Content Analysis as a Tool in International Relations Research," *Computers in Humanistic Research,* ed. Edmund A. Bowles (Englewood Cliffs, NJ: Prentice-Hall, 1967); "External Conflict and International Cohesion: The Sino-Soviet Case," *The General Inquirer: A Computer Approach to Content Analysis in the Behavioral Sciences,* ed. P.J. Stone (Cambridge: MIT Press, 1966); "Content Analysis in Political Research," *Computers and the Policy-Making Community,* ed. D. Bobrow and J. Schwartz (Englewood Cliffs, NJ: Prentice-Hall, 1968); *Content Analysis for the Social Sciences and Humanities* (Reading, MA: Addison-Wesley, 1969); *Crisis, Escalation, War* (Montreal and London: McGill-Queen's University Press, 1972); Ole R. Holsti, P.T. Hopmann, and J.D. Sullivan, *Unity and Disintegration in International Alliances* (New York: Wiley, 1973).

S25. Folk Medicine in Texas

Principal investigator: John Q. Anderson, Professor, Dept. of English, U. of Houston, Houston, TX 77004.

Objective: A system for handling a large amount of natural language data comprised of short separate items.

Scope: More than 6000 cures, remedies, and health practices have been collected over the years; to be arranged for comparative analysis. [*CHum,* May 1967]

S26. Salish Language and Culture

Principal investigator: Joseph G. Jorgensen, Associate Professor, Dept. of Anthropology, U. of Michigan, 221 Angell Hall, Ann Arbor, MI 48104.

Scope: A statistical analysis of internal cultural relations, internal language relations, and the relations between language and culture. Inferential hypotheses about culture and language evolution and history are tested. *Method:* Thirty Salish cultures and languages are compared and ordered into matrices and tree diagrams by a non-metric technique for finding the shortest distance in a euclidian space, a very fast cluster analytic method that makes many fewer assumptions than metric cluster analytic methods.

Type of computer: IBM 360/40, 360/67. *Language:* FORTRAN IV G. I will reproduce decks if anyone is interested.

Status: Correlations, coefficients of similarity, and coefficients of distance are ready for matrix ordering. A huge correlation project (with Harold E. Driver and William C. Smith) on the North American Indians is planned.

References: "Geographic Clusterings and Functional Explanations of In-Law Avoidance: An Analysis of Comparative Method," *Current Anthropology* 7 (1966), 161-169; *Salish Language and Culture: A Statistical Analysis of Internal Relationships, History and Evolution,* Language Science Monographs, Supplement to International Journal of American Linguistics, Bloomington, IN, 1969. [*CHum,* May 1969]

S27. Analysis of Prehistoric Pottery from Tonga, Western Polynesia

Principal investigator: Jens Poulsen, Senior Lecturer, Institute of Prehistoric Archaeology and Ethnography, U. of Aarhus, Moesgard, 8270 Hojbjerg, Denmark.

Objective: An archaeological time sequence based on studies of pottery development and seriation of shell midden horizons. *Method:* A statistical evaluation was made of the frequencies of pottery features on a background of sherd distribution in stratified shell middens excavated. The pottery features were organized in rim and decoration codes, each sherd described on a separate 80-column punch card. The distribution of the sherds in the sites was reconstructed by means of a record code which allows the retrieval of any detailed contextual information. Revision of this code and of the punch cards now permits also straightforward retrieval of the total horizon distribution of any individual pottery feature as well as of any combination of the pottery features coded in various kinds of tables. The statistical testing is based on use of the *chi*-square formula, the results of which are incorporated in the tables.

Type of computer: 1) IBM 1620 and 360, 2) GIER, 3) CDC 6400. *Language:* 1) FORTRAN, high, 2) ALGOL, SLIP (Assembler), 3) FORTRAN, high. *Disks or tapes:* 3) 841 and BM 101 disks. Program is available.

Status: Completed in 1972.

References: A Contribution to the Prehistory of the Tongan Islands, vols. I-II (Unpub. Ph.D. dissertation, The Australian National University, 1967); *On the Processing of Pottery Data,* Jysk Arkaeologisk Selskab, Handboger II, 1972, I kommission hos Gyldendalske Boghandel, Nordisk Forlag, Kobenhavn, 1972; J.I. Poulsen and L.M. Groube, *Early Tongan Prehistory* (forthcoming volume of the series *Terra Australis,* published by the Department of Prehistory, Research School of Pacific Studies, The Australian National University, Canberra, Australia).

S28. Urbanization at Teotihuacan, Mexico

Principal investigators: Rene Millon, Professor, Dept. of Anthropology, U. of Rochester, Rochester, NY 14627; George L. Cowgill, Associate Professor, Dept. of Anthropology, Brandeis U., Waltham, MA 02154.

Scope: A computer data file is being completed for data storage and retrieval and for statistical manipulations. This is one major aspect of Professor Millon's very extensive archaeological study of the ancient city of Teotihuacan, based on detailed mapping, very thorough surface reconnaissance, and excavations in a few structures of critical importance. *Method:* For each of several thousand structural units, counts are recorded for many categories of artifacts found on the surface, as well as ratings or states for many kinds of qualitative observations, and both are transferred to a tape data file. Statistical methods emphasized so far include cross-tabulation of pairs of variables or of variables versus locality, and studies of association patterns by means of Shepard-Kruskal multidimensional scaling.

Type of computer: IBM 7094. *Language:* FORTRAN II, FORTRAN IV. Program is available. Much of work uses the DATA-TEXT program system, presently specific to IBM 7094. Write Social Relations Department, Harvard University, Cambridge, MA.

References: G. Cowgill, "Computer Analysis of Archeological Data from Teotihuacan, Mexico," *New Perspectives in Archeology,* ed. Sally and Lewis Binford (Chicago: Aldine, 1968); G. Cowgill, "Evaluación preliminar de la aplicación de métodos a máquinas computadoras a los datos del mapa de Teotihuacan," *Teotihuacan: Onceava Mesa Redonda, México, 1966* (Mexico: Sociedad Mexicana de Antropologia, 1967, English with Spanish summary). [*CHum,* May 1969]

S29. Bibliographical Dictionary, with Statistical Sections, of British Colonial Service from about 1870 to 1929

Principal investigator: Vincent Ponko, Jr., Associate Professor, Dept. of History, Villanova U., Villanova, PA 19055.

Objective: A computer-produced bibliographical dictionary, with statistical sections, of the British Colonial Service. Data to be collated from many divergent sources in government and private depositories, punched on IBM cards to produce a printout which will give statistical and biographical information hitherto uncollected to aid in showing the decisive contributions these men have made to the growth of the areas in which they were stationed, as well as to the history of Great Britain. *Scope and Method:* The production of coded name IBM cards from various printed sources in London, Oxford, and elsewhere in the U.K. This procedure should provide biographical information such as name, place and date of birth, education, places of service, type and years of service, retirement, death, publications and awards; statistical information such as total time of service for each man, age groupings, number of men in any particular year in terms of service categories, geographical distribution of homes of people entering the service, educational institution distribution, listing of persons in service in each colony, and number of people who transferred to the Colonial Service from the military or the Foreign Service. [*CHum,* May 1967]

S30. Inventory of the Ethnological Collections in Museums in Missouri

Principal investigator: Mary Jane Schneider, Dept. of Sociology-Anthropology, U. of North Dakota, Grand Forks, ND 58201.

Scope: To compile an index of the ethnological materials in Missouri. These will have a world-wide range and will form a companion file to that at the University of Oklahoma. *Method:* 1) Record the items in each museum. 2) Make punch cards from each record sheet. 3) Sort.

Type of computer: IBM 360/75. *Language:* FORTRAN IV. Program, GYPSY, is available from University of Oklahoma Computer Center.

References: "A Museum Inventory for Missouri," *Midwest Museums Quarterly* (Fall 1968); "Indexing Ethnological Collections in Missouri," *Museum Graphic* (Winter 1969); "A Guide to Inventorying Ethnological Collections" (1970, available from author).

S31. Informational Language for Recording the Material of Historical Polythematic Funds

Principal investigator: Milan Svankmajer, Institute of History of the European Socialist Countries of the Czechoslovak Academy of Science, Praha 1, Thunovská 22, Czechoslovakia. *Associate:* Vl. Vanek.

Objective: Annotated bibliography of titles in the field of the nineteenth and twentieth centuries from scientific historical journals.

Type of computer: EPOS 1. *Disks or tapes:* 3 tapes. Program is available.

Status: Research in the selective quality of the informational language. *Future plans:* Elaboration of 1) the general computer program, 2) the informational language (indices: maximum 10 percent of redundance—maximum 1 percent of loss of material).

References: "Kotázkám stronjn ich resersi v historii," *CSCH* 1 (1965). [*CHum,* May 1967]

S32. The United States Population, 1800-1960

Principal investigator: Jack E. Eblen, Assistant Professor, Dept. of History, U. of Connecticut, Storrs, CT 06268.

Objective: To determine the type and nature of changes that occurred in a variety of specific population types during the past 160 years.

Scope: The study will be programmed for computer analysis and graphing of a number of key demographic characteristics, such as age and sex distribution, fertility, nativity, mortality, and so on, for each of some 210 specific types over the 160-year period. Program will be available.

Status: I am currently running a pilot project through which I expect to develop the program, techniques of computer analysis. *Future plans:* To run the general study within a year or two.

References: "An Analysis of Nineteenth-Century Frontier Populations," *Demography* II (1965); "American Historical Demography: A Survey of the Nature and Kinds of Changes that Occurred in the Last 160 Years" (tentative title), for *Journal of Social History* (Winter 1967). [*CHum,* May 1967]

S33. Methodology of Cross-Societal Analysis

Principal investigator: Jack Sawyer, Professor, Wright Institute, 2728 Durant Avenue, Berkeley, CA 94704. *Associate:* Robert A. LeVine.

Objective: To explore measurement problems involving the selection of variables and of units of analysis; analytic problems like testing models of evolutionary development, and separating the effects of historical diffusion and independent invention; and problems of establishing generality through replication and through comparison of cross-national and cross-societal analyses.

Type of computer: CDC 6400. *Size of storage:* 65,536 60-bit words in central memory. *Language:* FORTRAN, COMPASS. *Disks or tapes:* 85-million-character magnetic disk storage.

References: "Dimensions of Nations: Size, Wealth, and Politics," *American Journal of Sociology* 73 (September 1967), 145-172; "Cultural Dimensions: A Factor Analysis of the World Ethnographic Sample," *American Anthropologist* 68 (June 1966), 708-731.

S34. Social Mobility in Boston, 1880-1970

Principal investigator: Stephan Thernstrom, Professor, Dept. of History, Harvard U., Cambridge, MA 02138.

Type of computer: IBM 7094.

Status: Five random samples of the Boston population have been drawn from manuscript census schedules, marriage license records, etc., and sample members have been traced through city directories for as long as they remained in the city. Various types of social mobility have been measured and analyzed.

References: The Other Bostonians: Poverty and Progress in the American Metropolis (Cambridge: Harvard University Press, 1973).

S35. Voting Alliances and Voting Blocs in the United Nations

Principal investigator: Charles Wrigley, Professor, Computer Institute for Social Research, Michigan State U, East Lansing, MI 48823.

Objective: To find the changes in alliances in the twenty-year history of the United Nations. *Method:* Factor analysis and the pattern-analytic methods of McQuitty are being used to identify voting blocs.

Type of computer: CDC 3600. *Size of storage:* 65k. *Language:* FORTRAN IV. *Disks or tapes:* 10 tapes; 2 drums. Program is available.

Status: The exploratory analyses for examining the feasibility of the methods have been completed. [*CHum,* May 1967]

S36. Class, Corruption, and Politics in the French Chamber of Deputies—1846-1848

Principal investigators: Patrick Higonnet, Assistant Professor, Harvard U., Cambridge, MA. *Associate:* Trevor Higonnet.

Objective: Investigation of the background of political opinion in the French Parliament.

Type of computer: IBM 7090.

Status: Completed.

References: Article in *French Historical Studies,* Fall 1967. [*CHum,* May 1967]

S37. Museum Cataloging by Computer

Principal investigator: Robert G. Chenhall, Dept. of Anthropology, U. of Arkansas, Fayetteville, AR 72701.

Objective: To provide lists of ethnological and archeological museum collections by type of items, by culture, by location, and by material and method of manufacture.

Type of computer: CDC 3400. *Size of storage:* 32k. *Language:* Assembly (COMPASS) or FORTRAN. [*CHum,* May 1967]

S38. Content Analysis for Explaining Judicial Decision-Making

Principal investigator: Werner F. Grunbaum, Professor, Dept. of Political Science, U. of Missouri-St. Louis, 8001 Natural Bridge Road, St. Louis, MO 63121.

Scope: A 450-word dictionary with 125 sub-entries was classified according to 51 content categories and sub-categories. The resulting St. Louis Supreme Court Dictionary yields category tabulations in sentence and word units that distinguish structural and attitudinal variables between various justices' opinions.

Type of computer: IBM 360/50. *Size of storage:* 90,000 bytes static + automatic storage. *Language:* PL/I, version 4. *Disks or tapes:* 1 disk (optional). Program is available.

References: "Quantitative Analysis of the 'Presidential Ballot' Case," *Journal of Politics* 34 (February 1972), 223-243.

S39. Computer Aids to Historical Lexicography

Principal investigators: Richard W. Bailey, Associate Professor; James W. Downer, Professor; Jay L. Robinson, Associate Professor, Dept. of English, U. of Michigan, Ann Arbor, MI 48104.

Scope: A variety of pilot studies have been undertaken since 1967 to investigate: 1) the manipulation of large files of linguistic material, and 2) techniques to facilitate interaction between the stored material and the lexicographer. The eventual goal of our work is to establish techniques to assist in historical and commercial lexicography that make use of contemporary and future technology. *Method:* Computer-output microfiche; we anticipate using the Magnavox Plato IV Student Terminal for experimental editorial work.

Type of computer: IBM 360/67. *Language:* PL/I and COBOL. *Special equipment:* COM unit.

Status: In 1973 we expect to publish *Michigan Early Modern English Materials* which includes 1) a user's handbook and bibliograpy, 2) computer-produced microfiche containing nearly 20,000 citations of Early Modern English Usage (1475-1700), 3) a magnetic tape record of this material, 4) a "secondary index" referencing all usages in the collection.

References: R.W. Bailey and J.L. Robinson, "The University of Michigan Early Modern English Dictionary Project," *Shakespearean Research Opportunities* 4 (1968-69), 120-121; R.W. Bailey and J.L. Robinson, "Computers and Dictionaries," *Computers and Old English Concordances,* ed. Angus Cameron, Roberta Frank, and John Leyerle (Toronto: University of Toronto Press, 1970), pp. 94-102; R.W. Bailey and J.L. Robinson, "The Computer in Lexicography," *Festschrift Kurath,* ed. John Reidy and Harald Scholler (Wiesbaden: Franz Steiner Verlag, 1973), pp. 37-45; R.W. Bailey and J.L. Robinson, "Computer Produced Microfilm in Lexicography: Toward a Dictionary of Early Modern English," *Computers and Literary Studies,* ed. A.J. Aitken, R.W. Bailey, and N. Hamilton-Smith (Edinburgh: Edinburgh University Press, 1973); R.W. Bailey, "Reflections on Technology in Lexicography," *Proceedings of the International Conference on Lexicography in English* (New York: New York Academy of Sciences, to appear); R.W. Bailey, "Research Dictionaries,"

American Speech (to appear); R.W. Bailey, J.W. Downer, and J.L. Robinson, *Michigan Early Modern English Material* (to appear).

S40. Analysis of Chronological Change in Archeological Data

Principal investigators: Frank Hole, Associate Professor, Dept. of Anthropology and Sociology, Rice U., Houston, TX 77001; Mary Shaw, Project Scientist, Dept. of Computer Science, Carnegie-Mellon U., Pittsburgh, PA 15213.

Objective: To evaluate the theory for understanding systematic artifact change, to compare previously proposed techniques for deducing chronology, and to develop an improved technique. *Method:* The proportion of each of the artifact types in a site characterizes that site. The chronology for a group of sites is derived by ordering them in such a way that the more similar two sites are, the closer they are placed in time.

Type of computer: We are in process of converting the programs to run on the IBM 360. Program is available.

References: "Computer Analysis of Chronological Seriation," *Rice University Studies* 53 (1967). [*CHum,* November 1968]

S41. Lessico Rosminiano

Principal investigator: Michele Federico Sciacca, Professor, Istituto di Filosofia dell'Università di Genova, Genoa, Italy.

Scope: Opera Omnia di A. Rosmini. *Method:* Si procedrà ottenendo dapprima una lista di tutte le parole con le relative frequenze. I ricercatori sceglieranno in base a questa lista le parole rilevanti delle quali quali desiderano una elaborazione ulteriore, fino alle concordanze e alle schede-contesto. La Direzione dell'Istituto ha deciso di seguire nella elaborazione dei testi i metodi e i programmi dell'Accademia della Crusca. E in corso la preedizione e la perforazione dei primi testi.

Size of storage: 5 milioni di parole. [*CHum,* November 1967]

S42. Vocabolario Giuridico Italiano

Principal investigator: Piero Fiorelli, Professor, c/o Centro Studi IBM, Via S. Maria, 36, Pisa, Italy.

Scope: Codici, testi legislativi e dottrinali di diritto. *Method:* Si Vogliono ottenere concordanze e schede-contesto per la compilazione di un vocabolario storico della lingua giuridica italiana e, in linea subordinata, per la costituzione di un archivio di documentazione giuridica. Si è deciso di seguire lo stesso metodo messo a punto per l'Accademia della Crusca. Si svolge regolarmente il ciclo di perforazione-controllo lemmatizzazione.

Size of storage: Da stabilire. [*CHum,* November 1967]

S43. Index Thomisticus (Number changed to P1)

S44. A Computer Analysis of Chronologies in Genesis

Principal investigator: David E.Y. Sarna, IBM Israel, Bank Street, Haifa, Israel.

Scope: The 64 different ages/numbers in Genesis. *Method:* 1) List and keypunch original data. 2) Sort by various factorization methods. 3) Print out tabular form factors and keyed explanation. 4) Hand inspection for factors forming pattern.

Type of computer: Honeywell 1800. *Language:* MAC.

Status: Completed 1967. Computer output and the analysis reports are available for the cost of reproduction.

References: Proceedings of the ACM 22, 301-309.

S45. Swedish Voluntary Movements 1850-1950

Principal investigators: Carl Goeran Andrae and Sven Lundkvist, Docents, Dept. of History, Uppsala U., Kungsangsgatan 19, Uppsala, Sweden. *Associate:* Ulla Hedquist.

Scope: Material processed by IBM 551 data sheets: a) Membership of 20,000 local unions and associations 1850-1950; b) personal information of the inhabitants of three towns 1886, 1892, and 1902 in census, assessment registers, electoral registers, and union lists of members; c) newspaper announcements concerning the activities of the movements 1875-1900. *Method:* 1) Ecological analyses of the voluntary movements by material a and b. 2) Structural analyses of

some towns specially considering the members of voluntary movements by material b. 3) Functional analyses of the activities of the movements by material c.

Type of computer: CDC 3600. *Size of storage:* 32k. *Language:* 3600 FORTRAN. *Disks or tapes:* 2 CDC 603 and 4 CDC 604 tapes; 2 CDC 861 drums. *Special equipment:* IBM optical reader. [*CHum,* November 1967]

S46. Analysis of the Cajas of Lima and Mexico, 1519-1810

Principal investigator: John J. TePaske, Professor, Dept. of History, Duke U., 6727 C.S., Durham, NC 27708.

Objective: To determine the income and the outgo of the various *ramos* or sections of the two major treasuries in the Spanish New World both as an index to economic conditions in the two viceroyalties and as a basis for comparison of the two areas. *Method:* Compilation of the *sumarios* or summaries of the income and outgo (*cargo y data*) from archival sources in Peru, Mexico, and Spain. Data not yet fully computerized.

Type of computer: IBM 360. *Language:* SCATRAN and FORTRAN.

Status: Continuing.

References: "Quantification in Latin American Colonial History," *The Dimensions of the Past: Materials, Problems and Opportunities for Quantitative Work in History,* ed. V. Lorwin and J. Price (New Haven: Yale University Press, 1972), 431-476.

S47. Rose Bibliography [Suspended] For information write to Robert Walker, Dept. of American Studies, George Washington U., Washington, DC 20006. [*CHum,* May 1968.]

S48. A Historical Case Study of the Effect of Education Reform on an Underdeveloped Area: Scotland in the Eighteenth Century

Principal investigator: Vern L. Bullough, Professor, Dept. of History, California State U., Northridge, CA 91324. *Associates:* Bonnie Bullough, Martha Voght Barcus, Lucy Kluckhohn.

