

The Computer and Theological Materials

Theologians and theological librarians now have the computer to aid them with theological materials and research. In particular, the computer is being used in concordance and index production, in linguistics, in problems of biblical authorship, in archaeology, and in machine translation. A survey of the use of the computer with theological materials and basic computer techniques connected with the projects is presented.

IN A WORLD WHERE the computer with its diverse uses is taken for granted, the field of theology is doubtless considered by some as one of the last strongholds against this modern form of gadgetry. Because theological librarians, in general, have separated themselves from the scientific realm of the computer, Kathryn Luther Henderson in 1962 rebuked them for keeping to their isolation in an age when it takes longer to catalog a dozen books than it does to orbit a man around the earth.¹ A challenge was issued by George H. Bricker, past president of the American Theological Library Association, when he likened the world of computers to a foreign culture where the language is unknown. He cautioned the timid traveler that "if the art of theological librarianship is to advance, we must at least step into the new world of the machine and learn to know our way around."²

¹ Kathryn Luther Henderson, "Keeping and Casting Away: Cost Implications of the Library Development Program for Technical Services," *Proceedings of the Sixteenth Annual Conference of the American Theological Library Association* (Hartford, Connecticut: 1962), p. 152.

² George H. Bricker, "The Theological Librarian and the New Era of the Machine," *Proceedings of the Twentieth Annual Conference of the American Theological Library Association* (Louisville, Kentucky: 1966), p. 39.

Aside from the obvious uses of the computer in the theological library for circulation control, technical processes, and information storage and retrieval, there are rudimentary computer techniques and less well-known applications of the computer that should be of interest to theological librarians and their clientele. Two examples of areas in which such experimentation has been attempted are concordance production and literary analysis for authorship identification.

Concordance production has been the most successful application of the computer to theological materials. Historically, Father Roberto Busa of the Centro Automazione Analisi Linguistica in Gallarate, Italy, made one of the first important contributions in the use of mechanization for concordance-making. In 1949 Busa started experimenting with IBM accounting machines for some of his philological research.³ His initial project was to produce a concordance to the works of St. Thomas Aquinas, and by 1956 the work was converted to the electronic digital computer. In blessing the computer for the project, Archbishop Giovanni Battista Montini of Milan (later Pope Paul VI) stated that the computer had "become a modern means

³ Paul Tasman, "Literary Data Processing," *IBM Journal of Research and Development*, I (July 1957), 249-56.

Mr. Taylor is Southwest Regional Librarian for the Tulsa City-County Library System in Oklahoma.

of contact between God and man."⁴ The *Index Thomisticus* was to have been completed in 1965 by Busa's staff of sixty.⁵ Busa's main goal in preparing the concordance was that of "producing the larger base from which may be induced more valid quantitative information on language components. Our effort has entailed in particular the application of techniques to the preparation of indices and concordances which are the true materia prima for a historical diachronical dictionary of a language."⁶ The end result of this philosophy has been the transferral of fifteen million words of eight languages (Latin, Greek, Hebrew, Aramaic, Nabataean, German, English, and Italian) onto computer tape. Of this material 1,967,000 lines of three categories of St. Thomas' works are included: the concordance containing twenty-five major and forty-six minor authentic works of St. Thomas; sixty-three attributed works; and Aquinas' commentary on the Latin Vulgate, Peter Lombard, and Aristotle.

The use of the computer in the literary analysis of Aquinas' works was completed in five steps.⁷

STEP I

First, the scholar must examine the text in order to devise exact instructions for the transferral of the text onto punched cards by a keypunch operator.

STEP II

In the second step for producing the concordance, the clerk keypunches "phrase cards" which contain the individual lines or phrases of the text, with each card containing one phrase. Each phrase card is issued a serial num-

ber, reference number to location of the phrase in the text, and a "special reference sign." These punched cards are then verified; in other words, the clerk keypunches the same information onto each card in order to detect mistakes in the original keypunching.

