

Improving OPAC Subject Access: The ADFA Experiment

Alex Byrne and Mary Micco

The Australian Defence Force Academy (ADFA) library has embarked on an ambitious project to upgrade subject access in its online public access catalog (OPAC) by adding an average of twenty-one multiword terms from the table of contents and/or index to the 653 field in the MARC record for each book. After reviewing six months of work involving some 6,000 books, this study attempts a preliminary evaluation of the impact on the workload and resources, as well as on subject access.



The Australian Defence Force Academy (ADFA), a campus of the University of New South Wales, has embarked on an ambitious project to improve subject access to their book collection, which supports teaching and research in a full range of academic subjects. A relatively small academic library of 160,000 volumes, it recently completed a total retrospective conversion and came online with the integrated Urica Library System. Urica is supplied by AWA in Australia and McDonnell Douglas in Europe and North America as a turnkey package of hardware and software. It is able to store and access all the fields available in the MARC format and is designed to permit subject access by controlled headings and keywords.

OPAC PROBLEMS

User surveys of existing OPACs have indicated that the major problems in database access were

1. finding the right subject heading;

2. ambiguous codes and abbreviations;
3. currency and coverage;
4. indexing and database accuracy;
5. updating.¹

The first problem also appears to be the major weakness of the Urica system. A keyword capability attempts to overcome the difficulty by increasing searching flexibility, but it has not proved as successful as had been hoped, for retrieval is handicapped by both data and software limitations. In common with most other library catalog systems used in Australia and North America, Urica restricts subject searching to the MARC 650 fields and hence to Library of Congress Subject Headings (LCSH) recorded in those fields. Therefore the keyword capability can use only the words taken from those headings. That is, subject searching is possible only by "direct hit" on the exact LC heading including subdivisions or by combinations of terms taken from that artificial vocabulary.

As programmed, the system first

Alex Byrne is Deputy Librarian, Australian Defence Force Academy Library, Canberra, ACT, Australia. Mary Micco is Associate Professor of Computer Science, Indiana University of Pennsylvania, Indiana, Pennsylvania 15705.

ful information, to integrate it into the text, and to cite sources clearly. For their section of the course, the bibliographic instruction librarians reinforce their explanations with carefully labeled handouts, and short hands-on exercises that relate to the major research project from the ESL class. To facilitate interaction further, the bibliographic instruction librarians have, with student help, compiled glossaries of library terms in the dominant languages of our international student population.

In conclusion, awareness of how ESL students differ from other students would

help librarians modify their approaches, leading to improved communication with this growing segment of the university population. A more comprehensive introduction to the library than is presently employed is a necessity for such students. An integrated program involving bibliographic instruction personnel with intensive English, English department, or "Writing across the Curriculum" programs can be an ideal way to help ESL students acquire the knowledge and skills essential for college success.