Objective: To explain why so many eminent men appeared in eighteenth-century Scotland. It was hypothesized that one of the key reasons was the educational reform which had taken place. *Method:* A sample of 375 eminent Scotsmen was drawn from standard accounts in history of economics, philosophy, etc., and compared with Scotsmen listed in the *Dictionary of National Biography.* Biographical information was gathered upon the sample from source materials and then the information was coded as to social status, educational achievement, etc. The next step was to examine the parishes in which the eminent men either were born or received their primary education, and in addition a random sample of non-contributing parishes was drawn to make a total of some 246 upon which studies were made to determine the economic conditions in the parish, the nature of schooling available and so forth. The 246 sample parishes represented a little more than a quarter of the 938 in existence at that time.

Type of computer: 1) GE 225, 2) GE 415. *Size of storage:* 8k. *Language:* FORTRAN. *Disks or tapes:* 4 tapes. Program is available.

Status: Completed in 1970.

References: "Intellectual Achievement in Eighteenth Century Scotland: A Computer Study," *Comparative Education Review* (1970); "Longevity and Achievement," *Omega: Journal of Geriatric Psychology* (1970); "The Causes of the Scottish Medical Renaissance of the Eighteenth Century," *Bulletin of the History of Medicine* (1971); "Birth Order and Achievement in Eighteenth Century Scotland," *Journal of Individual Psychology* (1971); "Intellectual Achievers," *American Journal of Sociology* (1971).

Status: Book-length study is near completion.

S49. BEAR (Berkeley Elites Automated Retrieval)

Principal investigator: David Nasatir, Associate Professor, Dept. of Sociology, U. of California, Berkeley, CA 94720. *Associates:* David Apter, Robert A. Scalapino, Leo Rose, B. Jhoshi, Margaret Fisher, Joan Bondurant, Charles Yarbrough.

Scope: Gathering selected biographical materials on communist elites in Korea, Vietnam, and China; technocratic elites in Argentina, Chile, Peru; legislative elites in India and Pakistan; royalty in Nepal. *Method:* Using a form specially developed for the project, research assistants gather available data on selected individuals from all available sources. Materials are keypunched and stored on tape for retrieval and analysis.

Type of computer: IBM 360/40. *Size of storage:* 128k. *Language:* FORTRAN IV G. *Disks or tapes:* 1 2314 disk, 3 7-track tapes. *Special equipment:* Upper/lower case print chain. Program is available.

Status: Development of a machine-readable data base of biographical materials stored—in large measure—in textual form; data base was to be used by various scholars for diverse purposes. BEAR system of computer programs used for storage and retrieval purposes. Project discontinued.

S50. International Comparative Political Parties

Principal investigator: Kenneth Janda, Associate Professor, Dept. of Political Science, Northwestern U., Evanston, IL 60201.

Scope: 260 political parties in ninety foreign countries, information about these parties to be retrieved from available writings on party politics in each country, prepared for machine retrieval from microfilm. *Method:* 1) Read and index individual pages of material. 2) Keypunch index codes for computer tabulation. 3) Read index codes with modified IBM 026 keypunch and record codes in binary form on microfilm next to corresponding pages. 4) Develop microfilm and search for desired codes to locate information.

Type of computer: CDC 6400 and Eastman Kodak MIRACODE System. *Size of storage:* 65k. *Language:* FORTRAN. Program is available.

References: "Political Research with MIRACODE: A 16mm. Microfilm Information Retrieval System," *Social Science Information* 6 (April-June 1967), 169-181. [*CHum,* May 1968]

S51. Treatment and Representation of Archaeological Prospecting Results

Principal investigator: Richard E. Linington, Director, Fondazione Lerici Prospezioni Archeologiche, Via Veneto 108, I 00187 Rome, Italy.

Objective: Improving the handling and interpretation of the results from archaeological prospecting (mainly geophysical methods). In addition various special studies including extensive theoretical investigations aimed to help understand and develop the practical applications. *Method:* 1) Punch survey results on cards or paper tape or use other input methods. 2) Check the input which normally involves up to 10,000 values at a time. 3) Normally transfer values to tape. 4) Where necessary apply mathematical treatments ranging from simple averaging to complex filtering systems. 5) Print out the results using a wide range of possible representations including symbol diagrams as well as simulated graphs, contour diagrams, and isometric drawings. The project involves a large number of programs which are revised as required and include both those of regular use and those for special studies especially as regards the problems of interpretation.

Type of computer: IBM 360/50. *Size of storage:* 200k (normal maximum program size). *Language:* FORTRAN IV, Assembler (OS versions). *Disks or tapes:* 2 tapes, disk files.

References: "The Rome Computer System for Treating Archaeological Survey Results": First part, *Prospezioni Archeologiche* 3 (1968), 19-36; Second part, *Prospezioni Archeologiche* 4 (1969), 9-58; "Techniques Used in Archeological Field Surveys," *Phil. Trans. Royal Society of London A* 269 (1970), 89-108;"A First Use of Linear Filtering Techniques on Archaeological Prospecting Results," *Prospezioni Archeologiche* 5 (1970), 43-54; "Further Tests on Ono Non-symmetrical Filtering Systems," *Prospezioni Archeologiche* 6 (1971), 9-20; "A Summary of Simple Theory Applicable to Magnetic Prospecting in Archaeology," *Prospezioni Archeologiche* 7 (1972); "Topographical and Terrain Effects in Magnetic Prospecting," *Prospezioni Archeologiche* 7 (1972).

S52. Changes and Continuity in Political Leadership: A Comparison of Four U.S. Congresses, 1820, 1860, 1900, and 1940

Principal investigator: Ari Hoogenboom, Professor, Dept. of History, Pennsylvania State U., University Park, PA 16802. *Associates:* David Porter, Frank Tusa.

Scope: Comparison of biographical data of congressmen; analysis of roll call votes on certain key issues by section and political party. *Method:* 1) Collect relevant data from *Biographical Directory of American Congress.* 2) Code the biographical data. 3) Keypunch. 4) Sort biographical data. 5) Cross-tabulate biographical data by controlling for political party and geographical region.

Type of computer: IBM 360/67. *Size of storage:* 512k. *Language:* FORTRAN. Program is available.

References: "Industrialism and Political Leadership; A Case Study of the United States Senate," *The Age of Industrialism in America,* ed. Frederick Jaher (New York: The Free Press, 1967). [*CHum,* May 1968]

S53. Interaction and Content Analysis of Northern Rhodesia-Zambia Legislative Debates

Principal investigator: James R. Scarritt, Associate Professor, Dept. of Political Science, U. of Colorado, Boulder, CO 80302.

Scope: Random sample from Northern Rhodesia-Zambia Legislative Debates 1959-1967. *Method:* 1) Code interaction and themes. 2) Keypunch. 3) Sort by thematic categories. 4) Print in tabular form.

Type of computer: CDC 6400. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 1 disk. Program is available.

S54. Index of Australian Aboriginal Song Words

Principal investigator: Alice M. Moyle, Research Fellow (Australian Institute of Aboriginal Studies), Dept. of Music, Monash U., Clayton, Victoria, Australia 3168. *Associates:* Georgette Silva, Peter Hohaus.

Scope: Segmented song texts which have appeared to date in writings on all aspects of Australian Aboriginal culture (65,000 words). *Method:* 1) Edit for song region. 2) Transfer to tapes. 3) Make word index with page and source identification.

Type of computer: CDC 3200. *Size of storage:* 32k. *Language:* FORTRAN. *Disks or tapes:* 4 tapes. Program is available.

References: See G15. *Index of Australian Aboriginal Song Words,* 2nd edition, February 1973.

S55. Political Communications Behavior

Principal investigator: Edward C. Dreyer, Associate Professor, Dept. of Political Science, U. of Tulsa, Tulsa, OK 74104.

Scope: The communication activities of Americans is conceptualized as an important form of political activity which is affected by the individual's social situation and past experiences. Communication activities such as media exposure and interpersonal conversations about political affairs may be a source of variation in political perceptions which govern other political behaviors such as turnout and partisan choice in different electoral situations. *Method:* Secondary analysis of survey data available through the Inter-University Consortium for Research in Political Behavior (Ann Arbor, Michigan). *Special equipment:* IBM 360/40 and 360/67 systems developed by Inter-University Consortium. Program is available through Consortium.

References: "Media Use and Electoral Choices: Some Political Consequences of Information Exposure," *Public Opinion Quarterly* 35, 4 (Winter 1971-72), 544; "Change and Stability in Party Identifications," *Journal of Politics* 35 (August 1973).

S56. Political Socialization of Non-Party Groups in Canada

Principal investigator: Allan Kornberg, Professor, Dept. of Political Science, Duke U., Durham, NC 27706. *Associate:* Joel Smith.

Objective: To compare the process of political socialization of an elite and cross-sectional sample of adults in Vancouver and Winnipeg with previously obtained data from a sample of party leaders in these two cities. *Method:* Structured personal interviews.

S57. Correlates of War

Principal investigator: J. David Singer, Professor, Dept. of Political Science, U. of Michigan, Ann Arbor, MI 48103. *Associates:* Melvin Small, Michael Wallace, Stuart Bremer.

Objective: To ascertain which variables in which combinations best account for fluctuations in the frequency, severity, and magnitude of international war in the period since 1815. *Method:* Bivariate and multivariate analysis of time series data.

Type of computer: IBM 360/67 full duplex. *Size of storage:* 1,048,516 bytes. *Language:* FORTRAN IV. *Disks or tapes:* 8-track tapes.

References: J.David Singer and Melvin Small, *The Wages of War, 1816-1965: A Statistical Handbook* (New York: John Wiley & Sons, 1972); J.D. Singer and Melvin Small, *The Strength of Nations: Comparative Capabilities Since the Napoleonic Wars* (forthcoming); J.D. Singer, M. Wallace, and Stuart Bremer, *A Structural History of the International System, 1816-1965* (forthcoming); "The Correlates of War Project: Interim Report and Rationale," *World Politics* 24, 2 (January 1972), 243-270; J.D. Singer, S. Bremer, and J. Stuckey, "Capability Distribution, Uncertainty, and Major Power War, 1816-1965," *Peace, War, and Numbers,* ed. Russett (Beverly Hills, CA: Sage, 1972).

S58. Determinants of Land Tax Delinquency in Iowa, 1860-1900

Principal investigator: Robert P. Swierenga, Professor, Dept. of History, Kent State U., 305 Bowman Hall, Kent, OH 44242.

Objective: Determination of the factors in land tax delinquency and tax liens in fifteen selected counties in Iowa, 1860-1900, based on the Tax Sale Record Books. *Method:* Multiple regression analysis of the percentage of taxes liened per county per year (dependent variable) and eight independent variables: population density, land and livestock assessed valuation, horse density, railroad density, Warren-Pearson Wholesale Farm Price Index, the Agricultural Terms of Trade Index, and the time factor.

Type of computer: Burroughs B5500. *Size of storage:* 32k words. *Language:* FORTRAN IV. *Disks or tapes:* 3 9-6M character disks; 4 7-channel tapes. Program is available.

Status: To be completed in 1973.

References: "The Tax Buyer as a Frontier Investor Type," *Explorations in Economic History* 7 (Spring 1970), 257-292.

S59. Issues and Factions in the U.S. Congress, 1859-1877

Principal investigator: John L. McCarthy, Assistant Professor, Dept. of History, Yale U., New Haven, CT 06520.

Objective: To describe and explain the changing nature of issues, parties, Congressional institutions, etc., during the Civil War and Reconstruction. *Method:* Roll call vote analysis using various programs developed by the chief investigator, including simple cohesion measures, clustering routines, Guttman scaling, and programs for schematic presentation of results. Analysis of changes in committee structure, incumbency, and a limited number of biographical variables for all members of Congress during the period. Some analysis of constituency variables such as election returns, census data, and changes in Congressional districts.

Type of computer: IBM 370/155. *Size of storage:* 450k. *Language:* FORTRAN IV. *Disks or tapes:* 1 3330 disk (optional), 2 tapes (optional). Program may be available by December 1973.

Status: Continuing.

References: "Reconstruction Legislation and Voting Alignments in the House of Representatives, 1863-1869" (Unpub. Ph.D. dissertation, Yale University Department of History, 1970). (See S68)

S60. Historical Demography Data Archive

Principal investigator: Theodore K. Rabb, Professor, Dept. of History, Princeton U., Princeton, NJ 08540. *Associate:* Myron Gutman.

Objective: To investigate the feasibility of, and hopefully eventually to establish, a computerized data archive, consisting of tapes containing information about family life and population in history. We hope to work out standard ways of coding and analyzing such materials. *Method:* Conferences, inventories of projects completed and in progress, research assistance, and undertaking of pilot research projects with wide methodological implications.

References: "The Planning of a Historical Data Center," *Proceedings of the V International Congress of Economic History* (Moscow, 1970).

S61. Multiple Biography of Anglican Clergymen in Seventeenth- and Eighteenth-Century America

Principal investigator: James B. Bell, Lecturer, Dept. of History, Princeton U., 232 Dickinson, Princeton, NJ 08540. *Associate:* Shirley K. Gilbert.

Scope: An examination of the educational, social, economic, professional, and intellectual characteristics of the parsons. *Method:* 1) Biographic data is being coded and keypunched. 2) Data-Text (or a similar package) will be used for cross-tabulation of relevant variables and PL/I for the creation and analysis of "life patterns" which can subsequently be compared with basic background information.

Type of computer: IBM 7094 or IBM 360/67. *Size of storage:* 768k bytes. [*CHum,* November 1968]

S62. Country House Building and Ownership in Hertfordshire, Northamptonshire, and Northumberland, England, 1540-1879

Principal investigator: Lawrence Stone, Professor, Dept. of History, Princeton U., 131 Dickinson, Princeton, NJ 08540. *Associate:* Jeanne C.F. Stone.

Objective: To codify and analyze patterns of house building and ownership; to relate these patterns to each other as well as to political and economic changes. *Method:* Data is now being coded for keypunching. In analyzing house histories the unit of analysis is an event (positive, negative, change of ownership or residence). These events will be placed on a "house disk" which in turn will be keyed to an "owner disk" and, where relevant, to a "female owner disk." The latter disks will be organized in the form of multiple biographies.

Type of computer: IBM 360/67. *Size of storage:* 768k. *Language:* PL/I F.

Status: Expect to complete in 1974.

References: "Country Homes and Their Owners in Hertfordshire 1540-1879," *The Dimensions of Quantitative History,* ed. W.O. Aydelotte, A.L. Bogue, and R.W. Fogel (Princeton, 1972).

S63. The Structure of Influence Relationships in the International System

Principal investigator: Steven J. Brams, Associate Professor, Dept. of Politics, New York U., Washington Square, New York, NY 10003.

Scope: The theory of directed graphs (digraphs) is used to construct hierarchical influence structures based upon the symmetry/asymmetry of international visits between high-level government officials of all nations in the world in 1964-1965. Longitudinal comparisons are made of the turnover of nations at different levels in the hierarchies between 1964 and 1965, and some hypotheses linking the behavior of nations to their positions in the hierarchies are suggested.

Type of computer: IBM 360/50. *Language:* FORTRAN IV. *Disks or tapes:* 3 tapes. Program is available from Syracuse University Computing Center, Syracuse, NY 13210.

References: "DECOMP: A Computer Program for the Condensation of a Directed Graph and the Hierarchical Ordering of Its Strong Components," *Behavioral Science* 13 (1968), 344-345 (abbreviated version published in *DMG Newsletter* 2, 1968, 5); "The Structure of Influence Relationships in the International System," *International Politics and Foreign Policy,* ed. James N. Rosenau, rev. ed. (New York: Free Press, 1969), pp. 583-599; "Measuring the Concentration of Power in Political Systems," *American Political Science Review* 62 (June 1968), 461-475, reprinted in *Political Power,* ed.Roderick Bell, David V. Edwards, and R. H. Wagner (New York: Free Press, 1969), pp.346-359.

S64. Political Aspects of Urbanization in India

Principal investigator: Robert Sidwar Robins, Associate Professor, Dept. of Political Science, Tulane U., New Orleans, LA 70118.

Scope: 85 characteristics of members of the Indian upper house, 1952-1960. Attempts to determine nature of elite integration in India and its historical development. *Method:* 1) Documentary sources. 2) Statistical analysis using principally four-fold PHI and factor analysis.

Type of computer: IBM 7044. *Size of storage:* 32k words. *Language:* FORTRAN IV. *Disks or tapes:* 1 tape. Program is available.

Status: A new study has been undertaken on political and census (including urban) data, by district and state for all Indian elections to the State Lower House, 1957-1962; aggregate correlational analysis, at the initial stages. [*CHum,* November 1970]

S65. The Funding of the South Carolina Revolutionary Debt

Principal investigator: W. Robert Higgins, Fellow, Institute of Southern History, The Johns Hopkins U., Baltimore, MD 21218.

Objective: An attempt to examine the transactions involved between the state government and holders of revolutionary obligations from the date of issue after 1783 through final cancellations by means of exchange for land, purchase of confiscated property, payment of taxes, or redemption by cash. *Method:* Currently preparing material on worksheets for punching. Data consists of approximately 63,000 bonds and 150,000 to 200,000 separate transactions. Simple acquisitions of material will require several more years.

Type of computer: IBM 360/40. *Size of storage:* 131,000 bytes. *Language:* FORTRAN IV, PL/I.

References: "The South Carolina Revolutionary Debt and Its Holders, 1776-1780," *The South Carolina Historical Magazine* (January 1971).

S66. Collective Violence in Large-Scale Social Change

Principal investigator: Charles Tilly, Professor, Depts. of Sociology and History, U. of Michigan, Ann Arbor, MI 48104.

Objective: Testing and developing theories relating diverse forms of conflict (notably collective violence and strike activity) to urbanization, industrialization, and related large-scale structural changes. *Method:* Assembly of large samples of conflicts in West European countries, 1830-1960; analysis of variations in their forms, intensities, and loci as a function of structural changes occurring in the areas involved.

Type of computer: 1) IBM 360/40, 2) IBM 360/67. *Size of storage:* 1) 125k, 2) 1400k. *Language:* 1) and 2) FORTRAN F. *Disks or tapes:* 1) 1 2314 disk, 10 9-track tapes, 2) 2 data cells, 6 2314 disks, 2 drums; 1 9-track tape.

References: "Quantification in History, as Seen from France," *The Dimensions of the Past,* ed. Jacob M. Price and Val Lorwin (New Haven: Yale University Press, 1972); "How Protest Modernized in France, 1845 to 1855," *The Dimensions of Quantitative Research in History,* ed. W.O. Aydelotte, A.R. Bogue, and R. Fogel (Princeton: Princeton University Press, 1972); David Snyder and Charles Tilly, "Hardship and Collective Violence in France, 1830 to 1960," *American Sociological Review* (October 1972).

S67. Demographic and Social Structural History and Changes in Literacy in England 1550-1850

Principal investigators: Peter Laslett, R.S. Schofield, E.A. Wrigley, Cambridge U., Cambridge, England. *Associates:* K. Oosterveen, R. Davies, R. Wall.

Scope: Designed to provide a quantitative basis for the study of the relationship between demographic, social structural, and literacy variables and economic and social change in England between the sixteenth and nineteenth centuries, and in particular their relation to the Industrial Revolution. The study makes use of record linkage, family reconstitution and social structural analysis techniques.

Type of computer: IBM 360, IBM 370. *Language:* Assembly, PL 360, ALGOL. Some programs are available.

Status: Continuing. About 120 programs are completed making available data on England back to 1538. Bibliography available on request.

S68. Congress and Voters During Reconstruction

Principal investigator: John L. McCarthy, Assistant Professor, Dept. of History, Yale U., New Haven, CT 06520.

Objective: To trace the evolution of Republican and Democratic elites and electorates during reconstruction. It uses roll call analysis of votes in the House of Representatives, from the 36th through the 43rd Congresses, and aggregate analysis of county election and demographic data, 1850-1880.

Status: This project is an extension of S59, examining a wider range of questions in somewhat less detail; most of the analysis has already been completed; I will begin writing sometime in the fall of 1973.

S69. The Radical Right in Historical Perspective

Principal investigator: Seymour Martin Lipset, Dept. of Government and Social Relations, Harvard U., Cambridge, MA 02138.

Scope: An analysis, derived from both voting and survey data, of the social base and ideology of different extremist movements. [*CHum,* November 1968]

S70. Leadership in the New York Liberal Party

Principal investigator: Frank B. Feigert, Dept. of Political Science, Knox College, Galesburg, IL 61401.

