Research into the history of man’s culture and his institutions has always been conducted with procedures which have a basic grammar. Upon that basic grammar scholarship has developed, since Poliziano, ever more complex routines as the raw materials for research have proliferated. The provision of these raw materials has been, since the Renaissance at any rate, one of the prime functions of historical and enumerative bibliography. Oddly enough, the basic grammar has not often been formalized. I suspect that this is due to the fact that we are more often than not required to exercise autodidactic skills in learning it. And we have learned how to develop the special skills needed to deal with the documents of the past (printed and manuscript) through imitation of examples furnished by previous failures and successes.

Statements about the ‘grammar of research’ have necessarily been tentative, because the evidence upon which research operates is both diverse and fugitive. Frequently we fail to see it; as often we fail to recognize it. Which is why the discipline which most readily lends itself as a model is archaeology. The bibliographer, the scholar, like their colleague the archaeologist, must sift mountains of rubble to recover fragments of truth, and the objects of their discoveries are ungenerous witnesses revealing always less than could be wished. The interrogative procedures of the historian constitute the grammar, and this principle is one we unconsciously invoke when we condemn a piece of research as the mere accumulation of data destitute of interpretation. And, if we are honest, we must concede that our libraries are overflowing with books in which interpretation subordinates evidence and vice versa. The newcomer to research, whatever his discipline, can be forgiven dismay at the sheer bulk of the sources at his disposal, sources which seem to multiply with undisciplined rapidity. The threat to enlightenment lies less in the insufficiency of information than in its overwhelming flood.

That flood began almost exactly a hundred years ago. The 1870s witnessed developments in the technology of printing which encouraged the belief that, within a generation, co-operative effort could produce a ‘universal catalogue’. A scheme for the reduction of all knowledge of books in the world—within the framework of a usable catalogue—was born in the year preceding the Great Exhibition of 1851, and was actively promoted by the Society for Arts for twenty-five years. The simplicity and grandeur of the scheme was typical of Victorian optimism. That it came to nothing is, in retrospect,
hardly surprising. So often in history positive optimism has been frustrated by negative
innocence. What the promoters of the grand design failed to comprehend was that a ‘book’
is not always what it appears to be, and that the number of such ‘books’ in the
bibliographical universe was beyond the capability of existing techniques to handle. The
publication of the British Museum’s General Catalogue was an immense step forward, but
even that prodigious inventory (as we now know) was but a start along a road the length of
which is still undetermined. The tools for bibliographical research have, since 1870,
multiplied to the point where the great research libraries of the world face paralysis—and
very expensive paralysis at that. An accountant’s estimate of the ‘real’ cost of maintaining
and accommodating the card catalogue and the printed catalogues in the Yale University
Library suggests that these tools have become insupportable burdens. It is probably no
exaggeration to suggest that the British Museum Library spent more resources on the
maintenance of the various Reading Room catalogues between 1787 (when the first
printed catalogue of the collections appeared) and the creation of the British Library in
1973 than it did on acquiring the collections. Microreproduction, once regarded as an
inexpensive solution, is clearly insufficient: for, while its demands on space are acceptable,
its demands on patience are considerable, access to data is tedious, and comparative
consultation of more than one source is impractical.

Consequently, in the very period through which we are living, the perception that
traditional methods of storing and transmitting information (upon which all research
depends) are inadequate has led to the conclusion that the computer, with its astonishing
capacity to store huge quantities of data, and to provide access to such data in a time-frame
where a minute is an eternity, is our only hope. It is a bright hope: but let us remember
that our optimism will be of little account if it is partnered with innocence. It would be
futile to suggest that merely to computerize knowledge renders it more tractable to the
enquiring mind. The enquiring mind has always demanded a grammar to direct its
enquiries, and a subtlety to direct interpretation of what its enquiries have revealed.