REFERENCES AND NOTES

1. Office of Educational Research and Improvement, Center for Statistics, *Digest of Education Statistics, 1985-86* (Washington, D.C.: U.S. Dept. of Education, 1986), p.218-19.
2. Kwasi Sarkodie-Mensah, "In the Words of a Foreigner," *Research Strategies* 4:30-31 (Winter 1986).
3. Mary Alice Ball and Molly Mahony, "Foreign Students, Libraries and Culture," *College & Research Libraries* 48:160-66 (Mar. 1987); Manuel D. Lopez, "Chinese Spoken Here: Foreign Language Library Orientation Tours," *College & Research Libraries News* 44:265 (Sept. 1983); Terry Ann Mood, "Foreign Students and the Academic Library," *RQ* 22:176-80 (Winter 1982); and Sally G. Wayman, "The International Student in the Academic Library," *Journal of Academic Librarianship* 9:336-41 (Jan. 1984).
4. One Lebanese student, for example, was accused of cheating on his history exam because he had duplicated from memory, word for word, comma for comma, five pages of his textbook. However, he was vindicated when the teacher challenged him to repeat the performance and he did so with ease. Student and teacher were obviously working under different sets of assumptions. The student declared memorizing at length was easier for him than trying to rephrase new ideas and information into language that might not be as clear.
5. Robert L. Kohls, ed., "Resource 8," *Developing Intercultural Awareness* (Washington, D.C.: Society for Intercultural Education, Training, and Research, 1981), p.18-22.
6. Sarkodie-Mensah, "In the Words," p.31.
7. Wendy Moorhead, "Ignorance was Our Excuse," *College & Research Libraries News* 47:585-87 (Oct. 1986).
8. This practice can lead to disastrous misunderstandings, as when French Quarter customers, asking a Cambodian waitress if they could use the rest rooms in a fast food restaurant were told, "yes, yes," but none existed.
9. Louise Greenfield, Susan Johnston, and Karen Williams, "Educating the World: Training Library Staff to Communicate Effectively with International Students," *Journal of Academic Librarianship* 12:227-31 (Sept. 1986).
10. One might occasionally find students capable of the most amazing esoterica; for example, the Japanese western enthusiast who cries out, "Come back, Shane!" or the Arab lawyer who can recite the issues involved in the Goetz case.

searches for a direct exact subject heading match using a simple hashing algorithm. A keyword search is then initiated automatically if no direct match is found or at the user's request. The keywords are combined by an implied AND. If too many hits are retrieved, additional terms or truncated terms may be AND'ed to narrow retrieval. However, successful retrieval requires all terms to be in the same 650 field; searching across headings is not possible at this juncture. If many hits are found, an AND strategy can be used to combine additional terms (as character strings) from other 650 fields or most other data fields. Free-text searching is not available.

"Several methods were considered before the subject access approach method advocated by Pauline Atherton was chosen."

PRELIMINARY ANALYSIS

Frustration with these limitations led the ADFA librarian, Lynn Hard, and his staff to consider means of improving the subject data and the software capabilities. As it was decided to remain compatible with the standard Urica system for the present, software changes were minimized and efforts were concentrated on augmenting the subject data. Several methods were considered before the subject access approach method advocated by Pauline Atherton was chosen.² This method consists of selecting supplementary terms from the contents pages of books and, where needed, the index. ADFA changes were accomplished by selecting an average of twenty to twenty-five keywords from the chapter-level headings to be keyed into the 653 field for each book. There was some concern that not all books included the needed information but this proved to be a minor difficulty. The background study performed on a sample of 179 items showed that 104 items had both contents page and index, 52 had contents only, 3 had index only, and 20 had neither. Of the latter, twelve were fiction, four bibliographies, three in-

dices, and one a dictionary.

Before embarking on the actual processing, staff faced three questions:

Would the augmented headings be used separately or in combination with LC headings? A small trial with twenty-two queries confirmed that the augmented terms would be most useful if available for searching as an alternative to the LC headings. The LC terms would be regarded as the main subject of the book and the additional terms as "contents" terms. Care was taken to reduce redundancy by entering each term only once per document.

How would the augmented headings be structured? The best solution appeared to be the entry of all headings in a single field, the 653 field for local subjects. The semantic relationships between the terms were preserved. Separate headings were delimited with "@" with subheadings separated by means of a semicolon, punctuation that would not distort the retrieval process.

Would any system of weighting be employed? It was considered but rejected as possibly increasing the complexity of indexing by requiring value judgments. It was also felt that no reliable method was possible within the constraints of the current software.³

The inquiry software, particular ENQ, which is used by the public, needed two changes to the display options. The first was labeled "enhanced full" and included both LCSH and augmented headings. The other was an "enhanced MARC" option.

After six months, more than 6,000 books have been processed with enhanced subject access terms. This paper measures and describes the impact.