Scope: A study of status congruence/incongruence, ethnic group involvement, and ideologies of party leaders. Data is derived from a questionnaire given to all party leaders from the local to the state policy committee levels, and is compared to the Survey Research Center national data. [*CHum,* November 1968]

S71. Partisan Conflict and Cohesion in the United States Senate During the Twentieth Century

Principal investigator: Jerome M. Clubb, Inter-University Consortium for Political Research, Ann Arbor, MI 48106. *Associates:* Howard W. Allen, Aage R. Clausen.

Scope: The study attempts to assess Senate voting alignments and the factors that influenced them. It involves analysis of roll calls in selected Congresses and comparison of voting alignments with various constituency and other characteristics. The study uses Guttman analysis, correlation, and various measures of party cohesion. [*CHum,* November 1968]

S72. Alienation and Race

Principal investigator: Joel D. Aberbach, Institute of Public Policy Studies, U. of Michigan, Ann Arbor, MI 48104.

Scope: Designed to provide conceptual clarification and an empirical test of selected theories and subcultural comparisons. The study uses the Center for Political Studies national samples of blacks and whites, and utilizes factor analysis, analysis of variance, and rank order correlations.

S73. Evolution of Electoral Forces

Principal investigator: Donald E. Stokes, Survey Research Center, U. of Michigan, Ann Arbor, MI 48104.

Scope: An analysis of the historical evolution of forces on the mass electorate within the party systems of the English-speaking world. The study fits various statistical models to historical time series. [*CHum,* May 1969]

S74. Correlates of International War 1815-1945 (Superseded by S57)

S75. Coercion and the Outcome of Revolution

Principal investigator: Diana E.H. Russell, Center of International Studies, Princeton U., Princeton, NJ 08540.*

Objective: The relation between certain regime variables (namely, the resources for coercion at the disposal of a regime, and its willingness to use these resources) and the success or failure of revolution. Quantitative data from ten successful twentieth-century revolutions have been matched against data from ten unsuccessful revolutions. [*CHum,* May 1969]

S76. Western European Integration 1950-1962: Cooperation, Communication, Community

Principal investigator: Donald J. Puchala, Institute of War and Peace Studies, Columbia U., New York, NY 10027.

Objective: To assess movement toward Western European Regional unity in terms of quantitative indicators of intergovernmental cooperation, international social communications, and national and regional mass and elite attitudes, opinions, and outlooks. The study uses quantitative content analysis of intergovernmental interactions, analysis of communications flows in terms of indices of Relative Acceptance, and multivariate and factor analytic assessments of trends in mass and elite opinions. [*CHum,* May 1969]

S77. Lexicographical Analyses and Concordances of Historical Texts

Principal investigator: Dr.J. Untermann, Professor, Institut für Sprachwissenschaft der Universität, 5 Köln, Germany.

Scope and method: Es ist geplant, lexikographische Arbeiten und Konkordanzen historischer Texte maschinell erstellen zu lassen.

Type of computer: IBM 7094. *Language:* FORTRAN II.

S78. A Quantitative Study of Harry S. Truman's Senatorial Voting Behavior

Principal investigators: Gary Fink, Assistant Professor, Dept. of History, Mankato State College, Mankato, MN 56001; James Hilty, Dept. of History, Temple U., Philadelphia, PA 19122.

Scope: An analysis of Senator Truman's voting behavior on major issues in relation to all other members of the Senate from the time of his first election as Senator (1934) until his selection as Vice President (1944). *Method:* Guttman scalograms.

Type of computer: IBM 7040. *Language:* FORTRAN. Program is available.

Status: Research has been concluded (1972) and published in expanded form. More publications are projected.

References: "Metodi Statistici e Storia: Il 'New Deal' Di Truman Attraverso le sue Dichiarazioni di Voto," *Rassegna Economica* 36 (Maggio-Guigno, 1972); "Prologue: The Senate Voting Record of Harry S. Truman," *Journal of Interdisciplinary History* (1973).

S79. Voting Alignments on Tariff Legislation in the U.S. Senate, 1909-1934

Principal investigators: Howard W. Allen, Associate Professor, Dept. of History, Southern Illinois U., Carbondale, IL 62901; Jerome M. Clubb, Director, Historical Archives, Inter-University Consortium for Political Research, Ann Arbor, MI 48104; Aage R. Clausen, Dept. of Political Science, U. of Wisconsin, Madison, WI 53706.

Objective: To examine voting alignments on tariff legislation in the U.S. Senate, 1909-1934. *Method:* Analysis of roll calls using Guttman scaling, correlation, and measures of party cohesion. [*CHum,* May 1969]

S82. Division Lists of the British House of Commons, 1761-1835

Principal investigator: Donald E. Ginter, Associate Professor, Dept. of History, Sir George Williams U., Montreal, PQ, Canada.

Objective: To compile a data bank on the voting records of every member of the British House of Commons, 1761-1835.

Status: All extant division lists have been retrieved. Newspapers and collected debates have been systematically searched; additional data have been obtained from pamphlets, broadsides, and manuscripts. Lists are presently being collated, coded, and placed onto tapes. [*CHum,* May 1971]

S83. An Index to Vol. VI of the *Corpus Inscriptionum Latinarum*

Principal investigator: E.J. Jory, Senior Lecturer, Dept. of Classics and Ancient History, U. of Western Australia, Nedlands, Western Australia, 6009. *Associates:* D.W.G. Moore, Helen Zaliki.

Scope: Establishment of a word index and analysis of dating criteria with a view to further investigation and historical research. *Method:* Coding, punching of paper tapes, editing, transfer to magnetic tapes, more editing, and sorting by the computer.

Type of computer: PDP-6. *Size of storage:* 32k. *Language:* Macro 6. *Special equipment:* Remote-controlled teletype.

Status: Completed in 1972.

References: Revue 2 (1966), Organisation Internationale pour l'étude des langues anciennes par ordinateur; H. Krummrey, "Ein Computer-index zu CIL VI," *KLIO* 52 (1970), 235-239. Completed work to be published by De Gruyter, Berlin.

S84. Computer Analysis of Interaction [*CHum,* May 1969] (terminated 1971)

S85. Systems for Analytical Archaeology

Principal investigator: John G. Rhoads, undergraduate, Dept. of Anthropology, Harvard College, Claverly Hall 33, Cambridge, MA 02138.

Objective: Development of statistical and graphical computer routines useful in the analysis of archaeological data.

Type of computer: 1) IBM 360/65, 2) 360/50, 3) 360/30. *Size of storage:*1) 1024k, 2) 256k, 3) 32k. *Language:*1) ALGOL and OS Assembler F, 2) ALGOL and OS Assembler F, 3) DOS Utilities and DOS Assembler. *Disks or tapes:* 1) 4 2401 7-track tapes, 4 2401 9-track tapes; 2) 4 2401 7-track tapes, 4 2401 9-track tapes; 3) 4 2401 7-track tapes. 1) 3 2311, 1 2314 (9 drives) disks; 2) 4 2311 disks. *Special equipment:* CalComp plotter. [*CHum,* May 1969]

S86 (formerly S11). Pattern of Migration, Social and Economic Mobility of a German Elite Immigrant Society, Cincinnati, Ohio

*Principal investigators:*Zane Miller, Associate Professor, Dept. of History, U. of Cincinnati, OH 45221; Guido Dobbert, Associate Professor, Dept. of Sociology, Youngstown State U., Youngstown, OH 44503.

Objective: Extracting all relevant biographical facts listed in the obituaries published by the society of its deceased members onto IBM cards, for the purpose of analyzing and correlating the patterns as indicated in the title above. Secondary purpose is an analysis of the society's value system in regard to the amount and type of information published on its deceased members. Depending upon the results of the investigation, the scope of research may be

extended to biographical material and obituaries of other ethnic groups. *Method:* Methods consist first of abstracting the data on IBM cards; processing the latter by means of a proofreading program which tests accuracy of transcription and methods of categorization. After the proofreading step, the cards will be transferred onto magnetic tape for permanent storage and statistical analysis, though, for the actual processing, the data will be first transferred onto disks.

Type of computer: 1) IBM 360/40, 2) IBM 360/50. *Size of storage:* 1) 384k, 2) 512k. *Language:* PL/I F. *Disks or tapes:* 1 IBM 2314 disk, 2 IBM 2415 tapes. Program is available. [*CHum,* May 1969]

S87. Data Processing of Geophysical Measurements on Archaeological Sites

Principal investigators: Irwin Scollar, Department Head; I.D.G. Graham, research associate, Labor für Feldarchäologie, Rheinisches Landesmuseum, 53 Bonn 1, Colmanstrasse 14-16, Germany. *Associate:* H. Wauschkuhn.

Scope: Geophysical measurements are made on buried archaeological sites in the Rheinland with high sensitivity magnetometers. The results are recorded directly in the field on paper tape. These tapes are processed to obtain pictures of the buried structures which closely resemble excavation plans and provide a valuable guide to the field archaeologist prior to digging. In many cases the plans are sufficiently clear so that no excavation is needed and the monuments can be classified and protected without further expenditure. *Method:* Production of plots resembling half-tone pictures from two-dimensional numerical data. Processing of this data by Fourier transform techniques as well as nonlinear filtering methods.

Type of computer: IBM 370/165. *Size of storage:* 2 million k. *Language:* FORTRAN IV G and H. *Disks or tapes:* 12 3330 diskpacks, 3 2315 fixed head disks; 8 9-channel, 1 7-channel, 8-channel paper tapes. *Special equipment:* CalComp 584 drum plotter; CalComp 1675 microfilm plotter. Program is available for a fee.

References: I. Scollar, "Magnetic methods of archaeological prospecting-advances in instrumentation and evaluation techniques," *Philosophical Transactions of the Royal Society* 268 (London 1970), 111-121; I. Scollar, D. Gubbins, P. Wisskirchen, "Two dimensional digital filtering with Haar and Walsh transforms," *Annales de Geophysique* 27, 2 (1971), 85-104; I. Scollar, "Magnetic mapping of buried archaeological sites," *Endeavour* 31 (1972), 34-40.

S88. Northern Pacific Railway Company Archives Project

Principal investigator: Helen M. White, Associate Curator, Manuscripts Department, Minnesota Historical Society, Cedar & Central, St. Paul, MN 55101. *Associates:* Maureen Leverty, Deborah Neubeck.

Objective: Modification of the Spindex II program to inventory the Archives of the Northern Pacific Railway Company, in order to meet the needs of current study and evaluation of Company records; to effect their transfer to the custody of the Society; and to make possible their detailed analysis by historians. *Method:* Design data sheets, test on preliminary inventories, keypunch printout, reevaluate, and apply to records studied, department by department. [*CHum,* November 1969]

S89. The Selection of Representative Sites in Archaeology

Principal investigators: Ralph M. Rowlett, Assistant Professor; Michael C. Robbins, Assistant Professor; Dept. of Anthropology, U. of Missouri, Columbia, MO 65201. *Associate:* Richard B. Pollnac.

Objective: To determine which site(s), including both cemetery and habitation areas, of the Marne culture in Champagne, France, are the most typically representative of that culture during the La Tène Ia phase. Likewise, to determine which site(s) are most representative of each of the four constituent Sondergroups of that culture. *Method:* Archaeological sites' material inventories (artifacts) and the spatial relationships of the deposits of these artifacts are submitted to factor analysis to determine which factors are the most common for all sites. Output includes factor loadings.

Type of computer: IBM 360. *Language:* FORTRAN. Program is available.

References: M.C. Robbins, A.V. Williams, P.L. Kilbridge, and R.B. Pollnac, "Factor Analytical Techniques for Case Selection in Complex Societal Research: An Example from Buganda," *Human Organization* 28 (1969), 217-224.

S90. The Social and Economic Structure of Philadelphia, 1760-1775

Principal investigator: James A. Henretta, Assistant Professor, Dept. of History, Princeton U., Princeton, NJ 08540.

Objective: To analyze and describe, in quantitative terms, the economic and social structure of pre-Revolutionary Philadelphia; to determine rates of social and geographic mobility for all wealth and occupation groups within the community. *Method:* To put wealth, occupation, residence, and other material for each of the 4000 individuals listed on the tax lists for 1769 and 1774 onto a separate computer card; to determine the relationship between the above variables. [*CHum,* November 1969]

S91. Biographical Dictionary, with Statistical Sections, of British Colonial Service, 1870-1919

Principal investigator: Vincent Ponko, Jr., Dean of the School of Humanities, California State College, Bakersfield, 615 California Ave., Bakersfield, CA 93304.

Objective: A biographical dictionary with statistical sections of the British Colonial Service 1870-1929. *Method:* Keypunching of coded data IBM cards; machine collation; a printed presentation with appropriate statistical information. [*CHum,* November 1969]

S92. Applications of Numerical Taxonomy to Archaeology [Cancelled] [*CHum,* November 1969]

S93. Méthodes mathématiques et linguistiques en archéologie

Principal investigator: J.C. Gardin, Directeur, Centre d'Analyse Documentaire pour l'Archéologie, Centre National de la Recherche Scientifique, 31 Chemin Joseph Aiguier, 13 Marseille 9e, France. *Associates:* M. Borillo, R. Bourrelly, E. Chouraqui, F. Digard, W. Fernandez de la Vega, M.S. Larange, P. Landau, A. Laurier, N. Nivelle, M.R. Salomé.

Scope: Analyse descriptive des matériaux archéologiques et plus particulièrement de l'anthropologie culturelle. Analyse du contenu des documents écrits, mise au point de systèmes documentaires et de procédures de diffusion de l'information scientifique (bibliographies, index, etc.), construction de modèles linguistiques et surtout mathématiques pour les phénomènes culturels. *Method:* Linguistique, sur le plan syntaxique et surtout sémantique et/ou sémiologique. Mathématique, par des méthodes statistique (analyse factorielle, etc.) et algébrique (analyse combinatoire, graphes, etc.).

Type of computer: UNIVAC 1108, IBM 360/44. *Language:* FORTRAN V, FORTRAN IV. Program is available. Progr. DEND computes and prints a dendrogram of classifications obtained by average method–GNRAL (Recherche Bibliographique). Divers programmes de classification. [*CHum,* November 1969]

S94. Models of Coalition Formation in Voting Bodies

Principal investigator: Steven J. Brams, Associate Professor, Dept. of Politics, New York U., Washington Square, New York, NY 10003.

Scope: Using counting and probabilistic concepts, several theoretical models will be developed to illustrate the calculation of benefits and costs which members of voting bodies might make in decisions to join or not join a coalition. Some concepts from the models will be applied to data on voting in several multi-ballot U.S. national party conventions in an attempt to retrodict election outcomes and evaluate the "rationality" of the delegates' voting. *Method:* In order to extend and generalize the analysis of coalition dynamics to voting bodies with more than ten members (which is the size of a voting body for which hand calculations have been made), the extensive calculations of the models will be undertaken through computer simulation.

Type of computer: 1) CDC 6600, 2) IBM 360. *Language:* 1) and 2) FORTRAN.

References: S.J. Brams and W.H. Riker, "Models of Coalition Formation in Voting Bodies," *Mathematical Applications in Political Science,* VI, ed. J.F. Herndon and J.L. Bernd (Charlottesville: University Press of Virginia, 1972), pp. 79-124; S.J. Brams, "Positive Coalition Theory: The Relationship between Postulated Goals and Derived Behavior," *Political Science Annual: Conflict, Competition, and Coalitions* IV, ed. C.P. Cotter et al. (Indianapolis: Bobbs-Merrill Co., 1973); S.J. Brams and J.G. Heilman, "When to Join a Coalition, and with How Many Others, Depends on What You Expect the Outcome to Be," *Public Choice* (1973); S.J. Brams, "A Cost/Benefit Analysis of Coalition Formation in Voting Bodies," *Probability Models of Collective Decision-Making,* ed. R.G. Niem and H.F. Weisberg (Columbus: Charles E. Merrill Publishing Co., 1972), pp. 101-124.

S95. Role of Television in the Socialization of Children: A Cross-Cultural Survey

Principal investigator: Godwin C. Chu, Associate Professor, Dept. of Anthropology and Sociology, U. of Victoria, Victoria, BC, Canada. *Associate:* Jennifer Jones.

Objective: To assess the role of television in the socialization of children. Data are being collected from three countries: Canada, Australia, and Taiwan. *Method:* Multivariate analysis of questionnaire survey, depth interviews, life history, paragraph completion.

Type of computer: IBM 360/44. *Language:* FORTRAN. Some programs are available. [*CHum,* May 1970]

S96. Computer-Assisted Theory-Analyses

Principal investigators: K.B. Madsen, Professor; Svend Jorgensen, cand. pad.; Inst. of General Psychology, Royal Danish School of Education, Emdrupvej 101, 2400 Copenhagen N.V., Denmark.

Objective: To investigate the possibilities of using computer programs for metascientific analyses and comparison of psychological theories. *Method:* Creation of a computer program inspired by The General Inquirer.

Type of computer: IBM. *Size of storage:* 1800k. *Language:* FORTRAN. Program is available. [*CHum,* May 1970]

S97. Senate and House Voting, 1953-1964

Principal investigator: Aage R. Clausen, Assistant Professor, Dept. of Political Science, U. of Wisconsin, Madison, WI 53705.

Objective: To measure and validate dimensions of policy in roll-call voting with primary attention to dimensions that recur in the six Congresses. Such dimensions permit the analysis of change in the positions taken by Congressmen, including changes that occur when Congressmen are replaced. Additional analysis is concerned with correlates of voting: the impact of party, constituency, presidency, state party delegation. *Method:* Correlational-cluster analysis of roll calls and subsequent scoring of Congressmen on cluster-dimensions. Correlational analysis of relations between scales over time and between scales and other variables. Content analysis is one part of procedure establishing validity of policy dimension scales. (See S71) [*CHum,* May 1970]

S98. Dimensionality of Nations Project

Principal investigator: R.J. Rummel, Professor, Dept. of Political Science, U. of Hawaii, Honolulu, HI 96822.

Scope: The project is concerned with testing a field theory of international behavior. This has involved 1) determining the empirical dimensions of variation in cross-national and time attributes of nations, 2) the dimensions in their dyadic behavior, and 3) the linkages between the two sets of dimensions. Factor analysis, canonical analysis, and multiple regression have been the major techniques used.

Type of computer: IBM 360/65. *Size of storage:* 51k. *Language:* FORTRAN IV. *Disks or tapes:* 3 2314 disks, 5 7- and 9-track tapes. Program is available.

Status: Completed.

References: Applied Factor Analysis (Evanston: Northwestern University Press, 1970); *Dimensions of Nations* (Sage Publications, 1972).

S99. World Handbook of Political and Social Indicators, Second Edition

Principal investigators: Charles L. Taylor, Associate Professor, Dept. of Political Science, Virginia Polytechnic Institute, Blacksburg, VA 24060; Michael C. Hudson, Associate Professor, School of Advanced International Studies, Johns Hopkins U., 1740 Massachusetts Ave., N.W., Washington, DC 20036. *Associates:* John D. Sullivan, Katherine H. Dolan, Edwin H. Dolan, John Dow.

Objective: The presentation of quantitative indicators, relevant to comparative political theory, in the fields of economic, demographic, communications, political structure, governmental change, political conflict, and external relations. *Method:* Assembly of data from international and governmental sources, plus generation of new governmental change and conflict data from multiple press sources, all with extensive discussion of theoretical uses and reliability. A summary factor analysis provides a rough description of the patterns in the data.

Type of computer: IBM 7040, 7094. *Language:* FORTRAN IV. Data are available through the Inter-University Consortium for Political Research, Ann Arbor, MI. *The World Handbook* was published by Yale University Press, 1972.

S100. The Arkansas Archeological Data Bank

Principal investigator: Robert G. Chenhall, Assistant Professor, U. of Arkansas Museum, U. of Arkansas, Fayetteville, AR 72701. *Associate:* Sandra C.Scholtz.

Scope and method: The Arkansas Archeological Survey has professional archeologists assigned to each of the eight State-supported colleges and universities. The purpose of this project is to provide a mechanism for the storage, retrieval, and analysis of all data produced by these archeologists or by amateurs or other professionals working anywhere within

the State. The system and programs developed by the Museum Computer Network (New York City) are being adapted for the processing of archeological data on the computer at the University of Arkansas.

Type of computer: IBM 360/50. *Size of storage:* 512k main storage, 2 million k slow. *Language:* H/1. *Special equipment:* 1-1050, 1-2741, 1 2250, 2-2260 mag card/Selectric typewriter.

Status: Active; approximately 5000 records now in file.

S101. Computerized Content Analysis of Presidential Campaign Speeches and Newscasts

Principal investigator: John W. Ellsworth, Professor, Dept. of Government, Southern Illinois U., Edwardsville, IL 62025.