STEP III

The computer then performs the third step by producing a phrase-by-phrase copy of the text. In addition, "word cards," which contain each word of the phrase cards, are produced on separate punched cards. Each word card is given one of four identifying signs by the scholar:

- / for passages in which Aquinas refers to passages in his own work
- # for words Aquinas quoted from other authors
- for passages which cannot be considered to be Aquinas' because he quotes the doctrine of others, comments on them, or refutes them
- ' for references Aquinas makes to his own works

STEP IV

The computer performs the fourth step in concordance-making. Duplication of words is eliminated in this step with the production of "form cards" which contain each "graphically" different word on separate punched cards. On each form card is printed the frequency of occurrence of the word in the text and a number showing the position of the word in the alphabetical listing.

STEP V

The scholar takes the alphabetical list with its numerical notations and performs the fifth step. He *takes this list and* "groups all the different forms of one and the same *graphic-semantic* unit under the single expression or word

⁴ "Sacred Electronics and St. Thomas Aquinas," *Time*, LXVIII (December 31, 1956), 48.

⁵ "Aquinas on IBM Tape," *America*, CXI (October 17, 1964), 433-34.

⁶ Roberto Busa, "An Inventory of Fifteen Million Works," *Proceedings of the Literary Data Conference* (Yorktown Heights, New York: 1964), 64-65.

⁷ Tasman, *op. cit.*

which will serve as an entry listing (as well as) separates all the homographs into their various uses."⁸ Therefore, the words *were*, *are*, and *be* would be listed or entered under *to be*. The entry words are punched onto cards, and the computer alphabetizes them and numbers them in sequence. At this point, a lexicon based on graphic-semantic relationships of words can be printed. As the computer groups the form and word cards, the entry cards are punched with the frequency number of appearances in the text. The five-step process allows four forms of the material to be produced.

Laterculum Formarum—list of form words

Conspectus Lemmatum—list of all entry words

Rationarum Verborum—list of entry words followed by their forms

Prospectus Homographiae—list of homographs followed by entry words⁹

An opportunity for Busa to show the versatility of his concordance procedures was made possible in 1958 when the computer was used with the Dead Sea Scrolls. The scrolls were in a decayed condition with many loose pieces that needed to be identified. Paul Tasman, who called the situation a giant jigsaw puzzle, summarizes the principal use of the concordance with the Dead Sea Scrolls.

... the application of the principles to the indexing of the Dead Sea Scrolls had to take into account the many plates in which there were incomplete words or words totally obliterated. This presented no particular problem inasmuch as the program provided for an analysis of the text, taking into account the frequency use and sequence of words in a particular text and their context. It becomes apparent that this forms a rudimentary system for analyzing the writing style of an author and also a

tool to interpolate missing words or to detect foreign additions which are uncommon to the author. While it will not be absolutely certain that exact substitution will be made, a more accurate machine substitution will be possible than is conceivable by manual methods.¹⁰

Another pioneer in the field of computer concordances is the Rev. John W. Ellison of the Church of Saint Clement in El Paso, Texas. His *Nelson's Complete Concordance of the Revised Standard Version Bible* was published in 1957. It took the UNIVAC I computer only four hundred hours to produce the Bible concordance in comparison with the thirty years it took James Strong to compile his King James Version Concordance published in 1894.¹¹ The RSV Concordance consists of individual word entries with text entries up to thirty-five characters in length under each word entry with the location of the quotation to the right of the text entry. In the preface to the concordance, Ellison states that while the concordance could be called exhaustive, it is not analytical since the Hebrew and Greek words of the original text could not be listed.¹²

To produce the concordance, the 780,000-word text of the RSV Bible was punched onto two separate magnetic tapes.¹³ These tapes were fed simultaneously into the computer for comparison and correction of errors. The corrected master tape was again fed into the computer where the whole text of the Bible was broken down into single words. This produced sixty-five tapes with single words and their text locations registered. At this point the computer was instructed to eliminate all occurrences of 127 words from the text.

⁸ Tasman, *op. cit.*, p. 256.

¹¹ "Bible Labor of Years Is Done in 400 Hours," *Life*, XLII (February 18, 1957), 92.

¹² John W. Ellison, comp., *Nelson's Complete Concordance of the Revised Standard Version Bible* (New York: Nelson and Sons, 1957), p. v.

¹³ John W. Ellison, "Can a Computer Write a Book?" *The Living Church*, CXXXIII (September 23, 1956), 7-9, 30-31.

⁸ Tasman, *op. cit.*, p. 255.