ENHANCED SUBJECT PROJECT

Although exact titles and title words could be searched in standard Urica, it was apparent that loading that data into a subject field would enhance subject access.⁴ Consequently, the first step in the project was to load every record's title into the 653 field with a \$t tag to indicate it was a title. This was accomplished in a batch run using software developed in house. Drawing on the experiment in providing

enhanced subject access at Syracuse University by Pauline Cochrane and Barbara Settel, the next setup, begun in December 1986, was to add 653 fields containing additional terms chosen from the book's contents page or index, to the MARC record.⁵ The project was entitled the Enhance Subject Project (ESP), and the staff began referring to the additional terms selected as ESP terms. At the time of writing, after seven months of adding enhanced subject terms to all new acquisitions, the library has a collection of 6,139 books, more than 4 percent of the database, that have received this treatment. Library patrons have full keyword access to these terms, which are stored in the 653 fields with special tags to indicate up to four levels of subdivision. Standard *Urica* searching software is used at present but the library's systems staff developed input and maintenance software for the 653 field.

There were several reasons for selecting this as the best and most cost-effective method of enhancing subject access. It can be accomplished easily in a few minutes. It adds many access points and increases the likelihood of retrieving an item. Users have commented that it helps them to judge the relevance of the retrieved items. It uses the terminology of the author, ensuring currency and relevance, and also provides access to the chapter-level terms at acceptable costs. Minimal skills and intellectual effort are required, and the time for cataloging is approximately fifteen minutes per item, yielding an average number of 20.7 headings or 53.6 terms per book.⁶

IMPLEMENTATION DETAILS

A manual outlining the selection procedures to be used was drafted, based on Barbara Settel's work with a few minor adjustments. The first step in the process of adding terms is to determine whether the table of contents is usable and, if not, to select suitable terms from the index. Guidelines for difficult cases as well as many examples are provided. The catalogers work from the photocopies of the title page, table of contents, and index. They highlight the terms that they select and add the appropriate level codes. These sheets are

then sent on for keying into the system. Terms are selected that describe the contents of each substantial chapter, except where they are significantly repetitious or introductory. The table of contents must contain at least one heading for every thirty-five pages or the index is used to provide additional terms. Page number ranges and names of authors of contributed articles (except for poems and short stories) are omitted.

IMPACT OF PROJECT

Numerous studies on subject access have shown that LCSH performs poorly. Since it provides only 1.7 access points per book,⁷ this is hardly surprising. It was assumed that providing table of contents information would significantly increase recall and improve precision of searches without adding unduly to the cost of processing and storing the information. The entries would provide naturally occurring hierarchies of terms that could be displayed, preserving the syntactical structure of the contents or index pages, to assist users in refining their search strategies and evaluating the relevance of retrieved citations. The major question is whether the terms added are useful or simply meaningless clutter. In this study, we have attempted to view the impact of the Enhanced Subject Project (ESP) in both its quantitative and qualitative aspects.

IMPACT ON WORKLOAD AND RESOURCES

The first consideration was the impact of this additional load on personnel and the system. The cataloging librarian reports that it takes the equivalent of one full-time entry-level clerical person to photocopy the contents pages and key in the selected data. Most selecting and coding of terms for the 1,000 new items a month are performed by one paraprofessional. The cataloging section has absorbed the additional work with little difficulty. Relieved of former duties by the replacement of a totally manual system with a fully automated one, the section has taken on this activity as well as continuous upgrading of descriptive cataloging data while losing only one clerk and experiencing a 25 per-

cent increase in acquisitions.