Objective: To measure the varying verbal styles and substantive content of Presidential candidates through time (1928-1972), and between varying campaign situations. The research seeks to measure the influence of audience size, the mass media, and other campaign situations on speech content, and to compare speech content with news media reports on the campaigns studied. *Method:* The General Inquirer method of content analysis is being utilized with a special campaign dictionary developed by the investigator.

Type of computer: IBM 370/135. *Size of storage:* 240k. *Language:* PL/I. *Disks or tapes:* 6 1311 disks, 3 tapes. Program is available. Write to Washington University, St. Louis, MO.

S102. Epistemometrics (A Case Study in Epistemological Sociology)

Principal investigator: G.C. Ankerl, Professor, School of Planning, M.I.T., 86 Massachusetts Avenue, P.O.Box 47, Cambridge, MA 02139. *Associates:* S. Cote, M. Camus, B. Sokoloff, Six Hamard.

Objective: To make more specific and semiautomatic the testing technique of scientific works, intensive critique of books and other lengthy works. *Method:* 1) Transcription of Max Weber's *The Protestant Ethic* in machine-readable form. 2) Symptomatic analysis: Adaptation of KWIC 6000 for an epistemological preclassification of the works (localization, wordstock, frequency statistics in relation to the technical character of the terms). 3) Programming of criteria application (semiotic, cognitive, and noncognitive elements, logical analyses and verification methods) for semiautomatic book critique simulation.

Type of computer: CDC 6400. *Size of storage:* 65k. *Language:* KWIC 6000. *Disks or tapes:* 1 tape.

Status: Expect to finish in 1975.

References: "L'Economisme Historique et l'Imputation Causale Weberienne," *Les Sociologiques Neokantiens Allemands* (Paris-Montreal: PUF-PUM, 1969); "From Intuitive to Formal Book Critique: A Semiotic Approach," *Heuristics* 1 (1972), 35-44; *Sociologues Allemands, avec le dictionnaire de l'Ethique protestante et l'Esprit du Capitalisme de Max Weber* (Paris: Payot-Baconnière, 1972).

S103. Landmarks in American Scientific Research during the Nineteenth Century: A Bio-Bibliography

Principal investigator: Phyllis A. Richmond, Professor, School of Library Science, Case Western Reserve U., Cleveland, OH 44106.

Objective: To use a bibliography as a method of historical investigation by adding quantitative studies, aided by computer. For example, to determine which sciences progressed and which lagged behind European research; to discover in the educational, job, and other background of the scientific investigators why some fields were more productive than others; to determine to what extent financial support aided the more progressive sciences; to indicate publication patterns and deduce some indication of cultural factors involved which made such patterns necessary; to show by means of honors received the reputation and possible influence of the scientists whose works were considered outstanding. The method of selection is principally by consensus of scientists, drawing upon their views about which scientists were major contributors in their fields. *Method:* Collect data from histories of science and especially from sourcebooks compiled by scientists; look up the *exact* titles of books and articles in which research results were announced; collect biographical data on each scientist for purposes of analysis.

References: "The 19th Century American Physician as a Research Scientist," *International Record of Medicine* 171, 8 (August 1958), 492-506. [*CHum*, May 1970]

S104. Presidential Popularity 1945-1969

Principal investigator: John E. Mueller, Professor, Dept. of Political Science, U. of Rochester, Rochester, NY 14627.

Scope: Regression analysis of popularity times series.

Type of computer: IBM 360.

References: "Presidential Popularity from Truman to Johnson," *American Political Science Review* (March 1970); *War, Presidents and Public Opinion* (New York: Wiley, 1973).

S105. Computer Analysis of Presidential Texts

Principal investigator: John H. Kessel, Professor, Dept. of Political Science, Ohio State U., 154 North Oval Drive, Columbus, OH 43210. *Associates:* Evelyn Small, Gerald Stacy.

Scope: The current research extends an earlier analysis of State of the Union messages through 1973 and broadens the range of texts in the analysis to include 1972 convention and campaign statements. The results will be combined with survey data to detect effects of campaign strategy, and analyzed to determine whether the policy areas and concepts uncovered in the earlier analysis of State of the Union messages are also characteristic of nomination and electral politics.

Type of computer: IBM 360/75.

Status: Continuing.

References: "The Parameters of Presidential Politics," a paper delivered at the meeting of the American Political Science Association, September, 1972.

S106. United Nations Peacekeeping, 1947-1969

Principal investigator: H.R. Alker, Jr., Professor, Dept. of Political Science, M.I.T., E. 53-407, Cambridge, MA 02139. *Associates:* C. Alger, C. Christensen, W.J. Greenberg, L. Groff.

Scope: A variety of papers modeling UN peacekeeping successes and failures and their impact on constitutional innovations in the UN system. *Method:* 1) Computer simulation of certain crucial cases: Congo, Article 19 crisis, Middle East, etc.; 2) nonlinear causal model of success, activity levels; 3) simulated learning model of diplomatic practice.

Type of computer: IBM 360/65.

Status: Completed June 1973.

References: H.R. Alker, Jr., and W.J. Greenberg, "The UN Charter: Alternate Pasts and Alternate Futures," *The United Nations: Problems and Propects,* ed. Edwin H. Fedder (Columbia: Center for International Studies, University of Missouri, 1971); H.R. Alker, Jr., and C. Christensen, "From Causal Modelling to Artificial Intelligence: The Evolving of a UN Peace-making Simulation," *Experimentation and Simulation in Political Science,* ed. LaPonce and Smoker (Toronto: University of Toronto Press, 1972).

S107. The Archaeological Provenience of Amber Artifacts

Principal investigator: Curt W. Beck, Professor, Dept. of Chemistry, Vassar College, Poughkeepsie, NY 12601.

Objective: To determine the geographical origin of amber artifacts found in archaeological contexts and to use the results of these determinations to map trade relations in prehistoric Europe and Asia. *Method:* Physical analysis by a number of multiparameter methods, principally infrared spectroscopy and nuclear magnetic resonance, followed by computer classification into geographical groups and/or mineralogical species and varieties.

Type of computer: IBM 360/30. *Size of storage:* 32k. *Language:* FORTRAN IV. *Disks or tapes:* 2 2311 disks, 4 9-track tapes. Program is available.

References: C.W. Beck, G.C. Southard, and A.B. Adams, "Analysis and Provenience of Minoan and Mycenaean Amber. II. Tiryns," *Greek, Roman, and Byzantine Studies* 9, 1 (1968), 5-19; C.W. Beck, "Richerche sulla provenienza di manufatti archeologici d'ambra," *Quaderni di Merceologia* 7, 1 (1968), 1-23; C.W. Beck, C.A. Fellows, and A.B. Adams, "Analysis and Provenience of Minoan and Mycenaean Amber. III. Kakovatos," *Greek, Roman, and Byzantine Studies* 11, 1 (1970), 5-22; C.W. Beck, A.B. Adams, G.C. Southard, and C.A. Fellows, "Determination of the Origin of Greek Amber Artifacts by Computer Classification of Infrared Spectra," *Science and Archaeology,* ed. R.A. Brill (Cambridge: MIT Press, 1971), pp. 235-240.

S108. Rose Bibliography and Index of American Civilization 1865-1917

Principal investigator: Robert Walker, Professor of American Civilization, American Studies Program, George Washington U., Washington, DC 20006.

Scope: Indexing and bibliographic retrieval system in American social, economic, and political thought, c.1865-1915. Over 4000 articles from *The Voice of the Negro,* the *Colored American, Alexander's,* and the *Crisis* have been indexed and stored on disks. *Method:* Articles are indexed for subject matter, keypunched, data formatted on computer disks, disks are searched, results are printed.

Type of computer: IBM 370/145. Program is available, depending on the circumstances.

Status: Temporarily inactive.

S109. The Bargaining Process in International Arms Control Negotiations

Principal investigators: P. Terrence Hopmann, Assistant Professor; Charles Walcott, Assistant Professor; Dept. of Political Science, U. of Minnesota, Minneapolis, MN 55455.

Objective: To investigate the effects of a number of factors on the bargaining process in international arms control negotiations, including the external environment, the national goals of negotiating nations, the perceptions and attitudes of negotiators, and the bargaining strategies and styles employed in the negotiations. The study involves the use of comparative research in the "real world" based on the negotiations in the Conference of the Committee on Disarmament (formerly the Eighteen Nation Disarmament Conference) from 1962 to 1972 and in an experimental laboratory environment. *Method:* Methods for the analysis of real world negotiations include: 1) content analysis of samples of the verbatim texts of negotiations; 2) coding through Bargaining Process Analysis of interactions in all negotiations during the period; 3) interviews with negotiators and staffs of relevant international organizations; 4) scaling of international events to measure intensity of conflict and cooperation. The experimental work involves analysis of small group bargaining processes, so that hypotheses about the bargaining process itself may be tested under controlled conditions and so that new hypotheses may be generated. In all cases, computers are used to generate data tables, and then for analyzing statistical relationships as hypothesized among variables.

Type of computer: 1) IBM 360/50, 2) CDC 6600. *Size of storage:* 1) 356k, 2) 65k words. *Language:* PL/I, FORTRAN IV. Program is available.

References: P.T. Hopmann, "Internal and External Influences on Bargaining in Arms Control Negotiations: The Partial Test Ban," *Peace, War, and Numbers,* ed. Bruce M. Russett (Beverly Hills: Sage Publications, 1972); C. Walcott and P.T. Hopmann, "Interaction Analysis and Bargaining Behavior," *The Small Group in Political Science* (forthcoming), ed. R.S. Golembiewski.

S110. Computer Applications in Archaeology

Principal investigator: John D. Wilcock, Senior Lecturer, Dept. of Computing, North Staffordshire Polytechnic, Blackheath Lane, Stafford,ST18 0AD England. *Associate:* F.S.C. Celoria.

Objective: Information retrieval: Roman inscriptions, prehistoric carvings, British cave archaeological and palaeontological finds (in collaboration with the British Cave Research Association), pottery file, full catalog (for Ribchester Museum, Roman collection), archaeological sediments, projectile points, plant species for pollen analysis; statistics (MDSCAL, agglomerative and divisive procedures, k-means, dendrograms); pottery and projectile point profile analysis (using d-Mac pencil follower); graphics; analysis of field readings from proton gradiometer, etc. All these facilities are interlinked in the PLUTARCH (*P*rogram *L*ibrary *U*seful *T*o *ARCH*aeologists) System, which is controlled by interactive graphics, using a light pen, light buttons, function switches and console. *Method:* Information retrieval data will be compatible with the archaeological format of the Information Retrieval Group of the Museums Association (IRGMA), which the author helped to design. Field data has been transmitted by remote terminal from archaeological sites. Various graphical techniques are used in the presentation of results (scalograms, dendrograms, distribution maps, histograms, piecharts, various text styles, 3D perspective rotation of site block diagrams, sections, pottery and projectile point diagrams, pollen diagrams, instrument plots, permutation of artifacts to detect trend in style).

Type of computer: 1) ICL 4130, 2) ICL System 4-50, 3) DEC PDP-8. *Size of storage:* 1) 64k, 2) 128k, 3) 4k. *Language:* 1) ALGOL 60, 2) ALGOL 60, 3) PAL III. *Disks or tapes:* 1) 4 disks, 4 tapes, 2) 2 disks, 4 tapes, 3) 2 tapes. *Special equipment:* d=Mac pencil follower, line-drawing display unit with light pen and function switches, digital incremental plotter. Program will be available in *The Computer in Archaeology* (Academic Press, forthcoming).

References: J.D. Wilcock, *Transactions of the Cave Research Group of Great Britain* 12, 1 (1970), 96-98, Ledbury, Herefordshire, England; *Science and Archaeology* 2, 3 (1970), 12-16; *Science and Archaeology* 2, 3 (1970), 27-29, 58, 59; *Prospezioni Archeologiche* 4, 85-93, and 5 (1970), 55-58, Fondazione Lerici, Rome; *Mathematics in the Archaeological and Historical Sciences* (Edinburgh: Edinburgh University Press, 1971), pp.470-481; *Computing and Computers* PM 951 (The Open University Press) Unit 8, pp.82-30.

S111. Occupational and Geographic Mobility in Atlanta, 1870-1920

Principal investigator: Richard J. Hopkins, Instructor, Dept. of History, Ohio State U., Columbus, OH 43210.

Objective: To determine the occupational and geographic mobility (both persistence rate and residential movements within the city) of all males aged 20-29, and their sons, in the 1870 and 1880 U.S. census manuscripts, according to three racial-ethnic categories (native white, foreign-born, and negro); attempt to correlate occupational mobility (or lack of it) with residential mobility and patterns. *Method:* 1870 and 1880 groups are drawn from the Federal census schedules; city directories are used for determining continued residence, occupation and address for 1890, 1900, 1910, and 1920. All data coded on individual cards for machine processing. [*CHum,* May 1970]

S112. Sweden During the Second World War

Principal investigators: F. Lindberg, S.U. Palme, and G.T. Westin, Professors, Dept. of History, U. of Stockholm, Teknologgt. 8 E, 113 60 Stockholm VA, Sweden. *Associates:* P. Haas, A. Johansson, J. Melin, K. Amark.

Scope: Information retrieval system of articles (c. 150,000) in Swedish newspapers during the period 1938-1945. *Method:* 1) Transference of main articles' contents to optical page, 2) reading of optical page, 3) decoding and checking, 4) print in tabular form.

Type of computer: 1) IBM 1231 optical reader, 2) IBM 1130, 3) CDC 3600. *Size of storage:* 32k. *Language:* 3600-FORTRAN. *Disks or tapes:* CD 603 tapes. *Special equipment:* 2 magnetic drums. Program is available. [*CHum,* November 1970]

S113. Decision-Making in International Politics

Principal investigator: Ib Damgaard Petersen, Associate Professor, Dept. of Social Sciences, Institute of Political Science, U. of Copenhagen, Rosenborggade 15, Copenhagen K, Denmark. *Associates:* Karl Landberg, Iven Reventlow.

Objective: To develop new methods in quantitative international politics on the basis of investigations of single systems delimited by the historical categories. Material concerning frequencies of actions in selected dimensions of action. Initial project, some 500 historical data. Extension up to 4000 data stipulated for next stage. *Method:* Probability model employed.

Type of computer: GIER. *Size of storage:* 5k. *Language:* ALGOL. *Disks or tapes:* CDC disks.

S114. Roll-Call Analysis for Norwegian Storting (Parliament)

Principal investigator: Ottar Dahl, Professor, Historisk Institutt, P.B.1008, Universitetet i Oslo, Blindern, Oslo 3, Norway. *Associates:* Ivar Fonnes, P.M. Bryhn.

Scope: Collection of data for the period 1814-1940. Analysis of group structures in Storting. Analysis of debating themes, argument types, etc., in the postwar debates of the Norwegian Storting to determine differences among representatives and changes with time with regard to attitudes to foreign policy, economic activities of the government, etc. *Method:* Punching of complete text of selected debates, production of frequency word lists, searching for keywords, and printing out occurrences in context. Comparison of frequency and types of occurrences between debates and between representatives.

Type of computer: CDC 3300. *Language:* FORTRAN IV. Program is available.

Status: To be completed 1973.

References: O. Dahl, "Groups in the Norwegian Storting 1892-1897," *Historisk Tidsskrift,* 1971; Ivar Fonnes, "The Opposition in the Storting 1872," *Historisk Tidsskrift,* 1972; unpublished studies available.

S115. Danish Rural Population in the Eighteenth Century

Principal investigator: Hans Chr. Johansen, Professor, Dept. of Economic History, U. of Odense, DK-5000 Odense, Denmark.

Objective: Family reconstitution of Danish parishes comprising 1 percent of Danish rural population in the eighteenth century. *Method:* Family reconstitution of parish registers and information in the censuses of 1787 and 1801.

Type of computer: 1) IBM 1130, 2) IBM 360/75. *Size of storage:* 1) 8k, 2) 250k. *Language:* 1) and 2) FORTRAN IV. *Disks or tapes:* 2) 9-track tapes. Program is available.

S116. Social Change in the Finnish Society

Principal investigator: Paavo Seppänen, Professor, Institute of Sociology, U. of Helsinki, Franzeninkatu 13, Helsinki 50, Finland.

Scope: Time series concerning different aspects of Finnish society. The number of time series is about 100, data from each year since 1911. *Method:* Punch cards, correlations, factor analysis.

Type of computer: IBM 1620. *Language:* FORTRAN.

References: "Muuttuva yhteiskunta," *Sosiologia* 2 (1965); "Om samhällets centrala förändringsprocesser," 47 (1965), Institute of Sociology, University of Helsinki; "Historian tutkimus ja tilastot, Aikamme historian tutkimus," *WSOY* (1967). [*CHum,* November 1970]

S117. The Swedish Electorate, 1866-1968

Principal investigator: Leif Lewin, Assistant Professor, Dept. of Political Science, U. of Uppsala, Skytteanum, 751 20 Uppsala 1, Sweden. *Associates:* Bo Jansson, Dag Sörbom, Lena Marcusson, Dan Holmbäck.

Scope: Analysis of professional status—more than 150 professions are used—and some other socioeconomic predictors of electoral behavior in the Swedish electorate since the end of the last century. New method for classification of professions into social "classes." Multidimensional scaling of parties' position. *Kind of material:* Aggregate data. *Level of observation:* Counties. *Method:* Canonic classification, quadratic programming, factor and component analysis.

Type of computer: 1) IBM 360/75, 2) CDC 3600. *Size of storage:* 1) 1535k, 2) 192k. *Language:* FORTRAN IV. [*CHum,* November 1970]

S118. Profession and Voting Pattern; A Historical-Statistical Investigation of the Social Structure of the Liberal Movement in Sweden, 1911-1920

Principal investigators: Jörgen Weibull, Professor of Contemporary History; Christer Weibull, Docent, lecturer of statistics, Institute of Contemporary History, U. of Aarhus, 8000 Aarhus C, Denmark.

Scope: Comparing party affiliation and voting returns in different election districts and at different locations; determining the voting pattern of different professional groups at each election and the changes in that respect 1911-1920, by a distribution of the voting figures from a regional, social, and professional point of view. *Method:* Correlation methods and quadratic programming with linear restraints.

Type of computer: 1) FACIT Edb 3, 2) SAAB D21. *Language:* ALGOL. [*CHum,* November 1970]

S119. The Popular Movements in Sweden and the Elections to the Second Chamber of the Swedish Riksdag, 1899-1920

Principal investigator: Sven Lundkvist, Docent, Dept. of History, Uppsala U., Kungsängsgatan 19, Uppsala, Sweden. *Associate:* Ulla Hedqvist.

Objective: To find out the popular movement activity in the elections to the second chamber and influencing factors such as socioeconomic, political and ethical backgrounds.*Method:* All data are transformed to optical data sheets and stored on tape, sorted according to the election results in the municipalities of Sweden, to the party belonging, to the type of socioeconomic structure, to the popular movement structure, and then analyzed with fitting statistical methods.

Type of computer: CDC 3600. *Size of storage:* 32k. *Language:* FORTRAN IV.

References: C.G. Andrae and S. Lundkvist, "Folkrörelserna och den svenska demokratiseringsprocessen," *Svensk historisk tidsskrift* (1969), pp.197-214; C.G. Andrae, "Popular Movements in Sweden: Report on a Mass-Data Project," *Social Science Information* 8, 65-75. [*CHum,* November 1970]

S120. Ecological Studies in Danish Voting Behavior

Principal investigator: Jan Stehouwer, Assistant Professor, Dept. of Sociology, Institute for Political Science, U. of Aarhus, 8000 Aarhus C, Denmark. *Associates:* Ole Borre, Erik Lund.

Objective: To analyze differences in level of strength of political parties between communes and constituencies; to analyze historical trends and short-term fluctuations in electoral behavior. *Method:* Ecological: comparison between ecological units in tabular form, correlation between political and demographical attributes of ecological units and changes in these attributes.

Type of computer: 1) GIER, 2) IBM 7090. *Language:* 1) ALGOL 4, 2) FORTRAN. *Disks or tapes:* 12 IBM 729 tapes.

References: O. Borre and J. Stehouwer, *Partistyrke og social struktur 1960 (Party Strength and Social Structure, 1960),* Aarhus, 1968; *Fire Folketingsvalg (Four General Elections),* Aarhus 1969. [*CHum,* November 1970]

S121. Social Overhead Capital in Swedish Economic Growth, 1860-1914

Principal investigator: Lars-Erik Hedin, Fil. lic. (Ph.D.) Research Assistant, Dept. of Economic History, Finngatan 16, Lund, Sweden.