The discursive arrangement of much scholarship in the humanities is predicated upon
the leisurely, and essentially private nature of its transaction. This is true whether the
work is primarily bibliographical or explicatory. The methods by which a researcher
accumulates his material tend to be highly individual—we all have our idiosyncratic
methods of organizing information. And the framework within which the results of
research are communicated tends to follow traditional models. The editing of classical
texts is a case worth noting, for, while it is obvious that great advances were made between
the endeavours of Aldus and those of Lachmann it is a study of some intricacy to
determine between any adjacent editors the precise nature of the advance. Perceptions
may be noted; but what of the grammar underpinning those perceptions? Johnson’s
edition of Shakespeare seems so much better than Rowe’s; Steevens’s better still. It is sad
that we know so little about the methods which the great figures of the past used to
enhance their perceptions. What would one not give to comprehend the ‘grammar’ that
led Johnson to the Preface to his edition of Shakespeare. Scholars have always been
careless about revealing the grammar upon which their insights are based, which perhaps
explains the historian’s complaint that the discovery of truth has mostly to do with the discovery of error.

It seems clear to me that while we may be sensitive to the virtues and deficiencies of different approaches to scholarly procedures, admiring particularities in method or feeling discomfort with presentation, it is exceedingly difficult to determine why we respond to a work of scholarship with gratitude or disapprobation. The researcher, after all, has a distinct edge over the reader, unless the reader is prepared to retrace the steps which led to the work in question. Few reviewers are prepared to undertake such a task; and the reader, unless he happens to be expert in the subject-matter, will very probably accept unquestioningly the premisses and the conclusions. This is so, because in addition to ‘grammar’ research has ‘rhetoric’. I mention this not to excite the notion that I am bent upon some neo-Ramist interpretation, but merely as a reminder that when we have gained some insight into the grammar of scholarly method we have still some way to go before we can fully comprehend the reasons why Johnson’s Shakespeare is better than Rowe’s. A work of sound and persuasive scholarship can appeal as distinctly to our hunger for truth and beauty as a work of art. That such works are rare is due, no doubt, to the fact that scholarship, like politics, is altogether too rhetorical.

It is not my task here to reconsider some of the great monuments of the history of scholarship. I leave that to another: preferably one with the immense learning of a Lewis and the imaginative grasp of a Frye. It will be, if successfully undertaken, a work of scholarship of fundamental importance. It will be performed, I am certain, with mechanical aids of a kind which we can only vaguely comprehend. Unlike the casual, improvised apparatus to which we have become accustomed, the aids of the future, predicated upon a grammar as severe as anything dreamed of in the logic of Leibniz, will demand of us who will use them an uncompromising discipline in approach and method. To some this may seem a future for scholarship too bleak to embrace, but it is as inevitable as tomorrow.

Those who find it difficult to grasp the implications of the computer for the storage, manipulation, and transfer of information associated with the traditional scholarly disciplines can take some comfort from the fact that the professionals are as unsure about the future as the amateurs. Which is precisely why this is such an interesting period in which to be living and exercising our skills. We have a rare opportunity not only to benefit from the evolving technology, but actually to use that technology to enhance our methods of dealing with information now available in dimensions of unprecedented magnitude. In every discipline the threat exists of surfeit: too many books, too many theses, too many articles. A recent survey conducted by the Library of Congress reveals that seven thousand scientific articles are written every day, and that scientific and technical information doubles every five years. That this rate will increase can hardly be doubted. The ultimate irony may turn out to be that while we drown in information we starve for knowledge, a theme which would have appealed to Coleridge.

Universities must accept a portion of the blame for this state of affairs, for it is they, more than any other agency, which have transformed an activity into an industry. And it is
ironic that the very industry which they have created now threatens them. Our research libraries, upon the resources of which the future of research depends, can barely cope with the present flood, let alone conserve for posterity what they have accumulated in times past. For small and specialist libraries the problems are probably soluble, but for major university and national research libraries the problems seem beyond solution. Unless, that is, we grasp the opportunities provided by the computer, just as Europe, five hundred years ago, grasped the opportunities provided by printing with movable type. It has been suggested that the invention of printing led inexorably to the transformation of an agriculturally based civilization into one in which the industrial revolution could take place. That process was a slow one, and may be traced to Bacon’s vision of the ‘New Science’. The electronic revolution, the beginnings of which we are now witnessing, promises to repair some of the social damage which industrialism and metropolitan civilization inevitably wrought, and one of its most important benefits may be the fact that it will permit participation (the participant instinct is a powerful one) and individualism without the necessity to live and work as though members of a herd. Social gregariousness is one thing; the regimentation of the factory and the inflexible constraints of metropolitan life are quite another. The computer may, if we have the imagination to use it wisely, not only make the flood of information more manageable, we might even rediscover a delight in learning—what Coleridge called the ‘inquiring spirit’.