The systems manager reports that additional disc storage requirements have not been as great as anticipated. Before copying the titles so that they could be searched together with LCSH and ESP terms, the bibliographic database occupied 244 megabytes (MB) of storage (of a total of 350 MB in use). Copying the titles into the 653 fields caused an increase of 14 percent, to 279 MB. By August, after eight months of ESP, during which a total of 8,500 catalog records including 7,000 with ESP were added, there were minimal effects on all files except the file containing the subject/title/ESP keywords. That file increased from 1.8 MB to 5.5 MB when titles were added, and then to 7.2 MB over the eight-month period. The increase attributable to ESP alone was approximately 2.5 MB. The number of keywords increased from 30,153 to 71,338 with the addition of the titles, and then to 101,559 (see

table 1). Some 30,000 words in the past eight months are undoubtedly due to ESP. While steep at first, this rate of growth is now declining.

By separating the 6,139 titles that had been ESP'ed into a distinct database we were able to evaluate the growth of the vocabulary. During this time a total of 51,175 new words were added to the system. It was initially expected that growth would be very rapid until it reached 50,000 and then would level off, topping out at around 100,000. A sufficiently large database is not yet available to verify this opinion but it appears unwarranted and will be reevaluated after twelve months. Proper names and jargon are included in the vocabulary but there are few non-English-language terms. The distribution of the vocabulary was interesting in that a very large number of words, 89 percent, had fewer than ten occurrences, and 58 percent had only one hit (see figure 1).

TABLE 1
IMPACT OF ESP ON VOCABULARY GROWTH

	New Keywords Added	Records Added	Ratio
January	4,977	448	11.0:1
February	4,247	648	6.6:1
March	4,184	698	5.9:1
April	3,129	707	4.1:1
May	6,676	802	8.3:1*
June	5,717	1,452	3.9:1

*Additional specialist military history materials E.S.P.'ed during these two months as part of a special project (Australasian Military History Database) temporarily interrupted the steady decline.

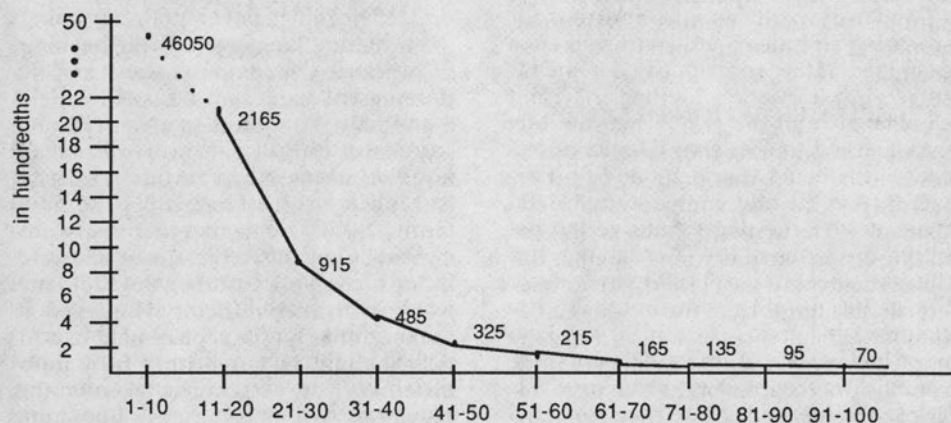


FIGURE 1
Distribution of Vocabulary-word frequency counts

It is not yet clear whether this pattern is a result of having such a small database and will hold as the database gets larger. It certainly has implications for searching. By putting in the lower frequency terms first, the user may avoid searching for high frequency words at all or may do so only after first culling a small pool of relevant documents. It will, however, be very important to provide an "OR" capability to ensure that a sufficiently large pool of documents is retrieved.

The vocabulary lists demonstrate again that the number of individual keywords generated from the LC headings is very restricted. We need a mechanism to map the natural-language terms put in by users to this very controlled vocabulary.

IMPACT ON INFORMATION RETRIEVAL

Another interesting question was the impact of the additional access points on information retrieval. In a quantitative analysis we selected 100 significant keywords in computer science identified by the faculty, and an additional 100 keywords in a stratified random sampling of the entire database. (Three or four keywords were randomly chosen from each letter of the alphabet, excluding the words with only one hit and those previously selected for the computer science segment.) The focus was on the extent to which the ESP terms are increasing recall in the system as opposed to adding clutter.