Objective: To measure the impact of investment in infrastructure on regional economic growth with special reference to investments in railroads, roads, canals, and ports. Tax figures are taken from unprinted material up to 1881 and after that from printed sources; 24 counties (län) and 100 cities/towns are included. No sampling. Other variables are population, income/capita, and workers. Investments in railroads and infrastructure in general are taken from printed statistics. *Method:* Regional economic growth is measured by using income tax figures for each county and city/town. The rate of growth is then compared with the investment in Social Overhead Capital. Correlation analysis will be used, but the important thing is to get some knowledge of regional economic growth.

Type of computer: IBM 360/50. Program is available. [*CHum,* November 1970]

S122. A History of Prices in Sweden, 1732-1914

Principal investigator: Lennart Jörberg, Associate Professor, Dept. of Economic History, Lund U., Finngatan 16, 223 62 Lund, Sweden.

Scope: A general description and analysis of price movement and change in price structure for 30 regions in Sweden. Sixty commodities are covered. *Method:* Sorting and tabulation, statistical manipulations; coefficient of variation, autocorrelation, indexes, range, weightings, etc.

Type of computer: ITC 1900, UNIVAC 1108.

Status: Book was to be published November 1972 in 2 volumes by C.W.K. Gleerup, Lund, Sweden.

S123. Railroad Transports and Business Cycles in Sweden, 1860-1914

Principal investigator: Jens Ahlfors, student, Dept. of Social Sciences, Institute for Economic History, U. of Gothenburg, Stora Nygatan 23-25, S-411 08 Göteborg, Sweden.

Objective: To study transport series from about 50 Swedish railroads in their relations to business cycles. *Method:* Seasonal adjustment with the seasonal adjustment program of the National Institute of Economic Research (Konjunkturinstitutet), Stockholm, Sweden.

Type of computer: IBM 360/75. [*CHum,* November 1970]

S124. Economic Growth and Population Change in Sweden, 1770-1880

Principal investigator: Gunnar Fridlizius, Associate Professor, Dept. of Economic History, U. of Lund, Finngatan 16, Lund, Sweden.

Scope: An analysis of the mutual interaction between demographic and economic variables in different regions and economies. Sample is 70 parishes. *Method:* Sorting and tabulation, cross-tabulation, etc.

Type of computer: CDC 3600. *Language:* FORTRAN. [*CHum,* November 1970]

S125. Migration in Southern Sweden, 1860-1910

Principal investigator: Gunnar Fridlizius, Associate Professor, Dept. of Economic History, U. of Lund, Finngatan 16, Lund, Sweden.

Scope: Description and analysis of the emigration from the south of Sweden 1860-1910 compared with the migration of the same region. *Method:* Sorting and tabulation, cross-tabulation, etc.

Type of computer: CDC 3600. *Language:* FORTRAN. [*CHum,* November 1970]

S126. Postwar Immigration to Sweden

Principal investigator: Oscar Bjurling, Professor, Dept. of Economic History, U. of Lund, Finngatan 16, Lund, Sweden. *Associates:* Rolf Ohlsson, Rune Rydén.

Scope: Description and analysis of the postwar immigration to Sweden, compared to the migration from Sweden; macro- and microstudies. *Method:* Sorting and tabulation, multiple correlation, matrix analysis, cross-tabulation, etc.

Type of computer: IBM 360/150 H. [*CHum,* November 1970]

S127. Certain Criteria of Social and Economic Assimilation Among Immigrants in Sweden, 1920-1945

Principal investigator: Sven Nordlund, MA in social sciences, Dept. of Economic History, U. of Gothenburg, Stora Nygatan 21-23, Göteborg, Sweden.

Scope: Sample 4500 individuals from a population of about 150,000 male immigrants.

Type of computer: IBM 360/50. *Size of storage:* 256k. *Language:* ISR. *Disks or tapes:* 1 1311 disk, 3 tapes. Program is available. Write to Computing Center of the University of Gothenburg. [*CHum,* November 1970]

S128. The Development of the Swedish Welfare, 1925-1960

Principal investigator: Karl-Gustav Hildebrand, Professor, Dept. of Economic History, U. of Uppsala, S:t Olofsgatan 10 A, Uppsala, Sweden. *Associates:* Margareta Engkvist, Guy Andersson, Karsten Lundequist, Allan Larsson, Uno Söderman.

Objective: Describe and analyze the development of income distribution among different groups in Sweden 1925-60, and investigate the effect of transfers on the family income available. *Method:* Analysis of correlation, print in tabular form, cross-tabulating.

Type of computer: CDC 1800. *Language:* INTEREST II. Program is available. [*CHum,* November 1970]

S129. Sweden and America After 1860

Principal investigator: Sune Akerman, Professor, Dept. of History, Section of American History, U. of Uppsala, S:t Larsgatan 2, Uppsala, Sweden. *Associates:* Torsten Augrell, Anders Haglund.

Objective: To investigate Swedish emigration to the USA and Canada. We have selected a certain number of years, i.e., 1869, 1887, 1894, 1903, 1908, 1918, 1923. This means that we are covering about 300,000 emigrants (25 percent of the total emigration from Sweden). We are especially interested in the following: Retrieval and registration; volume of the emigration; control of the sources (passenger lists and church registers, official statistics); structural analysis; test of certain theories in migration research. *Method:* We plan to investigate some selected years with many or few emigrants and to contrast these years to each other. This central investigation is also combined with different case studies which will allow us to complete our analysis of some special problems, and also to make very thorough investigations in a few small geographical units.

Type of computer: CDC 3600. *Size of storage:* 32k. *Language:* FORTRAN-63. *Special equipment:* 2 magnetic drums, RA-AREA. Program is available.

References: Planned presentation in 1970 winter issue of *American Studies in Scandinavia* (issued by Nordic Association for American Studies in Scandinavia, S:t Larsgatan 2, 752 20 Uppsala, Sweden). [*CHum,* November 1970]

S130. Membership Structure in a Trade Union in Västeras, 1895-1910

Principal investigator: Torgny Börjeson, Research Assistant, Dept. of History, U. of Uppsala, Kungsängsgatan 19, Uppsala, Sweden.

Objective: To find out factors which may have influence on workers joining a trade union but also on their resignation of membership. *Method:* The program gives in tabular form the joiners and those who resigned per month in respect to their age and time in work. It permits also calculations with percent, medians, and quartiles.

Type of computer: CDC 3600. *Size of storage:* 32k. *Language:* FORTRAN IV. Program is available. [*CHum,* November 1970]

S131. Eskilstuna and the Popular Movements, 1870-1920. Socioeconomic Structure and Political Activity

Principal investigator: Bo Ohngren, Research Assistant, Dept. of History, U. of Uppsala, Kungsängsgatan 19, Uppsala, Sweden. *Associate:* Ulla Hedqvist.

Scope: Total investigation of adults (over 15 years) in the city of Eskilstuna, 1881, 1890, 1896, and 1902. For the years of 1910 and 1920 a stratified sample will be taken. The purpose is to describe and analyze the impact of industrialization on the socioeconomic structure and its further influence on the voting behavior and the recruitment to the popular movements. To a great extent the analysis will cope with both geographic and social mobility. In order to obtain more generalized results as to the political behavior, we have collected individual election data from 8 other cities. Time span, 1890-1902. *Method:* All the data will be transferred to optical data sheets and stored on tape, sorted by ID, occupation, age and location categories. *Statistical methods:* Frequency distributions, cross-tabulation, *chi* square, diff. types of regression analysis, variance analysis and cohort analysis.

Type of computer: 1) IBM 1231, 2) IBM 1130, 3) CDC 3600. *Size of storage:* 32k. *Language:* FORTRAN IV. Program is available.

References: C.G. Andrae, "Popular Movements in Sweden: Report on a Mass-Data Project," *Social Science Information* 8, 65-75. [*CHum,* November 1970]

S132. Social and Geographical Mobility Among Workers in the Swedish Engineering Industry at the End of the Nineteenth Century

Principal investigator: Bo Ohngren, Research Assistant, Dept. of History, U. of Uppsala, Kungsängsgatan 19, Uppsala, Sweden. *Associate:* Ulla Hedqvist.

Scope: This data file Person 2 is composed of information drawn from a survey made by the Swedish Board of Commerce in 1899 and 1901. This file provides data on 35 variables of 23,000 workers all over the country. The aim is to construct mobility matrices for both geographical and social mobility, and also connect these results with the recruitment to the popular movements. Besides, we have aggregate data on geographical mobility from the Census 1870, 1880, 1890, and 1900. *Method:* Mobility matrices, factor analysis (program IMAGE and INTEREST), regression analysis, cross-tabulation.

Type of computer: 1) IBM 1231, 2) IBM 1130, 3) CDC 3600. *Size of storage:* 32k. *Language:* FORTRAN IV. Program is available. [*CHum,* November 1970]

S133. The Role of Technicians and Scientists in the Industrial Revolution in Sweden

Principal investigators: G. Eriksson, N. Runeby, R. Torstendahl, Dept. of History, U. of Uppsala, S:t Larsgatan 2, Uppsala, Sweden.

Objective: To explain the result of a technical education for the careers in private industry and governmental work. *Method:* Correlations between different variables concerning individuals.

Type of computer: CDC 3600. *Size of storage:* 32k. *Language:* FORTRAN II and FORTRAN IV. *Disks or tapes:* 4 CDC 603 tapes, 2 CDC 601 tapes. *Special equipment:* Magnetic drum 1M words. [*CHum,* November 1970]

S134. Family Reconstitution for Helligandssogn, Copenhagen, 1713-1804

Principal investigator: Poul Thestrup, Stud. mag. art., Institute of Economic History, Frederiksholms Kanal 4, 1220 Copenhagen K, Denmark.

Objective: To use the family reconstitution technique on a city material and to investigate the reconstituted families' representativeness of the city population. *Method:* Manual reconstitution of the families; computer calculations of the reconstituted families. [*CHum,* November 1970]

S135. Statistical Content Analysis of an Icelandic Population Census (three counties) from the Year 1729

Principal investigator: Hans Oluf Hansen, Assistant Professor, Institute of Statistics, U. of Copenhagen, Skt. Pedersstraede 19 I, 1453 Copenhagen K, Denmark.

Objective: Development of computer programs for automatic processing of similar censuses in other Scandinavian countries in the eighteenth and nineteenth centuries; production of data elucidating Icelandic demographic conditions in the first half of the eighteenth century. *Method:* Development of coding scheme and coding instructions. On request, these materials as well as computer programs and punched data will be available on payment of technical copying expenses.

Type of computer: IBM 7090/7094. *Language:* FORTRAN IV. *Disks or tapes:* 729 magnetic tape. Program is available.

Status: Completed 1971.

References: The Icelandic Population Census of 1729 (duplicated report 1971), forthcoming in *Hagskýrslur Islands* (Official Statistics of Iceland).

S136. Planning of Statistical Content Analysis by Electronic Data Processing of Icelandic Land Registers Originating Before 1860 on the Basis of the Land Register from 1702-1714 of Arni Magnusson

Principal investigator: Hans Oluf Hansen, Assistant Professor, Institute of Statistics, U. of Copenhagen, Skt. Pedersstraede 19 I, 1453 Copenhagen K, Denmark.

Objective: Elucidation of fundamental economic, social, and demographic structures in a closed pre-industrial agrarian society considering the many population censuses from the eighteenth and nineteenth centuries. *Method:* The planning aims at development of comprehensive coding instructions, coding schemes, and computer programs for sorting and counting the data; when completed these materials will be available for scientific institutions and interested single persons on payment of technical copying expenses.

Type of computer: IBM 7090/7094. *Language:* COBOL and FORTRAN IV. *Disks or tapes:* 729 magnetic tape. Program is available.

Status: Planning completed 1971.

S137. Mortality as a Determinant of Age Structures and of Population Growth: Selected Countries, 1936-1965

Principal investigator: Hans Oluf Hansen, Assistant Professor, Institute of Statistics, U. of Copenhagen, Skt. Pedersstraede 19 I, 1453 Copenhagen K, Denmark. *Associate:* Edda Sveinsdóttir.

Objective: To make a contribution to better understanding of population growth and dynamics of age structures in pre-industrial agrarian Scandinavian countries in the seventeenth-nineteenth centuries. *Method:* Preliminary sorting of age-specific patterns of mortality by aid of computed regression parameters; and further on, advanced numerical analysis of the subdivided materials originating from countries with rather homogeneous standards of economic and medical development.

Type of computer: IBM 370/165. *Language:* FORTRAN IV. *Disks or tapes:* 729 magnetic tape. Program is available and punched data as well in exchange for problem data (punched) from the fields of demography and sociology.

S138. Construction of Life Tables and Evaluation of Stable Population Analysis in a Pre-Industrial Closed Agrarian Scandinavian Society

Principal investigator: Hans Oluf Hansen, Assistant Professor, Institute of Statistics, U. of Copenhagen, Skt. Pedersstraede 19 I, 1435 Copenhagen K, Denmark. *Associate:* P. C. Mathiessen.

Objective: Development of methods of constructing life tables as a basis of advanced demographic and numerical analysis, e.g., of marriages, births, and different aspects of mortality. *Method:* 1) Computation of annual age distribution and age-specific rates of mortality 1768-1861; 2) investigation of levels of mortality expressed in crude rates and in age-specific rates 1768-1861; 3) projection of annual age-specific rates of mortality on the basis of the crude rate of mortality 1735-1767 and a population census from c. 1762.

Type of computer: IBM 370/165. *Language:* FORTRAN IV. *Disks or tapes:* 729 magnetic tape. Program is available.

References: Nations Unies, ed., "Le Concept de population stable, application à l'étude de populations des pays ne disposant pas de bonnes statistiques demographiques," *Etudes demographiques* 39, ST/SOA/Series A939, New York (1966).

S139. Internal Migration Patterns and Overseas Migration from a Danish Parish c.1870-c.1920

Principal investigator: Kristian Hvidt, Chief Librarian, Library, Danish Parliament (Folketinget), Christiansborg, 1218 Copenhagen, Denmark. *Associate:* Hans Oluf Hansen.

Objective: Analysis of internal and external migration patterns in the parish of Torslev c.1870-1920. The study is based partly on existing official statistics,and partly on lists and demographic statistics produced from nominal data in original census returns, parish registers, military conscription rolls, police registers, etc. *Method:* The interplay of the demographic background variables and empirical convergence towards stability is studied by population projection methods with computer techniques developed at the Institute of Statistics. Moreover, the social background of the individual migrant and his migration history is considered as far as reconstruction will allow. Finally, economic and

cultural conditions are scrutinized along with innovations in press, communication, economic production, social relief, etc., in order to reveal motives behind the Scandinavian migration to United States prior to c.1920. In association with the micro-study a data set consisting of overseas migrant records from all Denmark c.1870-1899 will be further analyzed statistically.

Type of computer: IBM 370/165. *Size of storage:* 2048k. *Language:* FORTRAN IV, PL/I. *Disks or tapes:* 5 3420 tapes, 5 3330 disks. Program is available on completion of project.

Status: Completed. Doctoral dissertation published March 1972.

References: "Danish Overseas Emigration 1868-1914. Trends and Problems," *Scandinavian Economic History Review* 14 (1966), 158-178; *Flugten til Amerika eller Drivkraefterne i Masseudvandringen fra Danmark 1868-1914* (Aarhus: Universitetsforlaget, 1972); an American version is under preparation, to be published by Seminar Press, New York); Hans Oluf Hansen, "Computer Methods for Production of Socio-Demographic Statistics," *Cuadernos de Historia Economica de Cataluna* 7 (1972), 319-333.

S140. An Investigation of the Migrations in Denmark in the Nineteenth Century

Principal investigator: Anna Thestrup, Archivist, Rigsarkivet, Rigsdagsgarden 9, 1218 Copenhagen K, Denmark.

Objective: To show how the population in Denmark migrated within the country in the nineteenth century, not only between country and city in different periods but also between rural parishes. The age, sex, profession, civil status, etc., of the migrant are mentioned. *Method:* The primary material includes census, church registers of mobility, registers of marriages, births and deaths. [*CHum,* November 1970]

S141. An Analysis of the Census of Norway of 1801

Principal investigators: Knut Mykland, Professor; Dagfinn Mannsaker, National Archivist of Norway, Historical Institute, U. of Bergen, Bergen, Norway. *Associates:* Jan Oldervoll, Georg Michael Gillow.

Objective: 1) To produce a new statistical handbook based on the census, 2) to make the census computer-available for researchers, 3) to be able to make printed lists of parts of the census. *Method:* 1) No coding; we are punching the whole census directly from the handwritten protocols, by means of a papertape-punching device; 2) semiautomatic categorization of the material.

Type of computer: IBM 360/50. *Size of storage:* 384k. *Language:* PL/I, GIS. *Disks or tapes:* 8 2314 disks. [*CHum,* November 1970]

S142. Maschinelle Analyse Altägyptischer Texte

Principal investigators: F. Schulte-Tigges, Leiter des Instituts f. Nichtnumerik; Rolf Gundlach, Doktor, Leiter der Gruppe Historische Wissenschaften, Institut f. Nichtnumerik, Ges. f. Mathematik und Datenverarbeitung mbH, Birlinghoven, Elisabethenstrasse 44 1/2, D-61 Darmstadt, W. Germany. *Associate:* C. Doelter.

Objective: Herstellung von lexikalisch-grammatischen Listen durch maschinelle analyse Altägyptischer texte auf der Grundlage einer graphematisch-morphematischen Textbeschreibung. *Method:* Die ägyptischen Texte werden beschrieben durch Angabe der Grapheme, einer Textkritik, eines Morphemindexes, und von Strukturkennzeichnungen. Ausgehend von einer graphematisch-morphematischen Analyse wird die phonematische Beschreibung der Morpheme eruiert und eine grammatische Formenbestimmung durchgeführt.

Type of computer: Telefunken TR440. *Size of storage:* 256k. *Language:* FORTRAN IV, TAS. *Disks or tapes:* WSP 414. Program is available.

Status: Completed 1972.

References: R. Gundlach and W.Schenkel, "M.A.A.T., Ein System zur lexikalischen und grammatischen Erschliessung altägyptischer Texte mit Hilfe einer elektronischen Datenverarbeitungsanlage," *Chronique d'Egypte* 42 (1967), 41-64; W. Schenkel, "Der Computer als Hilfsmittel für die lexikalische und grammatische Beschreibung des Altägyptischen. Möglichkeiten und Grenzen," *Zeitschrift der Deutschen Morgenländischen Gesellschaft,* Supplementa I, pp. 97-105; W. Schenkel, "Maschinelle Analyze altägyptischer Texte, Aufbau und Grundlagen des Systems M.A.A.T. und seine lexikographische Anwendung," *Schriftenreihe des Deutschen Rechenzentrums,* Heft S-7 (Darmstadt, 1969); W. Schenkel, "Texterschliessung mit Hilfe des Systems M.A.A.T., Uberblick über die verfügbaren Materialien," *Akten der Arbeitsgemeinschaft Dokumentation in den historischen Wissenschaften* 1, *Tagung Dokumentation ägyptischer Altertümer* 26 (July 1969) in Darmstadt 1970; R. Gundlach and W. Schenkel, *Lexicalisch-grammatische Liste zu Spruch 335a der alt ägyptischen Sargtexte LL/CT. 335A,* 2 vols., Schriftenreihe des Deutschen Rechenzentrums S-8/1, S-8/2, Darmstadt, 1970.

S143. Selektive Auswertung von Komponentendeskriptionen in den philologisch-historischen Wissenschaften

Principal investigators: F. Schulte-Tigges, Leiter der Abteilung Nichtnumerik; R. Gundlach, Doktor, Leiter der Gruppe Historische Wissenschaften, Deutsches Rechenzentrum, D-61 Darmstadt, Rheinstrasse 75, Bundesrepublik Deutschland. *Associate:* M. Biallas.

Scope: Informationen über Sachverhalte aus den Bereichen der Altertumswissenschaften und der Geschichte, abgefasst nach der Methode der Komponentendeskription, werden durch Untersuchung der Komponenteninhalte oder der Codes sowie durch konditionierte Befragung von Komponentenkombinationen ausgewertet und stehen so besonders bei Spezialrecherchen zur Verfügung. Das zu diesem Zweck in der Entwicklung begriffene System SELKOM (*Sel*ektive Auswertung von *Kom*ponentendeskriptionen) wird an das System HISDOC/KD (*His*torische *Dok*umentation/ *K*omponenten*d*eskription angeschlossen). *Method:* Zu den im Informationskatalog des Systems HISDOC/KD enthaltenen Informationen werden maschinell Spezialglossare zu den Komponenteninhalten Angefertigt und durch spezielle Retrieval-Prozeduren ausgewertet.