It has always appeared to me strange that innovation has a tendency to bring with it a rigid orthodoxy. Printing—the invention that brought mankind in Western Europe from darkness into light—rapidly developed an orthodoxy which has changed little in five hundred years. The significant features of a twentieth-century book differ very little from those of the books of the fifteenth century. Physical dimension depends as much today (in spite of international metrication) upon the basic geometry of an animal’s skin. And here we are, in the middle of a revolution as profound as we have ever witnessed, developing an orthodoxy to constrain the capability of microelectronics to enfranchise information. We have international standards, and we have a special brotherhood of well-meaning bureaucrats unable to see beyond the comfortable logic of that vertical thinking which made trains in the image of coaches. Thus, in America, we have seen in the past few years the almost universal subscription by libraries to bibliographical networks such as OCLC which, in spite of the enormous detail which its twelve million records contain, is just about as accessible as a respectable card catalogue, and the system produces, with ingenious and admirable efficiency, countless millions of catalogue cards every year. One is tempted to conclude from this that a printer is nothing more than an efficient and economical scrivener, and the computer nothing more than an efficient and economical typewriter.

The world, it sometimes seems, has always been divided between those possessing skill and those whose very existence depends upon the availability of skill. Substitute ‘mystery’ for ‘skill’ and you will perceive my message. The architects of the computer systems increasingly found in libraries throughout the world possess ‘secrets’. It is almost as if freemasonry has been granted a second chance. But it need not be so. For humanism has
been built upon agony and struggle; it has endured persecutions, tyranny, and every form of exquisite martyrdom. I believe it to be both enduring and durable.

Computers are, in fact, altogether without mystery—a 'miscellany of parts' as Swift might have described them. Certainly the Houhynyms would have understood them. It is precisely because they lack mystery that they should be so appealing to those responsible for encouraging research. Utterly docile, utterly stupid, they perform precisely what we ask them to perform. But they are unable to perform beyond the grammar they have been taught. And only we, who have created the instrument, possess that mystery.

One vitally important component in the construction of a computer program is clarity of purpose. We are generally tolerant with works of scholarship which seek to comply with this clarity but fail wholly to satisfy. With machine-readable data, whether bibliographical or textual, ambiguity of purpose is unacceptable. And this applies as much to the interrogators of data as to those responsible for the architecture of the system in which the information is stored. When the machine-system in which the *Eighteenth Century Short Title Catalogue* resides had to be designed it was necessary, from the start, to hypothesize the kinds of uses which the file might be expected to serve. The bibliographical record had to be structured in such a way as to facilitate different kinds of search, and every facility provided by the retrieval software used to advantage. The result has been the creation of a structure which conforms quite closely to the standards required by those international bureaucrats who govern the cataloguing of books in libraries, and yet goes some way to satisfying the expectations of an academic community for whom the quibbles of cataloguers are as comprehensible as the odder notions of the metaphysical divines. It is a model upon which numerous projects throughout the world are building bibliographical files, and has therefore justified the enormous resources which have been devoted to its construction. But it is merely a beginning: a model built upon a technology already ten years old, rapidly being overtaken by developments in microelectronics which will completely transform the way we handle information. And if, as has been suggested, information is the resource upon which our society increasingly depends for its existence then we ignore the technology which sustains that resource at our peril.