⁹ Busa, *op. cit.*

Words like *the*, *and*, and *of* were considered to be of little help to the user of the concordance. By doing this, 59.4 per cent of the text was eliminated. The trimmed set of tapes (now twenty-six in number) was sorted and merged and an alphabetical printout followed. The concordance contains 305,242 entries with about six thousand to seven thousand single words represented.

Rev. Ellison is also interested in using the computer in New Testament Greek studies. Because of the problem of so many variant Greek manuscripts of the New Testament, Ellison has been working on collating eight hundred of these variant manuscripts¹⁴ in order to produce a "neutral" text that will help him to place the manuscripts into families of common ancestry.¹⁵ Ellison has been working on a project to produce many computer concordances of a morphological (rather than etymological) nature of important writers.¹⁶ The *Talmud*, Septuagint, and early Greek literature will be among the first to be processed.

The early Church Fathers are the subjects for a concordance project by J. K. Siberz and J. G. Devine. The objective of their project from the outset was to produce an inexpensive concordance for scholars.¹⁷ In 1964, to test the feasibility of producing a concordance for the early Latin and Greek Church writers, three thousand words of Tertullian's works (*Ad Martyros* and *Ad Scapulam*) were punched onto cards and fed into the IBM 1620, Model I computer. Three concordances were generated: "a selective concordance containing only words

considered to be of interest to a theologian; a complete concordance, arranged in strict alphabetical order; and a complete concordance arranged so that inflected forms of the same word occur together under the principal form, with the principal form arranged alphabetically."¹⁸ Finding this experiment to be successful, Siberz and Devine employed the IBM 1401 computer which had a larger capacity for the input of the complete works of Tertullian, amounting to 300,000 words. The Tertullian concordance will contain 1,300 pages.

Another area in which the computer has been helpful in analyzing theological materials is that of various computerized indexes. In this connection John Wells has been using the computer to compile a word index of Old High German manuscripts covering the period 750 A.D. to 1065 A.D.¹⁹ The first step was to enter all of Elias Steinmeyer's work on *Die Althochdeutschen Glossen* onto three-by-five cards. From this Wells plans to prepare a word index that will be one of the most complete dictionaries of Old High German. The broadest part of the project is to normalize the text since so many variant glosses are in existence. Second in importance to the production of a dictionary will be the capability of the scholar to compare the computer printout of unmodified forms with the original printed glosses. Scholarly pursuits with Latin words in the glosses have all been made possible because of the computer.

Roy Wisbey of Cambridge University has been indexing Middle High German manuscripts of the eleventh and twelfth centuries. He started with the six thousand lines of rhyming couplets of the *Wiener Genesis* which is important for its Old Testament subject mat-

¹⁴ E. G. Fogel, "The Humanist and the Computer: Vision and Actuality." *Proceedings of the Literary Data Processing Conference* (Yorktown Heights, New York: 1964), p. 19.

¹⁵ Claire M. Cook, "Automation Comes to the Bible," *Christian Century*, LXXIV (July 24, 1957), 892-94.

¹⁶ "Computerized Research in the Humanities: A Survey," *ACLS Newsletter*, XVII (June 1966), 1-49.

¹⁷ J. K. Siberz and J. G. Devine, "Computer-Made Concordances to the Works of the Early Christian Writers," *Proceedings of the Literary Data Processing Conference* (Yorktown Heights, New York: 1964), 128-47.

¹⁸ Siberz and Devine, *op cit.*, p. 130.

¹⁹ John C. Wells, "A Word-Index and Dictionary to the Old High German Glosses," *Proceedings of the Literary Data Processing Conference* (Yorktown Heights, New York: 1964), 148-59.

ter as well as for linguistic considerations.²⁰ Besides the usual concordance or word list that can be produced, a reverse index will be produced which will allow the scholar to analyze the suffixes of the text in the same way he can analyze the prefixes with word lists. Looking ahead to future possibilities, Wisbey finds exciting the prospect of compiling a complete tape library of the important eleventh- and twelfth-century German works. By 1965 the *Speculum Ecclesiae* (a twelfth-century sermon collection) and volume one of Schönbach's *Altdeutsche Predigten* had been punched onto paper tape.