Each keyword was submitted for searching and the records retrieved were examined. If the word appeared in the LC subject heading, it was counted as LC and eliminated from the pool. Then the title was checked to see what additional records were being found by adding keywords from the title. In this study we did not measure the degree of overlap between the terms in the LC headings and the title, since our purpose was to evaluate the material being contributed by the chapter-level terms. Finally, if the keyword had not been found in either of these places, it was counted as a record that was being recovered by the chapter-level headings. All of this was made possible by the use of tags to differentiate the ele-

ments being stored in the 650 and 653 fields and by the programming skills of the systems analysts. See figure 2.

The results clearly demonstrate that the ESP terms are retrieving an average of 72 percent of the documents for terms in computer science and, over the whole database, an average of 75 percent.

"The terms drawn from the table of contents increased retrieval 300 percent."

The title contributed little more than 10 percent, while the LC subject headings retrieved only 15 percent of the available documents, even though this research procedure favors LCSH (see table 2). This means that effectively the terms drawn from the table of contents increased retrieval 300 percent.

ANALYZING THE QUALITY OF THE INFORMATION RETRIEVAL

Having determined, without a doubt, that the additional access points have increased document recall, the next question is how many irrelevant documents or false drops have also been retrieved. In a first effort to analyze the impact of the additional headings on precision, computer science terms were searched as pre-coordinated headings or compound headings (e.g., natural language). These headings were selected by computer science faculty to represent significant areas of the field. Admittedly computer science terminology is very difficult to keep up with since it is rapidly changing. In the future we hope to replicate the study with compound terms selected from other more stable disciplines. Only thirteen of the headings selected turned out to correspond to LC subject headings, while eighteen did not. In other words, for 58 percent of the terms picked by the faculty there was no direct match on a heading. Records were keyword searched under these headings and once again we first counted the records retrieved under the LC heading, then those

COMPARISON OF RECALL

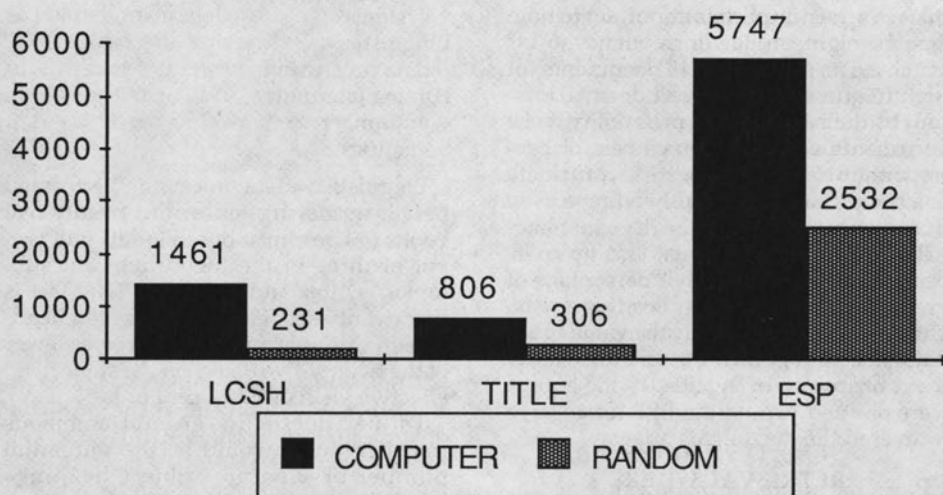


FIGURE 2
Comparison of Recall

TABLE 2

BOOK NUMBERS RETRIEVED BY NATURAL-LANGUAGE SEARCH STRATEGIES

Strategies	Book Numbers*
Artificial intelligence	1, 2
Computer-aided design	3
Expert systems (computer science)	1, 14
Grammar, comparative and general—syntax	8, 21, 23
Interactive computer systems	11, 13, 16
Language and languages—congresses	4, 22
Linguistics—data processing	5, 8, 9, 10, 11, 15, 16, 17, 21, 25
Linguistics—congresses	6, 19
Machine theory	7
Parsing (Computer grammar)	5, 12, 19, 21
Question-answering systems	16, 18, 20
Semantics	16, 18, 24