There has never been any doubt about the role which enumerative bibliography has to play in supporting research in the humanities. A recognition of that role has always been implicit in the activities of the Bibliographical Society, which has been responsible, over a period of fifty years, for bringing to completion a comprehensive listing of the printed sources for the period up to 1640. The separation between the activities of bibliographers and those of cataloguers—a source of complaint in most universities—is discernible in one of the earliest lectures delivered to the Society in December 1892. It was by Henry Wheatley, and its subject was 'The present condition of English bibliography, and suggestions for the future'. One of the aims of the Society, Wheatley argued, should be the promotion of universal bibliography:

What appears to me to be an ideal state is that in which a complete and full Bibliography of all literature should be produced, so that Catalogues might be less full in their descriptions, and contain references to the Bibliography for fuller particulars. This ideal state is far off, and it may
never arrive, but if we agree in believing it to be a desirable state, a body of workers united as the Bibliographical Society may do much to help us forward on the road that leads to this ideal state. It does seem a sad waste of time that cataloguers should continue to repeat long descriptions and collations of valuable works which might with advantage be done once and for all in a recognized Bibliography.

Two catalogues describing fifteenth-century books have come close to Wheatley’s ideal, even though neither is as yet completed: the Gesamtkatalog der Wiegendrucke and the British Museum Catalogue of Books printed in the Fifteenth Century. It is common practice in producing catalogues of collections of incunabula to give the briefest description followed by references to GW and/or BMC. But Wheatley’s most important observation concerns the need to reduce redundant effort in cataloguing books. A similar realization that a ‘once and for all’ description should be available through a complex structure of national and international computer networks to libraries throughout the world, has led to the creation of huge bibliographical databases financed by libraries on a ‘pay as you use’ basis. The Eighteenth Century Short Title Catalogue seems to me an interesting development, for though it can claim to be in the tradition of short-title catalogues it in fact goes further than either of the short-title catalogues for English printing to 1700 in fullness of descriptions and annotations. And because it is in machine-readable form it is susceptible to instantaneous correction and modification. Every ESTC record is, in a very real sense, a tentative description which can be modified as contributions from libraries throughout the world are assimilated. In order to appreciate why ESTC occupies a unique position in the contemporary scene, distinct in important respects from the databases established for modern books, it is necessary to remember that one of the most important discoveries of bibliography in recent times has been the fallibility of the individual copy as a witness of the testimony available for the edition.

Some books, especially contemporary books, composed, printed, and bound entirely by mechanical processes lend themselves to formalized description and itemization. Uniformity of output is, after all, a desideratum of modern book-production, and contemporary convention has made possible the concept of the International Standard Book Number (ISBN), or unique identifier. It follows, therefore, that a methodical and formalized set of descriptive principles stands a fair chance of success in identifying, beyond reasonable doubt, precise and universally accepted details for a book produced by precise and universally accepted methods. But a sizeable proportion of the books consulted in research libraries were printed in the period before 1850, in the era of the hand-produced book, and however urgent the need to simplify the cataloguing of books flooding from the presses of the world, there remains a distinct and determined community whose interests focus on the past. Recognition that books produced by hand can manifest peculiarities not normally encountered in books produced by machines is now taken for granted by those concerned with the editing of literary texts. If bibliography is to serve scholarship it follows that a descriptive format, capable of revealing the significant features of a hand-printed book, has to be developed. ESTC has,
within limitations imposed by the available funding, succeeded in devising a format which
does minimum violence to the evidence yielded by the original work; is compatible with
the cataloguing standards that have emerged in the wake of computer cataloguing; and is
hostile to books of different periods and printed in different languages. The medium
in which the growing body of information is stored is electronic, and that makes it possible
to interrogate the file in ways which have never previously been possible with printed
catalogues.

The traditional grammar of research necessitates the creation of what I shall call
‘disjunctive indexes’: these commonly take the form of separate files of cards, or slips,
arranged in an order which is meaningful to the compiler. Many bibliographies, and some
sophisticated catalogues, are furnished with several such indexes. If the catalogue
enumerates a few hundred, or even a few thousand items, then the labour of compiling
and consulting a variety of indexes is manageable, but when catalogues exceed a certain
size, indexing becomes impractical, and the labour required to discover items conforming
to particular determining features out of the question. The great benefit to be derived
from access to a machine-readable catalogue, supported by imaginatively constructed
software, is that we can ask multiple questions. The complexity of the questions which can
be asked is determined by the structure of the bibliographical record. This is the precise
point at which the grammar of conventional cataloguing, conditioned by the appearance
of print on a page, has to be abandoned in favour of a grammar more closely associated
with Ramist logic.