Type of computer: 1) IBM 7094/I, 2) IBM 1401. *Size of storage:* 1) 32k, 2) 8k. *Language:* FORTRAN IV, MAP. *Disks or tapes:* 1) IBM 1301 disk and 4 IBM 729 IV tapes. [*CHum,* November 1970]

S144. Automatische Herstellung von Korrelationen zwischen verschiedenen chronologischen Systemen

Principal investigator: F. Schulte-Tigges, Leiter der Abteilung Nichtnumerik; R. Gundlach, Doktor, Leiter der Gruppe Historische Wissenschaften, Deutsches Rechenzentrum, D-61 Darmstadt, Rheinstrasse 75, Bundesrepublik Deutschland. *Associate:* G. Stoll.

Scope: Umrechnung von chronologischen Angaben, die auf verschiedenen historischen Kalendersystemen fussen, in jede mögliche Richtung zu Informationszwecken und maschinelle Ermittlung von Konsequenzen, die eine Anderung eines chronologischen Systems zur Folge hat. *Method:* Die chronologischen Angaben werden durch die automatische Berücksichtigung von Bezugspunkten unter Einbeziehung der historisch zu erklärenden Ungenauigkeiten vergleichbar gemacht.

Type of computer: 1) IBM 7094/I, 2) IBM 1401. Unstellung auf Telefunken TR440 ist vorgesehen. *Size of storage:* 1) 32k, 2) 8k. *Language:* 1) FORTRAN IV, MAP. *Disks or tapes:* 1) 6 IBM 729 IV tapes, 1 IBM 1301 disk.

S145. Automatische Datierung archäologischer Objekte

Principal investigators: F. Schulte-Tigges, Leiter der Abteilung Nichtnumerik; R. Gundlach, Doktor, Leiter der Gruppe Historische Wissenschaften, Deutsches Rechenzentrum, D-61 Darmstadt, Rheinstrasse 75, Bundesrepublik Deutschland. *Associate:* A. Schwab-Schlott.

Scope: Archäologische Objekte, die nach der Methode der Komponentendeskription beschrieben sind, werden maschinell an Hand einer Datierungsliste, die für alle festgelegten Merkmale Datierungen enthält, untersucht und durch Vergleich der Datierungsangaben datiert. Eine Unvereinbarkeit der Angaben wird diagnostiziert. Zunächst wird das Verfahren angewendet bei der "Dokumentation ägyptischer Altertümer." *Method:* Die Datierungen der Merkmale werden maschinell in einen Anfangs- und einen Endcode überführt. Der Vergleich der Datierungsangaben wird rechnerisch durchgeführt.

Type of computer: 1) IBM 7094/I, 2) IBM 1401. Unstellung auf Telefunken TR440 ist vorgesehen. *Size of storage:* 1) 32k, 2) 8k. *Language:* 1) FORTRAN IV, MAP. *Disks or tapes:* 1) 3 IBM 729 IV tapes, IBM 1301 disks.

References: A. Schwab-Schlott, "Problem der Definition und Anwendung von Datierungsmerkmalen," *Akten der Arbeitsgemeinschaft Dokumentation in den historischen Wissenschaften* I (Darmstadt, 1970).

S146. Untersuchungen zur maschinengerechten Erfassung und automatischen Auswertung von Grundrissen in der Archäologie

Principal investigators: F. Schulte-Tigges, Leiter der Abteilung Nichtnumerik; R. Gundlach, Doktor, Leiter der Gruppe Historische Wissenschaften, Deutsches Rechenzentrum, D-61 Darmstadt, Rheinstrasse 75, Bundesrepublik Deutschland. *Associate:* G. Stoll.

Scope: Pläne von prähistorischen, antiken und mittelalterlichen Siedlungen und Bauten sollen beschrieben und der gezielten sowie zusammenfassenden historischen Auswertung durch automatischen Nachweis und Vergleich von Siedlungsformen, Gebäudetypen und Gebäudeelementen zugänglich gemacht werden. Ausgangsmaterial sind zunächst Grundrisse altägyptischer Tempel. *Method:* Die Strukturbestimmung des Materials wird mit Quellenbeschreibungen nach der Komponentendeskription verbunden. Maschinelle Verarbeitung und Auswertung erfolgt durch ein spezielles Teilsystem des Systems ARDOC (*Ar*chäologische *Dok*umentation).

Type of computer: 1) IBM 7094/I, 2) IBM 1401. Unstelling auf Telefunken TR440 ist vorgesehen. *Size of storage:* 1) 32k, 2) 8k. *Language:* FORTRAN IV, MAP. *Disks or tapes:* 1) 4 IBM 729 IV tapes, 1 IBM 1301 disk.

S147. Automatische Klassifizierung von Gefassformen in der Archäologie

Principal investigators: F. Schulte-Tigges, Leiter der Abteilung Nichtnumerik; R. Gundlach, Doktor, Leiter der Gruppe Historische Wissenschaften, Deutsches Rechenzentrum, D-61 Darmstadt, Rheinstrasse 75, Bundesrepublik Deutschland. *Associate:* M. Biallas.

Scope: Die Formen von Gefässen und Gefässbruchstücken prähistorischer und antiker Kulturen werden maschinengerecht beschrieben, automatisch analysiert und verglichen. Sie sollen damit bei Fragen nach Gefässen bestimmter Form greifbar sein und stehen dann einer maschinellen Klassifizierung zur Verfügung. *Method:* Die Formen werden nach der Kettenkodierungsmethode beschrieben (Elementendeskription), nach bestimmten formsyntaktischen Merkmalen analysiert und unter Berücksichtigung formkritischer Gesichtspunkte verglichen und klassifiziert (Teilsystem ACFORM [*A*utomatische *C*lassifizierung von *Form*en] des Systems ARDOC [*AR*chäologische *Dok*umentation]).

Type of computer: 1) IBM 7094/I, 2) IBM 1401, 3) Telefunken TR440. *Size of storage:* 1) 32k, 2) 8k, 3) 256k. *Language:* 1) FORTRAN IV,MAP, 3) FORTRAN IV, TAS. *Disks or tapes:* 1) 4 IBM 719 IV tapes. *Special equipment:* IBM 1627 plotter, 3) WSP 414 plotter. Program is available.

References: M. Biallas, *ARDOC: Dokumentation and Analyse von Gefässformen* (Programm-Veröffentlichung des Deutschen Rechenzentrums), Darmstadt, 1971; ACFORM-Automatische Klassifikation Gefässformen (Programm-Veröffentlichung des Deutschen Rechenzentrums), Darmstadt, 1971; Rolf Gundlach, "Zur maschinellen Erschliessung historischer Museumsbestande," *Museumskunde* 3 (1968), 140-143.

S148. A Computer System for Storage and Retrieval of Nevada Archaeological Site Records

Principal investigators: Mary Rusco, Acting Director; John E. Riehl, Research Assistant, Nevada Archaeological Survey, U. of Nevada, Reno, NV 89507.

Objective: To improve accessibility to data on Nevada archaeological sites, previously stored only in over 1600 separate file folders. The program will enable the researcher to retrieve records of sites which exhibit any one or any combination of over 50 attributes. *Method:* Data from the Survey Archives were recorded on simple precoded forms for punching on IBM cards. A computer program was written to permit rapid retrieval of all or any part of the data stored on cards. In addition, the variability of the program allows it to be used on a time-sharing system as well as on the standard system.

Type of computer: XDS Sigma 7. *Size of storage:* 32k. *Language:* FORTRAN IV. Program is available. *Special equipment:* Time-sharing capability.

References: Nevada Archeological Survey, Research Paper No. 111 (Reno: University of Nevada—in preparation). [*CHum,* November 1970]

S149. Communications, Power and Political Change in an Egyptian Community

Principal investigator: Iliya F. Harik, Associate Professor, Dept. of Political Science, Indiana U., Bloomington, IN 47401.

Objective: 1) Determine changes in the structure of influence and leadership, 2) determine the reception and effect of mass media on political awareness, 3) compare the effects of mass media with those of opinion leaders, 4) determine how successful the national government has been in inculcating the community with the regime's ideology, 5) determine how decisions are made and by whom. *Method:* A survey schedule was administered to 1) a random sample of ten percent of the adult males (ages 18 to 60) and to 2) a selected group of 44 leaders in the community. The second method has been participant observation and documentary.

Type of computer: CDC 3600. *Size of storage:* 35k. *Language:* FORTRAN IV. [*CHum,* November 1970]

S150. Comparative Analysis of Institutional Differentiation in Kentucky and Peru

Principal investigator: Allen C. Turner, Graduate Research Assistant, Dept. of Anthropology, U. of Kentucky, Lexington, KY 40505.

Objective: To amplify the understanding of the processes of urbanization, a comparative analysis of institutional differentiation in 200 Kentuckian and Peruvian settlements covering a broad range of population levels is undertaken. Three major objectives are addressed in this analysis: 1) to define the relationship between population and the sequence of institutional differentiation, 2) to define the relationship existing between sectors of institutions, and 3) to compare the institutional configurations of the two areas studied. Through the methods of correlation and regression analysis it is

shown that population alone is an insufficient predictor of the level of institutional differentiation. The analysis shows that the preexisting institutional base is important in determining the course of further differentiation. The correlations between the various institutional sectors provide a means for comparing models of institutional articulation in the areas studied.

*Type of computer:*IBM 360/65. Program is available. [*CHum,* November 1970]

S151. The Population of the Bishopric of Caracas, 1760-1820

Principal investigators: John V. Lombardi, Associate Professor, Dept. of History, Indiana U., Ballantine Hall 742, Bloomington, IN 47401; Trent M. Brady, Assistant Professor, Dept. of History, Toronto U., Toronto, 5, ON, Canada.

Objective: Process and analyze the annual parochial census returns from the Bishopric of Caracas, 1760-1820.

Type of computer: CDC 6600. *Language:* FORTRAN (CDC), SPSS.

S152. Institute of Colonial Studies Data Base

Principal investigators: Jackson T. Main, Professor; Keith Kavenaugh, Lecturer, Dept. of History, SUNY, Stony Brook, NY 11790.

Objective: To construct a computerized data base of bibliographic information about the microfilm store of Colonial documents being collected. From this store permuted indices, union type catalogs, frequency dictionaries, and selected indices will be constructed. *Method:* The GRIPHOS 06 (*G*eneral *R*etrieval and *I*nformation *P*rocessor for *H*umanities-*O*riented *S*tudies) programs are used to store, edit, and organize the input data (see G13).

Type of computer: IBM 360/65. *Size of storage:* 512k. *Language:* PL/I F. *Disks or tapes:* 1 data cell, 1 2311 disk; 3 60 KC tapes. [*CHum,* November 1970]

S153. Multivariate Analysis of the Hybrid Origin of the Arikara Indians of the Great Plains

Principal investigator: Ralph W. Alexander, Jr., Teaching Assistant, Dept. of Anthropology, Indiana U., Bloomington, IN 47401.

Objective: Elucidate the various contributions of several parental groups to the physical types present in the Arikara Indians. *Method:* Multivariate discriminant analysis of the metric and morphological characteristics of the hybrid Arikara group and the several parental groups proposed as ancestral, giving relative contributions and biological population "distance."

Type of computer: 1) CDC 3600, 2) CDC 3400. *Size of storage:* 1) and 2) 32k. *Language:* FORTRAN IV + 9 others. *Disks or tapes:* 1 500,000-word disk, 6 tapes. *Special equipment:* 1 high-speed drum. Program is available. [*CHum,* November 1970]

S154. Cluster and Principal Component Analysis of Ceramics

Principal investigator: Norman E. Wagner, Director, Graduate Studies and University Research, Waterloo Lutheran U., Waterloo, ON, Canada.

Objective: To test various kinds of cluster analyses on data gathered by scholars over the years. The purpose is to establish whether a single analysis can cope with data coming from numerous sites over a long time period. *Method:* Basing our model on Sokal and Sneath, *Principles of Numerical Taxonomy,* we calculated coefficients of correlation and taxonomic distances. A dendrogram based on the weighted pair group method was produced by means of a newly devised computer program. A principal component analysis program was developed which produced results more useful than the above, but consistent with them.

Type of computer: Honeywell 316. *Language:* FORTRAN IV. Program is available.

References: Papers presented to the Canadian Archaeological Association, currently available in mimeographed form.

S155. In Defense of the Western Sea: The Spanish Crown and the Defense of the Indies, 1535-1585

Principal investigator: Paul E. Hoffman, Professor, Dept. of History, Louisiana State U., Baton Rouge, LA 70803.

Objective: To determine what the Spanish crown spent for defense, what this bought in terms of types of military goods and services, where these forces were used, and when. Quantititative data covering this aspect of the story are supplemented by a narrative based on traditional (non-quantitative) sources. It is hoped this double study will permit the writing of a general history of the Spanish side of the hostilities in the Caribbean and the Eastern Atlantic, 1535-1585. *Method:* Quantitative: extracting and coding entries from the royal treasury accounts; these are then converted to a uniform monetary base (maravedi) and sorted by the computer, which compiles time series by type and geographic area. Non-quantitative: standard historical research techniques.

Type of computer: IBM 360/70. *Size of storage:* 5-10k. *Language:* FORTRAN IV. *Disks or tapes:* 1 I/O tape. Program is available.

References: "Royal Spending for the Defense of the Indies: 1535-1585," *Historical Methods Newsletter* 3, 3 (1970), 31; "The Program Budget as a Tool of Historical Analysis," ibid. 4 (1970), 14-18; "The Computer and the Colonial Treasury Accounts. A Proposal for a Methodology," *Hispanic American Historical Review* 50 (1970).

S156. The Ministry of Internal Affairs in the Russian Empire, 1905-1917

Principal investigator: D.K. Rowney, Associate Professor, Dept. of History, Bowling Green State U., Bowling Green, OH 43402.

Objective: To write a general history of the Ministry in the light of organization and system theory. *Method:* Analysis of career, demographic, and budgetary data.

Type of computer: 1) IBM 1620, 2) IBM 360/40, 3) IBM 360/50. *Size of storage:* 2) 512k, 3) 1112k. *Language:* 1) FORTRAN II, 2) and 3) FORTRAN IV G. *Disks or tapes:* 3) 1 9-track tape.

References: "Study of the Ministry of Internal Affairs in the Light of Organizational Theory," *The Behavioral Revolution and Communist Studies,* ed. Roger Kanet (New York: The Free Press, 1970). [*CHum,* May 1971]

S157. A Comparative Analysis of Social Change in Yorkshire and Massachusetts, 1690-1841

Principal investigators: Donald E. Ginter, Associate Professor; Robert E. Wall, Jr., Associate Professor; Dept. of History, Sir George Williams U., Montreal, PQ, Canada.

Scope: The project has selected two highly comparable societies during their period of critical transition from traditional to modern social systems. While initially focusing principally on the emergence of modern political parties as a dependent variable, the project is designed to gather the widest possible variety of social data and to test systematically hypotheses concerning the functional relationships between various forms of social change. *Method:* The Yorkshire phase of the project is presently focusing on the enfranchised level of the population during the period 1741-1807, utilizing pollbooks, participation in local and national political clubs and in the Wyvill association, reform petitions, land tax returns, and jurors lists. The Massachusetts phase is beginning with the census of 1790 and with political and tax records. Eventually the project will attempt demographic reconstitution throughout the period, pushing to the lowest strata of the population and utilizing such records as parish registers and probate files. The maximal number of available social indicators will be built onto the record of every individual identified. A detailed delineation of the changing social ecology of the two societies will provide a context for analysis of individual and group behavior patterns.

S158. Privacy and the Computer

Principal investigators: John M. Carroll, Assistant Professor, Dept. of Computer Science; J. Ivan Williams, Assistant Professor, Dept. of Sociology and Community Medicine; U. of Western Ontario, London, ON, Canada.

Scope: A study of computer data banks and their effect on individual privacy of information of subjects, including methods of collection, storage, manipulation, interchange, release, and dissemination of information. Categorization of types of data banks and types of information handled. Investigation of methods to ensure the security of information in real-time on-line systems. Special emphasis on privacy of student records in Canadian Universities. *Method:* Site visits, depth interviews, mail questionnaires, analysis of systems from detailed documentation and forms, computer simulation of the privacy environment, computer development of security techniques, and simulated attack on an actual system.

Type of computer: 1) PDP-10/50, 2) CDC CYBER 70/14, 3) PDP-8I. *Size of storage:* 1) 128k, 2) 64k, 3) 8k. *Language:* 1) FORTRAN, TECO, GPSS, 2) FORTRAN, SIMSCRIPT, 3) Assembler. *Disks or tapes:* 1) 3 1/2 M, 2) 1 M, 3) 100k disks; 1) 4 DEC tapes, 4MT, 2) 7 MT, 3) 2 DEC tapes.

References: "Bugging the 'Big Brains,' " *Executive* (December 1969); "The IE's Guide to Electronic Bugs," *Industrial Engineering* (May 1969); "Privacy and the Computer, and Interdisciplinary Enquiry," *Proceedings of Conference on Interdisciplinary Research in Computer Science,* University of Manitoba, 8-10 June 1970; "The Privacy Environment of Canadian Resource-Sharing Systems," *Infor Journal* (March 1971); "Privacy and Computers," Department of Communications, Department of Justice, Information Canada, Ottawa, 1972 (Section 2); Personal Records: Procedure, Practices and Problems, Department of Communications, Department of Justice, Ottawa, 1973; The State of Medical

Records in Canada, International Institute of Administrative Sciences, Working Group on Informatics and Administration, Nice, April 1973.

S159. Methods of Identifying Constituent Phrases in Statute Law

Principal investigator: Kenneth George Langland, Lecturer, Dept. of Computer Science, U. of Western Ontario, London, ON, Canada.

Objective: 1) To discover methods of extracting constituent phrases from data base of statute law. 2) To investigate ways of utilizing word search strategies and professionally created indexes with these constituent phrases. *Method:* Experimental design and evaluation of trial systems.

Type of computer: 1) IBM 360/65 OS, 2) IBM 7040, 3) PDP-10. *Size of storage:* 1) 512k, 2) 32k words, 3) 64k words. *Language:* 1) COBOL F, 2) COBOL, 3) ANSI COBOL. *Disks or tapes:* 1) 8, 2) 8, 3) 5 tapes; 1) 12, 2) 1, 3) 10 disks. Programs are patented.

References: Canadian Bar Association Journal (January 1971). [*CHum,* May 1971]

S160. Ancient Sumerian Cuneiform Economic Tablets

Principal investigators: Nels Forde, Associate Professor; Gary A. Thomson, Doctoral Candidate; Dept. of History, U. of Nebraska, Lincoln, NB 68508. *Associates:* Howard J. Cyr, Marie Craft.

Scope: Analysis of the approximately 30,000 Sumerian cuneiform economic tablets of the Ur III Dynasty (2050-2000 B.C.). Statistical comparisons will produce a new historical reconstruction of the Sumerian middle or mercantile class, the world's first. *Method:* The content of each tablet is transliterated in Sumerian cuneiform on punch cards. The program then uses the same keywords and context that would be used in translation to identify types of information and produce classified records which are then sorted on pertinent fields and subject to further analysis.

Type of computer: IBM 2065. *Size of storage:* 1000k. *Language:* COBOL F. *Disks or tapes:* 3 2314 disks, 7 tapes. Program is available. [*CHum,* May 1971]

S161. The Energies of Henry Adams

Principal investigator: Charles H. Holbrow, Associate Professor, Dept. of Physics and Astronomy, Colgate U., Hamilton, NY 13346.

Objective: To study Henry Adams' theory of history as expressed in his writings, especially in "A Letter to American Teachers of History" and in the last chapters of *The Education of Henry Adams. Method:* Selective concordance of substantial portions of these writings to discern the different meanings and uses of certain important words.

Type of computer: 1) IBM 360/91, 2) UBN 360/65, 3) IBM 1130. *Size of storage:* 1) 300k, 2) 200k, 3) 32k. *Language:* 1) and 2) SNOBOL4, 3) FORTRAN. *Disks or tapes:* 2314 disk, 2 tapes. Program is available. [*CHum,* May 1971]

S162. Cluster Analysis and Cultural Stratigraphy in Archaeological Sites

Principal investigator: Jack D. Nance, Assistant Professor, Dept. of Sociology and Anthropology, Murray State U., Murray, KY 42071.

Objective: To assess the value of cluster analysis (numerical taxonomy) in defining taxa of artifacts which have temporal significance within a stratified achaeological context. *Method:* 1) Average-link cluster analyses of individual artifacts (projectile points). 2) Statistics-aided interpretation of phenograms. 3) Average link cluster analyses of stratigraphic distribution of taxa of artifacts. 4) Statistics-aided discovery of taxa of artifacts which show significantly different stratigraphic position.

Type of computer: IBM 360/40. *Size of storage:* 100k. *Language:* FORTRAN IV G. *Disks or tapes:* 3 2314 disks. Program is available.

References: "Classification and Analysis of Artifacts" (Unpub. Ph.D. dissertation, University of Calgary, 1972).