*The books were numbered 1-25.

recovered by the title. The remainder were attributed to ESP headings. Faculty were asked to judge the relevance of the retrieved articles to the topic as they understood it. However, they were told to interpret relevance in its broadest sense, that is, articles that could be of interest to someone working on that general topic as it relates to the discipline of computer science.

When an LC heading exists there were only 12.5 percent irrelevant documents, almost half of which are attributable to the LC heading. Surprisingly, the title and ESP terms combined contributed only 16.8 percent to the relevant pool here when the terms have been precoordinated. This appears to vindicate the librarians who feel that LCSH does provide an invaluable grouping function. This is only

half the story, however. The LC headings that exist in computer science are of a very general nature, as can be seen from the chart in appendix A. It is important to note that the eighteen headings where no LC match exists recovered 318 documents, of which 220 or 69 percent were deemed relevant by the faculty. In all probability these documents would not have been recovered otherwise. Title appears to contribute much more where no LC heading exists, that is, only 6 percent of the relevant material when a heading exists, and up to 49 percent (as well as a higher percentage of irrelevant hits) where no heading exists. The study indicates clearly the value of additional access points, but failed to clarify the contribution of the title. If title entries were omitted, would the ESP terms have recovered the documents anyway?

RETRIEVAL WHERE NO LCSH HEADING EXISTS

Another issue is what happens when no LCSH heading exists, as in the case of *natural language*, the term chosen for *more detailed study*. Thirty books were retrieved using a keyword search for *natural* and *language*. In this group, five (17 percent) were considered irrelevant.

In this small collection of twenty-five relevant books we looked for the possibility of grouping by call number but found twelve different call numbers spread across three schedules, as shown:

Science	Q335 QA76.7 QA76.9 (5 books)
Literature and Language	P23 P25 P98 P123 P302 P325 PE1074.5
Technology	TA174

A similar scattering occurred under the LCSH headings. There were twenty-seven different subject headings, with a total of fifty-five keywords, giving a book average of 2.2, a slightly higher average yield than has been reported in other studies. Certainly plenty of access points are provided. What is alarming is that no one heading provided any sort of useful grouping function for retrieval purposes. The headings that retrieved three or more books in this group were as follows:

Grammar, comparative and general—syntax	4
Interactive computer systems	3
Linguistics—data processing	10
Parsing (computer grammar)	3
Semantics	3

Linguistics—data processing performed best in retrieving ten of the twenty-five books (40 percent), but failed to pull two out of three in the highly relevant sub-topics, syntax and semantics. The fact is that twenty-six of the twenty-seven headings in this area will locate three or fewer of the twenty-five books on the subject, an effective yield of 12 percent or less.

The next question is, given an ambitious searcher, what would be the minimum number of separate subject headings needed to retrieve all twenty-five documents? It took twelve different strategies (see table 2).

The most alarming point is that a keyword search restricted to LC headings with the words *natural* and *language* will yield no hits even though natural-language processing is a well-recognized topic within artificial intelligence.

If we follow the "snowball" strategy discussed in the literature (i.e., drawing on subject headings provided by those retrieved books deemed relevant), it is clear that the reader would recall vast amounts of irrelevant literature (see table 3). In all, eight of the twenty-seven headings can be definitely considered relevant and six marginal, while thirteen can be rejected as virtually useless because they are too broad.