A computer can index information in a variety of ways. In what is termed ‘keyword’
indexing every word is keyed to the fundamental referent—the anchor—which is a record
number. The difficulty with keyword indexing is that the computer is unable to make
semantic discriminations, since words are stored as strings of binary data. So a search on
the word ‘garden’ will certainly retrieve a substantial number of books on horticulture,
but it will also retrieve titles with other uses of the word. In ESTC a substantial number of
elements within the bibliographical record are indexed in this way. ‘Phrase’ indexing is
narrower and is generally used for particular sequences of letters—symbols for libraries,
for example. The narrowest (and incidentally the simplest) indexing is reserved for the
alpha-numeric code where a letter or number stands for a specific bibliographical feature.
Extended use of such codes might well lead to a system for detailed and sophisticated
description of numerous features of early printed books.

The development of bibliographical analysis during the past twenty-five years has led
to a recognition that many features of early printed books deserve particular notice in
methodical descriptions. Such features including binding (material, colour, construction,
finishing, ornament, provenance, date); paper (sheet size, quality, mould, watermark);
typography; layout; illustration. For books printed after 1800 we will have to take into
account method of printing and printing techniques used for illustrations (over a hundred
processes are known). There are many others that I shall pass over here. Why, it might be
asked, go to all that effort? The answer is simple. Print on the page (whether it is a
bibliography, a catalogue, or an edition of a text) is disposed in a manner which assumes
that the reader will wish to have certain questions answered, and the compiler or editor predicates his arrangement on those questions. Bibliographies, manually compiled, can answer some questions. In machine-readable form they are capable of answering questions which require imaginative effort to predict. Some of the questions which have been put to ESTC during the past few months will illustrate my point:

1. How many separate poems did Tonson print in folio/quarto/octavo?
2. Did the Dublin book trade suffer the same reduction in output between 1730 and 1750 as the book trade in London?
3. How many books published in the eighteenth century have the word 'consciousness' in the title?
4. How many poems printed between 1701 and 1760 claim to be written in imitation of Milton or Spencer?
5. Which travel books concerned with Scandinavia and Iceland were in the library of Sir Joseph Banks?

Unfortunately, ESTC has not found it possible to code the elements in a bibliographical record in such a way that all questions can be answered. This is because of the sheer number of items to be dealt with (a universe of some 400,000 separate items surviving in over two million copies), and the fact that copy-specific information submitted varies enormously from library to library. In some cases (records submitted from the Houghton Library at Harvard or the Humanities Research Center at the University of Texas, for example) the information supplied is meticulous and detailed; in others it is barely sufficient to enable editors to assign the copy to a particular edition. As a compromise between the elaborate bibliography and the traditional short-title catalogue, ESTC is no more than a first step in the development of bibliographical techniques in the computer age.

In order to indicate how comprehensively it is possible to retrieve combinations of features, I offer the following profile of a bibliographical record. It is divided into three sections: the first describes all features of the item which are common to the edition; the second, those features which are common to some (but not all) copies of the edition (usually described as variants); the third, those features which pertain to individual copies.

*Features common to all copies of the edition*
Title-page details; collation; imposition; illustration; dimensions; typography; paper; layout; signatures; running-titles; separate title-pages; press-figures; advertisements; etc.
Contents; chapter-headings; dedication; preface; etc.
Intellectual attributes; subject and generic classification; country of origin; language(s); etc.

*Features common to sub-families of the edition*
Paper; collation; press-figures; cancels; advertisements; etc.
Features of individual copies

Binding; provenance; manuscript notes or corrections; extra illustrations; etc.

Some of these will be indexed by keyword, but numerous features of printed books lend themselves to the simplest form of coding. Typographical features, for example, where a single-letter code can be assigned to various types of letter-forms and sizes. The same is true of layout and design. Bindings can be classified (admittedly on a rough and ready basis) according to material, method of construction, whether produced by hand or by machine, period, provenance, and finishing. For fifteenth-century books a vast amount of such information already exists in various catalogues, and if consolidated into one machine-readable file would make it possible for incunabulists to study more effectively the history of book production in its infant period.