S163. The Use of Sampling in Archaeological Survey

Principal investigator: James W. Mueller, Assistant Professor, Dept. of Anthropology, Bridgewater State College, Bridgewater, MA 02401. *Associates:* Lawrence Manire, Philip Strongin.

Objective: 1) To determine which sampling designs are most economical and representative of an empirical population, and 2) to discuss the relative advantages of each sampling scheme in various archaeological situations, and 3) to discuss the grave implications of the similarity between survey-sampling and cluster sampling. *Method:* A comparison of 326 simulated, sampling-based surveys to a known population of 488 sites in north central Arizona. The experiment tests four major sampling factors—sampling scheme, fraction, and unit, as well as population—and archaeological variables by the use of *chi*-square, Kish's concept of economy, the Spearman correlation coefficient, and simple percentages. Six variables—environmental locale, site type, unpainted and painted ceramics, age, and lithic function—are used to describe the samples and the population. For chi-square, frequencies of occurrence of these variables constitute the observed, while the population frequencies constitute the expected. Economy is based on actual survey costs divided by the sample variance in site density. The simple random and all variants of the systematic and stratified schemes were generated by computer program, while the cluster scheme and others were manually selected.

Type of computer: CDC 6400. *Language:* FORTRAN IV. *Disks or tapes:* 4 6603 disks, 4 607 tapes. Program is available. Write to the University of Arizona Computer Center.

Status: Completed August 1972.

S164. Studies of Regressive Imagery

Principal investigator: Colin Martindale, Assistant Professor, Dept. of Psychology, U. of Maine, Orono, ME 04473.

Scope: Ethnographic and child-rearing correlates of regressive imagery in a sample of folktales from 45 primitive societies. Validation study of Regressive Imagery Dictionary comparing schizophrenic vs. non-schizophrenic texts. Study of regressive imagery as a function of level of arousal in texts elicited from normal subjects. *Method:* Analysis of texts with the Regressive Imagery Dictionary, a coding scheme for use with The General Inquirer. Correlation and cross-tabulation of resultant indices with non-text variables.

Type of computer: IBM 370. Program is available.

References: The Romantic Progression: On the Psychology of Literary Change (in preparation).

S165. Quantitative Analyses of Eighteenth-Century Medical Historical Material

Principal investigator: Oivind Larsen, M.D., Assistant Professor, Dept. of Anatomy, Universitetet I Oslo, Karl Johans Gt. 47, Oslo 1, Norway.

Scope: Analyses of medical texts from the eighteenth century. Textbooks and handwritten medical records from medical institutions are compared to see to what extent the complicated medical treatment recommended really was practised. *Method:* Selected medical textbooks, and selected medical records from the period 1775-1800, all of Scandinavian origin, are transferred to magnetic tape. Searching for keywords, printing of occurrences in context, and comparisons of the context are the basic principles of the investigation.

Type of computer: CDC 3300. *Language:* FORTRAN IV. *Disks or tapes:* Part of 1 disk. Program is available (see G57).

S166. Factor Analysis and Archaeology

Principal investigator: Andrew L. Christenson, Archaeological Survey, U. of California, Los Angeles, CA 90024. *Associate:* Dwight Read.

Objective: To investigate the applicability of the various factor-analytic models for the study of prehistoric cultural systems. We are particularly interested in the use of factor analysis for describing and explaining formal, temporal, and spatial variability in archaeological materials. *Method:* The BMD series factor analysis program (BMDX72) is being used. We are testing and comparing the classical factor and principal component solutions on a wide variety of archaeological data. We are also reexamining many of the previous archaeological applications of factor analysis.

Type of computer: IBM S 360/91. *Size of storage:* 250k. *Language:* FORTRAN IV. [*CHum*, May 1972]

S167. Roll-Call Voting in German Parliaments in the Revolution of 1848-1849

Principal investigator: Donald J. Mattheisen, Assistant Professor, Dept. of History, Lowell State College, Lowell, MA 01854.

Objective: By using the Guttman scale analysis of roll-call votes in the German and Prussian constituent assemblies of 1848-1849, to see if our sketchy information on parliamentary parties can be enhanced by comparing the behavior of party members with that of the independents. Other statistical devices, such as the Rice Index of Cohesion, will be used where appropriate questions suggest themselves.

Type of computer: IBM 360.

References: "A Guttman Scale Analysis of the Prussian Parliament of 1848," *Research Reports in the Social Sciences* (Notre Dame, IN) 3 (1970), 1-8; "Voters and Parliaments in the German Revolution of 1848: An Analysis of the Prussian Constituent Assembly," *Central European History* 5 (March 1972), 1-22.

S168. Russian Imperial Bureaucracy, 1762-1881; French Old Regime Bureaucrats, 1661-1790

Principal investigator: John A. Armstrong, Professor, Dept. of Political Science, U. of Wisconsin, Madison, WI 53706. *Associates:* Brian Silver, James O'Connor.

Scope: Establishment of biographic data banks on Russian Imperial elite administrators, especially governors, for last century of old regime; a similar bank for governors *(intendants de province)* only for last century of French old regime. To compare asynchronously the phenomena of bureaucratic regression to patrimonial features during the final period of a semi-traditional empire and to provide more incidental bases for comparing these periods with post-industrial periods. *Method:* Search of major biographic directories, especially *Russkii Biograficheskii Slovar,* codification of data with special emphasis on periods of incumbency and age at which these incumbencies were held so as to obtain patterns of turnover, career expectancies, and duration of position occupancy. (Data abstract available from Social Science Data and Program Library Service, University of Wisconsin. Either or both of the above titles should be requested specifically by title.)

Type of computer: UNIVAC 1108. Program is available.

References: "Old-Regime Governors: Bureaucratic and Patrimonial Attributes," *Comparative Studies in Society and History* (January 1972); "Tsarist and Soviet Elite Administrators," *Slavic Review* (Spring 1972). [*CHum,* May 1972]

S169. QUIC/LAW

Principal investigator: Hugh Lawford, Professor, Law Faculty, Queen's U., 140 Beverley Street, Kingston, ON, Canada. *Associates:* Richard von Briesen, Bruce Walter.

Objective: Development of computerized systems of retrieval, text-editing, and printing of Canadian laws, international treaties, and other government documentation.

Type of computer: 1) IBM 360/65, 2) IBM 360/50. *Size of storage:* 1) 768 + LCS (1000k), 2) 512 + LCS (1000k). *Language:* Assembler, COBOL, PL/I. *Disks or tapes:* 1) 32 2314 disks, 8 tapes, 2) 8 2314 disks, 4 tapes. *Special equipment:* CRT terminals. Some programs available for exchange.

S170. "Hot" and "Cool" Catechesis: Technological Determinism Over 400 Years

Principal investigator: Ralph Dengler, S.J., Assistant Professor, Dept. of Communications, Fordham U., Fordham Road, Bronx, NY 10458. *Associate:* Don Goldhamer.

Scope: Content analysis of sixteenth-century and twentieth-century catechisms to test theories of technological determinism regarding the effect of printing and electronic media upon language as proposed by Marshall McLuhan, Harold Innis, and Walter Ong, S.J. Four hypotheses were selected from their writings. Results of the analysis were correlated with the hypotheses. Other documents now being processed. *Method:* The General Inquirer System was followed: 80,000 words selected by random sampling of sentences; words were "tagged" into 96 categories. The Harvard Psychosociological Dictionary III and a modification were used for two runs; tests of validity, consistency, and significance; clusters of relevant categories formed; results graphed; contingency analysis by retrieval; interpretation. Four additional documents processed in a follow-up study supported the original results.

Type of computer: 1) IBM 7094, 2) IBM 360/65. *Size of storage:* 2) 100k. *Language:* 1) FAP, 2) FORTRAN IV. *Disks or tapes:* 1) and 2) 1 tape, 2314 disk (temporary). Program is available.

Status: Completed February 1973.

References: Ralph Dengler, "A General Inquirer Analysis of Sixteenth-Century and Contemporary Catechisms," *Computers and the Humanities* 8, 1 (1974).

S171. Folklore

Principal investigator: P. Richard, Centre de Mathématiques Appliquées et de Calcul, 54, boulevard Raspail, Paris 6e, France.

Scope: Edition de résumés des contes du catalogue Delarue-Ténèze pour les contes français. *Method:* A partir de la liste des éléments d'un conte-type et de l'extrait d'une variante définie par Delarue ou M. Ténèze, le calculateur imprime le texte où chaque symbole a été remplacé par l'élément qui lui correspond. Les fragments entre parentèses sont adjoints aux éléments. Les ajouts non codés sont imprimés avec un format qui permet de les reconnaître. Les textes de variantes

codées peuvent être frappées sans presque aucune préparation. Seuls les ajouts sont repérés par l'adjonction d'un signe spécial. Un programme annexe permet de vérifier la correction de l'écriture et note les principales erreurs de format ou typographiques.

Type of computer: IBM 370/165 and IBM 370/135. *Size of storage:* 220k and 80k. *Language:* SNOBOL4/SPITBOL. Program is available.

S172. Analysis of the Florentine Catasto

Principal investigators: David Herlihy, Harvard U., Cambridge, MA 02138; Christiane Klapisch, Centre de Recherches Historiques, Ecole des Hautes Etudes, VIe section, Paris, France.

Scope: Analysis of the Catasto archives, a combined census and property survey first redacted in 1427 and repeated several times in the course of the fifteenth century, to study the family in Florence and its surroundings; to correlate characteristics of the family with such factors as wealth, residence, occupation, and social position. *Method:* The Catasto survey of 1427 is contained on approximately 120,000 card images (an average of two per household) written on seven tapes. Supplementary samples include a 10 percent random sampling of the later Catasti of 1458-59 and 1480, for the city of Florence only.

References: D. Herlihy, "Editing for the Computer: The Florentine Catasto of 1427," *ACLS Newsletter* 22 (1971), 1-7; David Herlihy, "Mapping Households in Medieval Italy," *Catholic Historical Review* 58 (1972), 1-24; David Herlihy, "Marriage at Pistoia in the Fifteenth Century," *Bollettino Storico Pistoiese* 74 (1972), 3-21; Christiane Klapisch, "Household and Family in Tuscany in 1427," *Household and Family in Past Time*, ed. Peter Laslett (Cambridge, 1972), 267-281.

S173. Religious Enthusiasm and Sainthood in the Thirteenth Century

Principal investigators: Michael Goodich and Michael Radow, Columbia U. Computer Center, Columbia U., New York, NY 10027.

Scope: These are computerized data consisting of all the essential information concerning all European Christian saints who flourished c. 1215-1330. *Method:* Collation on social class origins, geographical distribution, birth, death, age of conversion, family attitude, education, religious order, nature of religious conversion experience, marital status, sex, year of conversion of saints.

Type of computer: 1) IBM 360/91, 2) IBM 360/75 (ASP). *Language:* SNOBOL B. *Disks or tapes:* 9-track tapes. [*CHum*, May 1972]

S174. Collective Biography of Tudor Oxford

Principal investigator: J.K. McConica, c.s.b., Professor, Dept. of History, Pontifical Institute of Mediaeval Studies, 59 Queen's Park Crescent, Toronto, ON, Canada. *Associate:* T.K. Rabb.

Objective: A collective biography of Tudor Oxford. *Method:* A pilot project was run on one College. The experience gained from that experiment has suggested a procedure based on three samples: a very limited amount of data about a wide range of graduates (all known graduates), more extensive data about members of a biographical particular group of colleges, and extensive biographical data about all known members of one College. Another year will be required to finish assembling the needed data.

S175. The Bruges Money Market in 1400

Principal investigator: Hyman Sardy, Graduate Center, City U. of New York, New York, NY 10036.

Objective: Using as data business letters in the Datini Archives, Archivio di Stato, Prato, study of the money market on the basis of actual foreign exchange rates. More than 800 observations and 4000 exchange rates *(côtes du change)*: the exchange rates in Bruges on five other banking places (Barcelona, Genoa, Paris, Venice, and London).

References: Results published in *Verhandelingen* 63 of the Koninklijke Vlaamse Academie voor Wetenschappen, Letteren en Schone Kunsten van België (1968).

S176. Rural Economy of the Duchy of Wroclaw in the Later Middle Ages

Principal investigator: Richard C. Hoffmann, Professor, Dept. of History, York U., 043 McLaughlin College, Downsview, ON, Canada.

Scope: Secular course of the rural economy of a central Silesian principality during the period when its primary institutional form was the peasant village under so-called German law. *Method:* Time series of settlement types, landlordship, and land use distribution from archival records of land transfers and tax rolls. Computer used as data file and to construct contingency tables for various sets of characteristics of almost 400 villages at five points in time.

Type of computer: IBM 360. *Size of storage:* 250k. *Language:* MAD. *Special equipment:* Point plotter.

References: "Warfare, Weather, and a Rural Economy: The Duchy of Wroclaw in the Mid-Fifteenth Century," *Viator* 4 (1973).

S177. The Theories and Practices of Representative Government in Colonial Australia

Principal investigator: Dean H. Jaensch, Lecturer, Dept. of Politics, U. of Adelaide, Adelaide, S. Australia. *Associate:* L.G. Veitch.

Scope: An analysis of electoral apportionment and legislative behavior in Australia. A major project on elections and legislators in colonial South Australia is in the final stages. Following this, the intention is to carry on the research into the post-1901 period for Australia as a whole. *Method:* Various techniques in the analysis of apportionment. The analysis of legislative behavior is being carried out by use of three major programs: 1) program MULTBET developed by the CSIRO, 2) a cluster method based on that of Stuart Rice, 3) a factor method recently developed.

Type of computer: CDC 6400. *Size of storage:* 140k. *Language:* FORTRAN IV. *Disks or tapes:* CDC 6603/854 disks. *Special equipment:* CalComp 10" and 30" plotters. Program is available. [*CHum,* May 1972]

S178. A Social History of the Mexican-American Community of Los Angeles, 1850-1900

Principal investigators: Richard Alan Griswold, Instructor, Dept. of Social Science, Los Angeles Trade Technical College, 400 W. Washington, Los Angeles, CA 90008; Richard Weiss, Professor, Dept. of History, UCLA, 405 Hilgard Ave., Los Angeles, CA 90024.

Scope: A study of social changes within the Mexican-American community 1850-1900 with an emphasis on the family, social mobility, and leadership. A study of the effect of urbanization, industrialization, and minority status as affecting these institutions. *Method:* Use of census data, assessment lists, Spanish-language newspapers, city directories, and impressionistic primary sources. Analysis of variables using SPSS and reconstruction of family types using PL/I.

Type of computer: 1) IBM 360 DOS, 2) IBM 360 OS. *Size of storage:* 1) 400k, 2) 400k. *Language:* 1) PL/I, 2) SPSS. *Disks or tapes:* 1) and 2) 2 7-track tapes. Program is available.

References: Historical Methods Newsletter 2 (1969). [*CHum,* November 1972]

S179. Dutch Immigration Patterns to the United States 1845-1877

Principal investigator: Robert P. Swierenga, Professor, Dept. of History, Kent State U., Kent, OH 44242. *Associate:* Harry S. Stout.

Scope: Social, economic, religious, and demographic data on all 18,000 heads of families and single individuals who emigrated from the Netherlands for the period 1845-1877. Will focus on emigrants to the U.S. and trace in U.S. Immigration Passenger Lists 1820-1900 and censuses 1850-1880. *Method:* 1) Coding of data on name, age, sex, number of women and children, occupation, economic status and tax class, religion, destination, year of emigration, original province and city, and reason for emigrating; 2) card punching; 3) sort and merge program; 4) cross-tabulations of pairs of variables; 5) correlation analysis.

Type of computer: Burroughs B5500. *Language:* FORTRAN IV. [*CHum,* November 1972]

S180. A Formal-Functional Analysis of Ceramic Distribution at Kaminaljuyu

Principal investigator: Joseph J. Lischka, Assistant Professor/Survey Archaeologist, Dept. of Social Science, Arkansas Archaeological Survey, Box 3087, UAM, Monticello, AR 71655.

Objective: Test hypotheses concerning the uses and functions of various types of ceramic vessels by investigating the spatial distributions of the vessel types and the relation of the distributions to other cultural materials, specifically architecture. Isolate and identify sets of vessel types representing specific types of activities in which ceramic vessels were used. *Method:* 1) Define set of vessel types based on attributes assumed to be functionally significant, e.g., shape, dimensions, surface finish, decoration, wear, etc. 2) Tabulate frequencies of each vessel type in series of samples representing different archaeological contexts at same time level in same site. 3) Perform factor analysis or key-cluster analysis with vessel types as variables and samples as objects to obtain activity clusters of vessel types.

Type of computer: CDC 6400. Program is available, from Tryon and Bailey Associates in Boulder, CO. Name of program system is BC TRY.

References: Tryon and Bailey, *Cluster Analysis* (description of BC TRY system and statistics involved). [*CHum,* November 1972]

See also G22, G32, G57, G65, G67.

Visual Arts

V1. Line Drawing Manipulation by Computer

Principal investigator: Leslie Mezei, Associate Professor, Dept. of Computer Science, U. of Toronto, Toronto 5, ON, Canada.

Objective: To develop a set of routines to manipulate line drawings, particularly as required for artistic purposes. *Method:* A set of generalized subroutines including various manipulations involving position size, orientation, distortions, smoothing, etc.

Type of computer: 7094 II. *Size of storage:* 32k. *Language:* FORTRAN IV and subroutines. *Special features:* CalComp plotter. Program is available.

Status: 25 routines exist, with more being added and modified. *Future plans:* Development toward a general language of visual information processing. Concurrently, a study of artistic design assisted by computer, and study of formal design principles. [*CHum,* May 1967]

V2. Computer Art

Principal investigator: Charles A. Csuri, Professor of Art and Director, Computer Graphics Research Group, Ohio State U., Columbus, OH 43210. *Associates:* Manfred Knemeyer, Thomas DeFanti, Gerard Moersdorf, Mark Gillenson, Leslie Miller, Robert Schwartz, Wayne Bennett.

Scope: A real-time interactive computer system is used to produce art films. An animation/graphics programming language on a minicomputer has been developed which makes the computer system natural to use and more powerful in terms of itself. Specific computer techniques and algorithms are programmatically combined to make complex interactive programs that are general enough to be applicable to any structure. Structures, in turn, can include any drawing or motion sequence the artist creates. Special analog devices are also used to make the interaction between the artist and the computer a relatively natural one.

Type of computer: PDP-11/45. *Size of storage:* 16k MOS memory + 32k core memory. *Disks or tapes:* System disk and User disk, dual DEC tapes, DECwriter and VT05 alphanumeric display terminal. Vector General Graphics display with 3-D rotation hardware, CRT with z-axis cueing, variable intensity, subroutine stack. CRT input devices: 1 lightpen, 3-D joystick, 10 dials, 16 function switches, SAC 3-D Grafpen (sonic pen). Details of the techniques used and the language were to be published in a technical report to National Science Foundation, April 1973.

V3 [See M8].

V5. Maschinelle Dokumentation prähistorischer Keramik

Principal investigator: Cornelius Ankel, Dr. phil., Institut für Vor- und Frühgeschichte, U. Bonn, 53 Bonn, Liebfrauenweg 7, Germany; Rolf Gundlach, Dr. phil., Deutsches Rechenzentrum, 61-Darmstadt, Rheinstrasse 75, West Germany.

Scope: Gefässe und Scherben prähistorischer Zeitstellung. *Method:* Details durch Dr. Gundlach.

Language: ARDOC.

References: "Maschinelle Dokumentation prähistorischer Keramik," *Umschau in Wissenschaft und Technik* 3 (1967), 90; "Zur Dokumentation und Auswertung vorgeschichtlicher Keramik," *Keramische Zeitschrift* 19, 4 (1967), 221ff; R. Gundlach, "ARDOC: A System of Information Storage and Retrieval in Archaeology" (manuscript, Deutsches Rechenzentrum Darmstadt). [*CHum,* November 1967]

V6. A Hardware-Independent Computer Drawing System: The Quick-Draw Graphics System

Principal investigator: Jeff Raskin, Associate in Music, U. of California, San Diego (mailing address: Box 109, La Jolla, CA 92037).

Scope: A generalized computer graphics software system designed originally to facilitate work in music typography. (However, it is being used at the Pennsylvania State Computation Center for all types of graphics.) *Method:* A list structure is used to build hierarchically structured pictures which can be manipulated in a wide variety of ways. Any geometrical pattern can be used as a picture element. The symbols used in music typography are available separately from the author.

Type of computer: IBM 360/50, IBM 360/67. *Language:* FORTRAN IV G and H. *Special equipment:* CalComp plotter (for printer and 2250 CRT). Program is available.