A computerized index in the truest sense of the word has been produced by Sister M. Claudia of Marygrove College. When Sister Claudia became the index editor for the *New Catholic Encyclopedia* she consulted with some of the leading publishers of encyclopedias to see if there was a way to automate the production of the index for the *New Catholic Encyclopedia*.²¹ She was repeatedly told that none existed, but by July of 1964 she had contracted Documentation, Inc. to develop an automated indexing system. The main idea was to index the galley sheets as they came from the McGraw-Hill publishing house. All of this would be coordinated by a computer program. After biographical entries had been established, article headings compiled, and filing rules decided upon, a style manual was written. By July of 1966 all galley indexing was done except for the illustrations and maps. In November of the same year, the three thousand edited single-column pages of printout were sent to the printers. It was found that although it took as much personnel for automated indexing as for manual indexing, the computer was invaluable in helping to keep

²⁰ Roy Wisbey, "Computers and Lexicography," *The Use of Computers in Anthropology*, ed. by Dell Hymes (The Hague: Mouton and Co., 1965), p. 215-34.

²¹ Sister M. Claudia, "Automated Techniques in Comprehensive Indexing," *The American Archivist*, XXX (April 1967), 287-94.

to the rigid timetable. In the future, the index tapes will be used for the updating of the future editions of the index to the *New Catholic Encyclopedia*.²²

Two other interesting applications of the computer to theological materials are those of archaeology and machine translation. In the field of archaeology, much has been done in North American archaeology with the computer, but J. C. Gardin of the National Center of Scientific Research in Marseille, France, has worked with archaeological data and has devised codes which can be transferred to punched cards that will describe the archaeological artifact or monument.²³ John McDonough, Jr., predicts that in twenty years a three-dimensional scanner will be used to describe the shape of an artifact placed on a grid.²⁴ A beam of light would describe the item and would store the description in the computer. With all of this stored on computer tape, many items could be pieced together. This work could have many implications for the future of biblical archaeology.

Of concern to theologians is the work being done by the American Bible Society with the machine translation of natural languages. Eugene Nida has been directing a program to "develop a pedagogically oriented approach to the problems of the theory and practice of translation, with particular reference to the Bible."²⁵ Nida further states that results have shown that machine translation is applicable to single documents of limited subject material and that with the improvement of computers, linguis-

²² "Encyclopedia's Index to be Assembled by Computer," *Automation Reports*, No. 71, Part I (May 4, 1956), 1.

²³ *Current Research and Development in Scientific Documentation*, No. 14 (Washington, D.C.: National Science Foundation, 1966), p. 306.

²⁴ James T. McDonough, Jr., "Homer, the Humanities, and IBM," *Proceedings of the Literary Data Processing Conference* (Yorktown Heights, New York: 1964), 34-35.

²⁵ *Current Research and Development in Scientific Documentation*, No. 10 (Washington, D.C.: National Science Foundation, 1962), p. 162-63.

tic investigation will greatly benefit. A look into the future by John K. Mickelson prophesies the following for the evangelical user of the computer:

The language-handling abilities of computer programs may soon advance to the point where they can aid the Christian missionary who is reducing a tribal language to writing. Word frequency counts and structural statistics would provide the basis for a pedagogically sound method of teaching the language. The computer may become able to expedite the translation of large portions of scripture, by rapidly producing tentative translations for the missionary to improve until they are ready for publication. . . .²⁶

The most controversial use of the computer in the theological field has been its application to authorship problems. One such application raised quite a furor when the Rev. Andrew Q. Morton, a Scottish Presbyterian minister, in November of 1963, proclaimed that with the computer he had ushered in the Industrial Revolution in New Testament studies with his work on Paul's Epistles.²⁷ Morton had been working with G. H. C. MacGregor for seven years with a £800 grant from the Carnegie Trust to buy the machine for their work.²⁸ They proposed to devise a test for authorship by detecting unconscious habits of Greek writers. Morton concludes from his computerized study that of the fourteen epistles examined, five (Romans, First and Second Corinthians, Galatians, and Philemon) were written by one person, namely Paul, and that the remaining nine epistles were written by five different authors. Morton further states:

Now by the use of modern scientific techniques, including the operation of a computer . . . it has been proved beyond doubt that Paul himself wrote only four

major Epistles. Theologians all over the Christian world have now to face the implications of his discovery. They must change their view of the life of Paul, they must revise the history of the early Church and must jettison doctrines that have been shown to be without foundation. . . . Once again authority is called upon to yield to knowledge.²⁹