IMPACT ON USERS

When the ESP procedure had been operational for six months, a structured interview survey of sixty-seven users showed that 38 percent had found the enhanced terms useful, 19 percent had not, and 43 percent were not sure or were unaware of the facility. The users, a stratified random sample of those registered with the library, volunteered the following reasons: 41 percent had no need to use the

ESP CONTRIBUTION TO RECALL

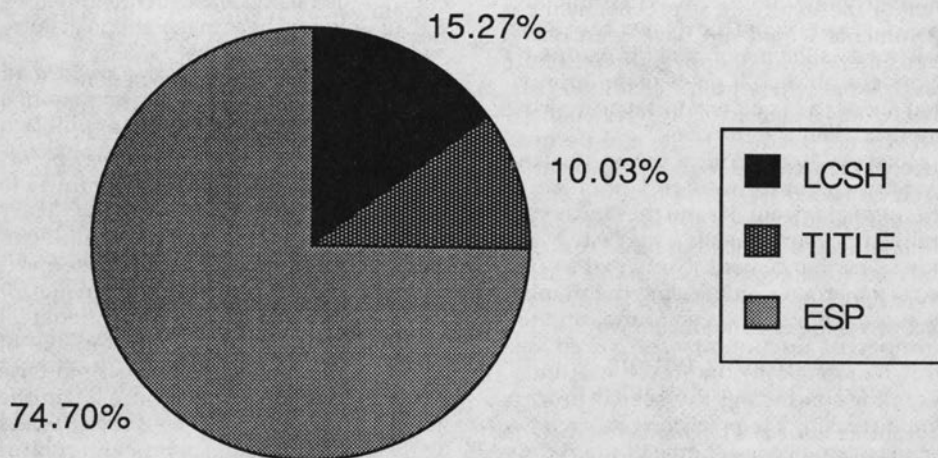


FIGURE 3
Extent to which ESP Increased Document Recovery

TABLE 3
RANKING OF LCSH HEADINGS AVAILABLE

High Relevance	Relevant	Marginal	Useless
Ambiguity Anaphora		Artificial Intelligence	
Discourse analysis	Cleopatra		Electronic digital computers—programming
Generative grammar	Grammar—Comparative and general —Syntax		Engineering design
Parsing—Computer grammar		Linguistics—Data processing	English language—Data processing
Question-answering systems		Logic—Symbolic and mathematical	Interactive computer systems
Machine translation		Machine theory	Language and languages Languages—philosophy Programming languages—Electronic computers—semantics Visual communication
Semantics			

terms, for seven percent they facilitated access to more material, but for 52 percent they gave more information about individual books. Besides the clear implication that more publicity and reader education are needed, the responses indicate that the improvement in recall was taken for granted while the increased information available to evaluate the relevance of items was seen as the major benefit. As one respondent put it: "It gives more information about the book which saves me having to actually get the book."

However, some critical comments suggested areas for software enhancement, in the areas of synonyms and thesaurus control. Truncation was also suggested as a possible enhancement. Now, that a satisfactory method of increasing the number of access points has been found, and the number of documents recovered has grown dramatically, improved user interfaces for conducting the search process and increasing the precision of search vocabulary are needed. Some of the recommendations that emerged from this study were as follows:

1. It is important to provide keyword searching in LC headings, also to display scope notes to indicate where they differ. It is not always clear to the uninformed why a distinction is made between information retrieval (the activity) and information storage and retrieval systems (the software that permits the activity).

2. Since the number of records retrieved even in this small database of 160,000 titles is unmanageable, further study is needed of methods of assisting users to narrow their searches once they have found the topic of interest. In this way they can focus on the books that best address their information needs.

CONCLUSION

It is important to clarify the issues in-

involved in providing subject access to books. For years the number of subject headings, or access points, was restricted by the size of the catalog card and a tradition that demanded, for economy of card catalog maintenance only, that subject access should only be provided for the main topic of the book. Consequently, an average of only two headings per book is provided even today.