But if the advent of the computer augurs an end to much of the drudgery associated with research, there will be some loss. I remember something Howard Nixon said to me many years ago, when he was Superintendent of the North Library in the British Museum: 'Don’t be in too much of a hurry: it is only when you have worked in about a hundred libraries, and handled a few thousand books, that you begin to understand the special language with which books communicate.' He approved, I know, of the fact that by the time I came to undertake ESTC I had already accumulated some experience of books and the uses to which they are put. Research, detached from those serendipitous processes associated with handling a wide variety of the ‘witnesses’ I referred to at the beginning, will probably have less excitement. It will certainly be less perilous. But can we really afford the time and energy to pursue knowledge as Poggio did, on the back of a mule from Italy to St Gallen, in order to examine a ‘rumoured’ manuscript of Aristotle?

We might do worse than remember the self-evident fact that research is a highly sophisticated form of play. When we play we observe, as sociologists are continually reminding us, a subtle obedience to a grammar which is as unmistakable as it is difficult to define. I am confident that the new technology will assist us in the revelation of the past. The extent of that revelation will depend, in large measure, on the contribution which all of us, whatever our preoccupations of the moment, can make with the tools to hand. Those tools have always been subject to evolution. Which is why I conclude as I began by urging the metaphor of the scholar as archaeologist. The Muse reveals what she will, and all is a gentle coaxing. No wonder we think her the stubborn shrew she is.

1 Little evidence is available for the research methods of the period up to the nineteenth century. One would give a great deal to know how scholars like Todd, Skeat, Furnivall, Bradshaw, for example, achieved the rare combination of perception and massive accomplishment. For more contemporary studies the lack of precision in determining the ‘grammar’ is evident in works such as J. Barzun and H. F. Graff, The Modern Researcher (New York, 1977); J. M. Morris and E. A. Elkins, Library Searching: Resources and Strategies (New York, 1978); M. S. Batts, The Bibliography of German Literature (Bern, 1978); P. Wilson, Two Kinds of Power (Berkeley, 1968).

2 The ‘Universal Catalogue Project’ was still being discussed in 1879, when its precarious history (since 1850) seemed at last close to resolution. In The Library Journal (iv, 1879, pp. 158–9) it was reported that evidence given by George Bullen to the Council of the Society of Arts indicated that
the first step, the printing of the catalogue of the British Museum Library, could be accomplished 'in a rough and ready way' within two years, though the task was not completed until 1900. It is a fact not generally known that the Society of Arts was responsible for producing the first specimen page of the proposed catalogue of the British Museum Library. For a brief account of the 'Universal Catalogue' see The Library Journal, iii (1878), pp. 61-2.

3 Bacon's Instauratio Magna (London, 1620) was intended as a statement, covering several disciplines, of the 'new science'. The Novum Organum is the fundamental source for much that has been credited to Bacon as the architect of modern industrial society. Victorian acknowledgement of Bacon as an apostolic hero of industrialism and capitalism can be seen from the numerous editions and translations of the Novum Organum published between 1840 and 1870.


See also A. Toffler, The Third Wave (London, 1980).

5 The classic debate on this is to be found in The Library, 4th ser. xiii (1932), pp. 225-58, between S. Gaselee and W. W. Greg, with A. W. Pollard's arbitration.


7 Transactions of the Bibliographical Society, i (1893), pp. 61-90.

8 The most obvious feature which ESTC does not include is provision of subject headings. However, intelligent use of keyword-searching can produce remarkable results. For a brief guide to searching ESTC see R. C. Alston, 'Searching ESTC Online', Factotum, Occasional Paper, i (1982).

9 For a lucid exposition of the tension which exists between the bibliographer and the cataloguer see G. T. Tanselle, 'Descriptive Bibliography and Library Cataloguing', Studies in Bibliography, xxx (1977), pp. 1-56.
Copyright of British Library Journal is the property of British Library Board 96. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder’s express written permission. However, users may print, download, or email articles for individual use.