References: "A Hardware Independent Computer Graphics System" (forthcoming in *SICGRAPH Newsletter).* [*CHum,* May 1968]

V7. Dimensions of Film: An Experiment in the Application of Psychometric Methods to the Study of Film as Art

Principal investigator: Mieczyslaw Choynowski, Director of the Psychometrical Laboratory of the Polish Academy of Sciences, Warsaw PKiN, Poland.

Objective: The development of a quantitative method for the description, classification, and analysis of films from the psychological and aesthetical point of view. *Method:* In the first stage of the project, presently under way, as many feature films as possible are being analyzed with a checklist of 372 terms describing, characterizing, and evaluating them. The collected data will be factor analyzed in order to discover the basic dimensions of a film, whereupon the checklist will be substantially shortened, and in the second stage a number of chosen prominent films, representing major film makers and major schools, will be analyzed.

Type of computer: This project does not demand special programs. Standard programs for different methods of factor analysis and multidimensional scaling can be used with any computer which is able to handle large matrices of data. [*CHum,* May 1968]

V8. Dimensions of Painting: An Experiment in the Applications of Psychometric Methods to the Study of Art

Principal investigator: Mieczyslaw Choynowski, Director of the Psychometrical Laboratory of the Polish Academy of Sciences, Warsaw PKiN, Poland.

Objective: The development of a quantitative method for the description, classification, and analysis of paintings. *Method:* To date three sets of contemporary paintings (Polish and Western originals) were evaluated in a pilot study with three kinds of instruments (semantic differential, ranking in incomplete block designs, and judgments of similarity) by groups of artists and laymen. The analyses, already done and still to be done, comprise factor analyses and multidimensional scaling. Further collection of data with a revised instrument on samples of pictures representing various periods of Western Art is in preparation.

Type of computer: This project does not demand special programs. Standard programs for different methods of factor analysis and multidimensional scaling can be used with any computer which is able to handle large matrices of data.

References: "Dimensions of Painting," *Perceptual and Motor Skills* 25 (1967), 128; "Wymiary malarstwa" (Dimensions of Painting), *Studia Estetyczne* 4 (1967), 199-232. [*CHum,* May 1968]

V9. Attribution and Identification of Suspect Works of Art by Computer-Oriented Pattern Recognition (To Include a Quantitative Analysis of Style)

Principal investigator: Malcolm H. Preston, Professor, Dept. of Fine Arts, Hofstra U., Hempstead, NY 11550.

Scope: To arrive at a system for quantifying an artist's work, with respect to those highly individual characteristics which mark his "handwriting style." These characteristics, including such variables as slope, angles, main lines, proportions, retraces, etc., are being studied in relation to general characteristics to produce a "personal pattern" of an artist's production.

V10a. Museum Computer Network

Principal investigators: David Vance, Registrar, The Museum of Modern Art, 11 West 53 Street, New York, NY 10019; Jack Heller, Professor, Dept. of Computer Science, SUNY; Stony Brook, NY 11790.

Scope: The user tagging system for museum data has been designed and specified to run with the GRIPHOS information retrieval and data dissemination system. Since June 1969 the following organizations have been building a portion of a data base from their accession records: Arkansas Archeological Survey; Florida State Division of Archives, History and Record Management; Museum of the American Indian; Metropolitan Museum of Art; Museum of Modern Art, Museum of Modern Art Film Department; National Gallery of Art. The data bank has been used to produce global reports about the holdings of the various organizations and selected retrieval based on logical questions about the data.

Type of computer: IBM 360/50, IBM 360/65, IBM 370/145 or 370/155. *Language:* PL/I F and BAL.

References: Everett Ellen, "An International Survey of Museum Computer Activity," *Computers and the Humanities* 3 (November 1968), 65; E. Ellen, "Information Systems and the Humanities: A New Renaissance," *Computers and Their Potential Applications to Museums* (New York, 1969), pp. 323-335; Jack Heller, "The Value of a Computerized Data Bank as an Adjunct of a Museum Card Catalogue," *Computers and Their Potential Applications in Museums* (New York, 1969), pp. 307-322; David Vance and Jack Heller, "Structure and Content of a Museum Data Bank," *Computers and the Humanities* 6 (November 1971), 67.

V10b. The Museum of Modern Art Computerized Catalog

Principal investigators: David Vance, Registrar; Teri Varveris, Associate Registrar; The Museum of Modern Art, 11 West 53 Street, New York,NY 10019. *Associate:* Eileen Bowser.

Objective: Transfer of museum collection of about 25,000 descriptions to machine-readable form; storage as depth-indexed, direct-access, open-ended data set (according to the data structure developed by the Museum Computer Network, Inc.); production of computer-generated card files, ledgers, and selective listings—complete. Transfer of Film Collection catalog of about 3800 descriptions, as above—60 percent complete.

Type of computer: IBM 370/145. *Size of storage:* 256k. *Language:* PL/I F level. *Disks or tapes:* 3 2314-type disks, 3 9-track tapes. *Special equipment:* TN chain with accents. Program is available.

References: David Vance and Jack Heller, "Structure and Content of a Museum Data Bank," *Computers and the Humanities* 6, 2 (November 1972).

V11. Information Aesthetics and Picture Generation

Principal investigator: Frieder Nake, Postdoctoral Fellow, Computer Science, U. of Toronto, Toronto 5, ON, Canada.*

Scope: Analysis of some well-known works of art (Mondrian) and generation of new pictures according to prescribed criteria of information aesthetics. Investigations towards some knowledge of "structure." *Method:* Application of some mathematical and aesthetical theories. 2) Development of rather general programs.

Type of computer: IBM 7094 II, IBM 360/65. *Size of storage:* 32k, 390k. *Language:* FORTRAN IV, ALGOL, PL/I, SPARTA. *Special equipment:* CalComp plotter, display unit. [*CHum,* November 1968]

V12. Automation of Accession and Catalog Records

Principal investigator: Registrar's Dept., Metropolitan Museum of Art, Fifth Avenue and 82nd Street, New York, NY 10028.

Objective: To determine physical characteristics of catalog, extent and type of catalog utilization by museum staff, scholars, etc., for purpose of eventual computerization. The study itself was to culminate in tentative systems designs and hardware, personnel and software recommendations, and, presumably, the actual setting up of the systems as recommended. *Method:* Primarily, systematic examination of actual catalog cards for objects varying widely in type, importance, complexity, etc., to determine overall and specific characteristics of data, and interviews and consultations with users.

Status: As a result of studies of data characteristics, we converted approximately 15 years of accession files for computerization in cooperation with the Museum Computer Network. Since basic software was already available to the Network, we focused on the problems of conversion. Data was entered into the Museum Computer Network data bank at the computer center of State University of New York at Stony Brook. Source data was edited--this means interpreted and in some cases researched--directly on the source document, and then keyboarded on an IBM 2741 terminal hooked up to a time-sharing system. The time-sharing system allowed for proofing and correcting of entered data, which was then put on tape and sent to Stony Brook. This project was terminated June 30,1972.

V13. Documentation Center of Modern Swedish Art and Art Life

Principal investigator: Sven Sandström, Senior Research Fellow, Institute of Art History, U. of Lund, Box 1135, 221 04 Lund, Sweden. *Associates:* Hjördis Kristenson, Per Paulsson.

Objective: The establishment of as broad as possible a base for any sort of research within the field of twentieth-century art, art life, architecture, environment, design, etc. *Method:* Extracting material from press, letters, and other sources and systematization of this material.

Type of computer: UNIVAC 1108. *Size of storage:* 128k (36 bits). *Language:* COBOL EXEC-2.

References: S. Sandström, "A System for Multifarious Analysis and Cataloging of Information on Art," *ARIS* (1971). Available from the Institute of Art History.

V14. Analysis of Greek Vases from 530 B.C. to 500 B.C.

Principal investigator: Paolo E. Arias, Istituto di Archeologia e Storia dell'Arte Greca e Romana, Facoltà di Lettere, U. di Pisa, Via del Collegio Ricci 10, 56100 Pisa, Italy. *Associates:* Marinella Montaga Pasquinicci, Gabriella Orselini.

Objective: To analyze the mythological, religious, artistic, historical themes of the Greek vases following the books and CVA file of Beazley.

Type of computer: IBM 7090. Program partially available.

References: "Mise en fiche des vases grecs; problèmes et discussions," *La Section Linguistiques du C.N.U.C.E.*, ed. Antonio Zampolli (Firenze, 1973), pp.243-247.

V15. A Study of Methods of Presenting Information about Art

Principal investigator: Nancy A. Risser, Assistant to the Director, Computer-based Education Research Laboratory, U. of Illinois, 252 Engineering Research Laboratory, Urbana, IL 61801.

Objective: To develop a model and method for providing visitors to museums with a means of gathering information about art objects. A single art object will be used to test the model before expanding the scope of the project. *Method:* A sophisticated and flexible computer-assisted instruction system is being used to present the information. A model is currently being developed and put on this system.

Type of computer: CDC 1604. *Size of storage:* 32k. *Language:* CATO TUTOR. *Disks or tapes:* Approximately 500 tapes, 9 disks. *Special equipment:* The PLATO system, a computer-assisted instruction facility at the University of Illinois. [*CHum,* May 1969]

V16. Computer Art Language Development

Principal investigator: David Caulkins, Senior Scientist, Jacobi Computation Center, 12061 Wilshire Boulevard, Los Angeles, CA 90025.*

Objective: To develop a language which will allow artists relatively unfamiliar with programming or computers to use a computer and an associated plotter to generate two-dimensional drawings. *Method:* Write and debug the language translation program; allow artists to experiment with it and alter it to fit their requirements.

Type of computer: UNIVAC 1108. *Size of storage:* 131k. *Language:* FORTRAN V. *Disks or tapes:* 8 drives. *Special equipment:* 150 x 10 million characters of drum storage. Program is available. [*CHum,* May 1970]

V17. Computer Animation

Principal investigator: Leslie Mezei, Associate Professor, Dept. of Computer Science, U. of Toronto, Toronto, ON, Canada. *Associate:* Ron Baecher.

Objective: Developing techniques for dynamic data display, graphic simulation, educational movies, and cartoon animation. *Method:* Produce a number of specific movies, then use the experience gained to develop appropriate software for combined interactive and preprogrammed modes of operation.

Type of computer: IBM 360/44. *Size of storage:* 256k. *Language:* ARTA language, available by negotiation. *Disks or tapes:* 5 2314-A2 disks, 2 2401 tapes. *Special equipment:* IBM 2250 graphic terminal, CalComp 835 offline microfilm plotter.

References: L. Mezei, "SPARTA, A Procedure-Oriented Programming Language for the Manipulation of Arbitrary Line Drawings," *Proceedings,* IFIP 1968; L. Mezei and A. Zivian, "ARTA, An Interactive Animation System," *Proceedings,* IFIP 1971; L. Mezei and T. Britten, "Artists to Computers," *Proceedings,* National Research Council of Canada Man-Machine Communication Seminar, May 1973; (film produced) "Art from Computers," 8-minute color, commentary, music; (film)"Educational Sampler," 15-minute black and white, silent; (film) "Art Sampler," 10-minute, black and white with sound.

V18. Pattern Recognition

Principal investigator: E. Robert Ashworth, Instructor, Design and Computer Science Dept., Southern Illinois U., Carbondale, IL 62901.

Objective: To identify patterns in different alphabets of the world. *Method:* To write an algorithm to scan the given pattern and to create the vectors for use on an *xy* plotter or CRT. Roman and non-Roman alphabets will be investigated.

Type of computer: 1) IBM 7044, 2) IBM 360/50. *Size of storage:* 1) 32k, 2) 512k. *Language:* 1) FORTRAN IV, 2) FORTRAN G. *Special equipment:* CalComp plotter. Program is available. [*CHum,* May 1970]

V19. Automatic Analysis of Photomicrographs and Other Pictures

Principal investigator: Robert S. Ledley, President, National Biomedical Research Foundation, 11200 Lockwood Drive, Silver Spring, MD 20901. *Associates:* Louis S. Rotolo, Marilyn Belson.

Objective: To do automatic pattern recognition on pictures of biomedical importance for diagnostic and research purposes. *Method:* The film is scanned and pattern recognition programs are applied for the analysis.

Type of computer: IBM 360/44. *Size of storage:* 128k bytes. *Language:* FORTRAN. *Disks or tapes:* 2 disks, 4 tapes. *Special equipment:* FIDAC (Film Input to Digital Automatic Computer), DRIDAC (Drum Input to Digital Automatic Computer).

References: F. Ruddle, S. Smith, R. Ledley, and M. Belson, "Replication-Precision Study of Manual and Automatic Chromosome Analysis," *Annals of the New York Academy of Sciences* 147 (March 31, 1969), 400-423; Robert S. Ledley, "Automatic Pattern Recognition for Clinical Medicine," *Proceedings of the IEEE* 57, 11 (November 1969). [*CHum,* May 1970]

V20. Structural Description of Chinese Characters as Graphic Patterns

Principal investigator: Osamu Fujimura, Professor, Research Institute of Logopedics and Phoniatrics, U. of Tokyo, Faculty of Medicine, Hongo, Tokyo, 113, Japan. *Associate:* Ryohei Kagaya.

Objective: To obtain a structural description of the set of formal patterns representing Chinese characters and Chinese character-like patterns, through its definition by a rule system of the generative processes using elementary strokes and operators. *Method:* A hierarchical system for generating character patterns is proposed and the system is tested by actually producing the character shapes by use of an on-line computer which accepts as inputs, strings of stroke identifiers and operators.

Type of computer: PDP-9. *Size of storage:* 8k. *Language:* Assembler. *Disks or tapes:* Paper tape. *Special equipment:* Oscilloscope display. Program is available.

References: "Structural Patterns of Chinese Characters," Annual Bulletin No.3, Research Institute of Logopedics and Phoniatrics, University of Tokyo; "The University of Electro-Communications," *Current Research and Development in Scientific Documentation* No.14, National Science Foundation (1966), p.516; "A Pattern-Structural Code for Kanji," First USA-Japan Computer Conference, 1972, pp. 287-290.

V21. Automatic Picture Processing and Object Recognition

Principal investigators: Dr. Helmut Kazmierczak, Dr. Fritz Holdermann, FIM, Forschungsinstitut für Informationsverarbeitung und Mustererkennung, 7500 Karlsruhe, Engesserstrasse 5, West Germany.

Scope: Image processing, pattern recognition, scene analysis, picture preprocessing.

Type of computer: CDC 3300. *Size of storage:* 32k. *Language:* COMPASS. *Disks or tapes:* 4 CDC 854 disks, 2 CDC 604 tapes. *Special equipment:* Several scanners, displays.

References: "The Karlsruhe System for Automatic Photointerpretation," Pictorial Pattern Recognition (Washington, DC: Thompson Book Company, 1968), pp. 45-61; F. Holdermann, "Classification by Cascades Threshold Elements," *Pattern Recognition* 3 (1971), 243-251; "Preprocessing of Gray-Scale Pictures," *Computer Graphics and Image Processing* 1, 1 (1972).

V22. Sèvres Porcelain Data Bank

Principal investigators: Jack Heller, Professor, Dept. of Computer Science, SUNY, Stony Brook, NY 11790; Carl Dautermann,Curator Western Decorative Arts, Metropolitan Museum of Art, and Dept. of Computer Science, SUNY, Stony Brook, NY 11790.

Objective: To construct a computerized data base of information about the major Sèvres porcelain pieces in the major collections around the world and Sèvres workers' data. From this data base produce permuted indices, frequency dictionaries of the information collected. *Method:* The GRIPHOS 06 (*G*eneral *R*etrieval and *I*nformation *P*rocessor for *H*umanities-*O*riented *S*tudies) programs are used to store, edit and organize the input data (see G13).

Type of computer: IBM 370/155. *Size of storage:* 512k. *Language:* PL/I F. *Disks or tapes:* 1 2314 disk, 3 60 KC tapes.

References: C.C. Dautermann, "Sèvres Incised Marks and the Computer," *Computers and Their Potential Applications in Museums* (New York: Arno Press, 1968).

V23. Catalog of American Portraits

Principal investigator: Wilford P. Cole, Keeper, Catalog of American Portraits, National Portrait Gallery, Smithsonian Institution, 8th and F Streets, N.W., Washington, DC 20560.

Scope: Operates data bank on portraits of or by Americans in collections throughout the country. The files include portraits in all fine arts media and some decorative arts, but exclude photography. Supplementary files include "Minifile" (an index to portrait references in publications, art auctions, and art dealer advertising) and in-house indexes to a graphics collection of 30,000 portrait engravings, etchings, and lithographs (by subjects, makers, and media). *Method:* Photographs of portraits, physical and historical data about them, and biographical and geographical data about sitters and artists are filed manually. Analysis of these data is manipulated by computer to produce indexes to the whole file and selected listings in response to search requests. Secondary files give additional leads to the searcher. Uses computer-output microfiche, instead of printouts, for file dumps and indexes.

Type of computer: Honeywell 1250. *Size of storage:* 128k. *Language:* COBOL I. *Disks or tapes:* 2 9-million-character disks, 6 7-track tapes. *Special equipment:* Dura 1041 papertape punch and typewriter, reader with computer, Datagraphix COM unit; readers and reader-printer. Program is available.

References: Daniel J. Reed, "Catalog of American Portraits," *American Archivist* 30, 3 (July 1967).

V24. Visual Structures 1,2,3

Principal investigator: Peter Struycken, Lecturer, Dept. of Monumental Art, Akademie voor beeldende Kunsten, Onderlangs 9, Arnhem, Netherlands. *Associates:* S. Temperlaars, G. Vermeulen.

Objective: To investigate the relation, conditions, and possibilities of visual elements (2-dimensional as yet). The relationship of element-structure is important and the establishment of a norm impression according to which the differences between structures can be determined. *Method:* Restriction of chance or prevention of repetitions in order to arrive at various types of structures.

Type of computer: Electrologica X8. *Size of storage:* 24k. *Language:* ALGOL. *Disks or tapes:* 1 disk. Program is available.

References: Electronic Music Reports 2/1970 (Utrecht State University); *Bulletin of the Computer Arts Society* (June 1970), p.8; *Computer in the Creative Arts,* NCC, Quay House, Manchester M3 3HU, England; *Geluid-Kikjen* (a catalog of the Stedelijk Museum), Amsterdam, Netherlands. See also L466. [*CHum,* November 1971]

V25. Literaturbericht zur Kunstgeschichte

Principal investigators: Stephan Waetzoldt, Generaldirektor, History of Art, Staatl. Museen Preuss., Kulturbes, Staffenbergstr. 41, 1 Berlin 30, West Germany; Christina Thor, History of Art, Kunstbibliothek, Staatl. Museen, Jebensstrasse 2, 1 Berlin 12, West Germany.

Scope: Current bibliography (with abstracts) of German language periodicals on art history. *Method:* Analogous to International Repertory of Music Literature (RILM).

Special equipment: Computering made by Zentralstelle für maschinelle Dokumentation, Frankfurt-Niederrad, Herriotstrasse. Program is available.

References: Berichte aus den Staatl. Museen der Stiftung Preuss. Kulturbesitz, Berliner Museen, N.F., 22, H1 (1971), 55-56; *H. Kunstchronik* 5 (1971), 88.

V26. Advanced Computer Graphic Characters

Principal investigator: David D. Cranch, Research Fellow, Dept. of Graphic Design, Royal College of Art, Kensington Gore, London SW7, England. *Associates:* R. Guyatt, H. Spencer, Brian Coe.

Objective: 1) To investigate the behavior of characters when aspects of their design are altered, and to deduce numerical rules (e.g., shape changes required if the weight of the character is increased). 2) To develop a program which can generate characters in a range of continuously variable styles and output them to a high-accuracy plotting table to annotate maps drawn by a research unit in the College. *Method:* Characters are defined relative to 9 x 7 mesh and are made up from straight lines and (initially) quarter circles. By adjusting the rows and columns of the mesh the style of a character is altered, and the quarter circles become arcs of ellipses. Thus, for example, an "O" can be drawn in any shape from a circle to a square and the effect of this on spacing can be studied and reduced to a numerical rule. The rule can be then built into the program. A tachistoscope will be used where necessary to examine the output and ensure that the "readability" of the characters is at an optimum.

Type of computer: 1) PDP-9, 2) PDP-15. *Size of storage:* 1)16k, 2) 24k. *Language:* 1) and 2) FORTRAN IV. *Disks or tapes:* 1) 1 RF15/RS09 disk, 3 TU56 tapes; 2) 2 RF15/RS09 disks, 3 TU56 tapes. *Special equipment:* AEG Geograph flatbed plotting table, storage tube display, and joystick. Program is available. Project is funded to February 1973 or 1974 and program will be updated continuously. [*CHum,* November 1972]

See also G23, G36, M8.

Index