A week after Mr. Morton presented his case in the November 3, 1963, issue of the London *Observer*, five leading churchmen were asked to comment on Morton's claims. Rev. John Robinson, author of *Honest to God*, said that Morton should have spent more time giving facts and figures than *odium theologikum*.³⁰ Rev. Leslie Weatherhead said: "I was delighted with Mr. Morton's article. . . . Frankly, who cares whether Paul wrote this book or that? All that matters is not authorship but truth and relevance."³¹ Rev. Thomas Corbishley said Morton had shown the value of the computer in discovering Paul's literary composition methods and that scholars had long accepted the fact that Paul had ghost writers and there was no reason why those writers could not have put Paul's ideas into their own words.

The controversy dropped out of the news until December 1964 when *Christianity Today* gave its opinion of Morton's attitude:

What is biting Mr. Morton so hard? It is very likely more than a problem of personal relations. Apparently it stems from the argument that if it could be proved that Paul did not write one or another of the epistles attributed to him by the Bible, the trustworthiness and the authority of the Bible would fall.³²

The writer goes on to wonder if Morton considered the facts that Paul was trilingual, may have dictated some of the epistles, and quoted much from the Old Testament.

²⁶ Morton, *op. cit.*

²⁷ "The Church and the Computer," London *Observer* (November 10, 1963), 21.

²⁸ "The Church and the Computer," *op. cit.*

²⁹ Reinier Schippers, "Paul and the Computer," *Christianity Today*, IX (December 4, 1964), 7-8.

²⁶ John K. Mickelson, "Computers in God's Service," *Christianity Today*, IX (January 15, 1965), 49.

²⁷ "The Rev. A. Q. Morton Replies," London *Observer* (November 10, 1963), 21.

²⁸ A. Q. Morton, "A Computer Challenges the Church," London *Observer* (November 3, 1963), 21.

Rev. John Ellison entered into the debate and stated at a conference at Yale University that Morton's method was unsupported and "an abuse of both computers and scholarship."³³ Ellison had applied Morton's method to Morton's own articles that appeared in the London *Observer* and concluded that two different authors had written the articles. Ellison also used Morton's technique on James Joyce's *Ulysses* with the result that he said he could prove it had five authors. In addition, Ellison had analyzed the Declaration of Independence and letters written to Mrs. Jefferson and has concluded that either Thomas Jefferson did not write the Declaration of Independence or that some other man was writing love letters to Jefferson's wife!³⁴

What method had Morton used that was so objectionable? Morton starts with the assumption that within any class of writers (in this case classical Greek writers) there are unconscious habits that can be used as a test for authorship.³⁵ He finds that common words like *the* are good tests for authorship since these types of words do not depend upon the subject matter of a work. He believes that a writer places these common words in an unconscious yet consistent manner. In order to apply these tests for authorship Morton uses what he calls purely objective and scientific methods which are not matters of personal opinion as is literary criticism. Therefore, the principles of statistics with the use of probability and sampling become important. Morton asked classical scholars to select samples of classical Greek writers which were ex-

amples of a great span of time and a variety of subject material. As a result such writers as Herodotus, Thucydides, Demosthenes, Plato, Isocrates, and Plutarch were chosen. Morton found that six common Greek words (*Ho, He, To, Kai, Autos, De, En, and Einai*) comprised about 31 per cent of the whole text of these Greek writers. These words were used for testing authorship. By instructing the computer to count the number of common words contained within each sentence in text samples, Morton presents tables of these calculations in which he shows the correlations of an author's use of these common words in most of his works. By establishing that there is a consistent use of the six common words in all types of classical Greek prose, Morton concludes that it is a valid test for authorship. Upon applying this technique to Paul's fourteen epistles and by using Galatians as a "real" work by Paul, Morton tested each epistle and compared the use of the six common words with the pattern established in Galatians. These tests led Morton to state that Paul wrote only five of the fourteen epistles and that the other nine were written by various hands. More experimentation by theological scholars is needed before the validity of the use of the computer in this kind of biblical analysis can be substantiated.