For periodical articles, experience with providing keyword access has shown that recall and user satisfaction are increased by providing thirty to fifty access points per document. The cost has not proven to be prohibitive.

The Enhanced Subject Project has demonstrated that the use of contents terms is a viable and cost-effective technique for dramatically increasing the number of subject access points to the contents of books without a serious increase in false drops. It has also shown once again the failure to provide adequate subject access of the accepted system of providing one or two controlled headings to each book. The controlled headings do help to provide a grouping function and should not be abandoned but rather enhanced, as suggested here. Much further work is needed on the vocabulary distribution patterns we identified and the implications for searching.

The ADFA library continues to input subject access terms, thereby providing an increasing pool of research data. It is hoped that additional studies will be implemented as more data becomes available.

The research reported here is the first step of a larger project to develop an expert system that can handle authority control by mapping the terms selected by the users to the very restricted vocabulary of LCSH.

REFERENCES

1. *Books Are for Use, Final Report to the Council on Library Resources, Syracuse University, School of Information Studies, Subject Access Project*, (Syracuse, N.Y.: Syracuse Univ., 1978), p.177. (ED 156 131).
2. *Ibid.*
3. Sue Beatty, "An Experiment in Enhanced Subject Access," *Cataloging Australia* 11, no.4:86-95

- (Dec. 1985); Alex Byrne, "Life Wasn't Mean to Be Whimsical: Painless Subject Augmentation," *Australasian College Libraries* 4, no.2:83-90 (June 1986).
4. AUSMARC, the Australian MARC specification, which is used in the academy library, has no 653 field. The 650 field is used for topical subject headings, 600 and 610 for name headings, while 651 contains geographic subject headings with data, according to AACR2 and LCSH. However, the National Library of Australia advised that the 653 field would probably be specified for data similar to that generated by ESP and that its use would be consistent with LCMARC. American practice would dictate the use of the 690 field for local headings.
 5. Barbara Settel, ed., *Subject Description of Books: A Manual of Procedures for Augmenting Subject Descriptions in Library Catalogs* (Syracuse, N.Y.: Syracuse Univ., School of Information Studies, Research Study #3, 1977).
 6. *Ibid.*
 7. H. Mary Micco, *An Exploratory Study of Three Subject Access Systems in Medicine: LCSH, MeSH, PRECIS* (Ph.D. diss., Univ. of Pittsburgh, 1980).

APPENDIX A: ANALYSIS OF RETRIEVAL

Comparison of Precision of LCSH, Title, ESP

Search Terms	LCSH	Relevant Title	ESP	LCSH	Irrelevant Title	ESP
LCSH Exists						
Comprehension	4	1	2	8	7	3
Computer interfaces	17	3	5		2	1
Critical path	35	0	2			
Expert systems	58	2	20			3
Generative grammar	20			6		1
Information retrieval	222	9	8			3
Machine learning	8	2	4		1	3
Machine translation	4		3			1
Parsing (computer grammar)	8	1	7			1
Pattern recognition	56	15	12		1	2
Question answering systems	4		2	1	2	1
Semantics	55	9	12	30	2	9
Speech recognition	9	4	6		2	3
Total	500	46	85	45	17	31
No LCSH						
Automatic classification		1	4			2
Computational linguistics		1				
Computer communication		9	3		10	22
Computer memory		3	12			
Computer speech		4	8			3
Computer translation			1		1	2
Decision support systems*	1	11	13			4
Human interface			9			
Intelligent tutoring			9			
Laser disk						2
Natural language		24	1		2	3
Optical storage		1	3			2
Query †	1	1	9		1	
Risk analysis ‡	3	8	22			
Software engineering		43	7		2	7
Text analysis		3	3		21	3
Text recognition			3			5
Video systems			6	3		3
Total	5	110	105	3	37	58

*The LC match was Decision making—Data processing.

†The LC term was Query languages.

‡The LC terms were Risk, risk management.