The intent of the above discussion and its accompanying table has been to point out the great activity that is now in progress with materials of interest to the theological community. Theological librarians must be prepared to make available these and other new computerized tools to their patrons. Theologians themselves need to be aware of these new modern resources and techniques that deal with their work. Even more important, the talents and enthusiasm of brave theologians are needed for serious computer research. The judicious use of the computer can greatly advance theological scholarship to the

³³ McCandlish Phillips, "Computer Flouts Test by Another—Study on St. Paul's Epistles Questioned at Yale Parley," *New York Times* (January 23, 1965), 17.

³⁴ Philip H. Smith, Jr., "The Computer and the Humanist," in *Computers in Humanistic Research*, ed. by Edmund A. Bowles (Englewood Cliffs, New Jersey: Prentice-Hall, 1967), p. 22-23.

³⁵ A. Q. Morton and James McLeman, *Paul, the Man and the Myth* (New York: Harper and Row, 1966).

point where the machine will be called upon to perform time-consuming duties, thus freeing the scholar to a more creative use of his intelligence.

TABLE 1*

VARIOUS PROJECTS IN WHICH THE COMPUTER HAS BEEN USED WITH THEOLOGICAL MATERIALS WITHIN THE LAST FOUR YEARS

Category	Researcher	Affiliation	Subject
Authorship	Alastair McKinnon	McGill Univ.	Kierkegaard's pseudonymous works compared with established works
Collation of Variant Manuscripts	James Halporn	Indiana Univ.	Study of the text-history of Casiodorus' <i>Psalm Commentary</i>
	Mark Lansburgh	Colorado College	Division of 8th, 9th, and 10th century Visigothic manuscripts into stylistic schools and ascribing these schools to specific scriptoria
	David Sarne	Brandeis Univ.	Computer-generated critical edition of Rabbinic texts
Concordance Production	Ford Battles	Hartford Seminary Foundation	Computerized concordance to John Calvin's <i>Instituto Christianae Religionis</i>
	Lawrence Berman	Dropsie College	Semitic Language Concordance from computers
	Miguel Civil	Univ. of Chicago	Sumerian Computer Concordance
	Melvin Emmons, et al.	Texas A & M Univ.	Concordance to <i>Book of Common Prayer</i>
	Alastair McKinnon	McGill Univ.	Concordance to Kierkegaard's <i>Samlede Vaerker</i>
	Alastair McKinnon	McGill Univ.	Hymns from the <i>Canadian Frontier</i>
	Paul Pillsbury	Eastern Michigan Univ.	A concordance to the West Saxon Gospels
Indices	J. Gordon Spaulding	Univ. of British Columbia	Chronological rearrangement of John Cooke's <i>The Preacher's Assistant</i> , a collection of English sermons from 1660 to 1783
Linguistic Studies	Nils Enkvist	Abo Akademi, Abo, Finland	Graphotactic Analysis of the Old English St. Chad
	Dewey Faulkner	Univ. of North Carolina	Study of phonology of Aelfric's Catholic Homilies, First Series
Miscellaneous	David Sarne	Brandeis Univ.	Structural analysis of Hebrew roots
	J. Arthur Baird	College of Wooster	Audience tradition in the Synoptic Gospels
	M. F. Bott	Univ. of Cambridge	<i>Speculum Ecclesiae</i>
	Charles Colbert	Antioch College	Corpus of the Shekal coinage in first century Judea
	Stephen Foster	Northern Illinois Univ.	(1) The Puritan social ethic (2) The good old cause in Old and New England, 1630-90
	David Freedman	San Francisco Theological Seminary	Study of the Leviticus Scroll from Qumran Cave II
	Finn Hansen	Univ. of Aarhus, Denmark	Statistical investigation of tonal properties of the Gregorian repertory
	Sherman Hawkins	Univ. of Rochester	Studies in religion and Renaissance tragedy
	Lee Ohringer	Univ. of Pittsburgh	<i>New English Bible</i>
	Richard Vann	Wesleyan Univ.	Demographic history of English Quakerism 1655-1837 using birth, marriage and death records
			Book on the interrelations between psychological and ethical concepts
Psychology	Raziel Abelson	New York Univ.	Computer-aided study of the emotional factors in religious motivation
	Louis Gaffney	Seattle Univ.	

* See footnote 16

