Subject Searching Using Controlled Vocabulary Versus Uncontrolled Vocaburary in Online Catalog System: Focusing on Multilingual Environment

Heeyoon Choi*
(최 희 윤)

ABSTRACT

The purpose of this paper is to investigate search efficiency of controlled vocabulary versus uncontrolled vocabulary subject access in online catalog systems. The question of the effectiveness of controlled versus uncontrolled vocabulary in information retrieval has been raised in many literatures. A debate continues in the Library and Information Science Professions over the relative merit, appropriateness, and efficiency of uncontrolled vocabulary subject access in online catalog systems. Actually users used to combine uncontrolled vocabulary subject searching with controlled vocabulary subject searching. But the success of user's subject search depends on his choice of search terms. Also the technical developments that facilitate cooperation among information services in general make it increasingly possible for such cooperation to take place on an international level. In this study, several common types of vocabularies on online catalog systems are described and compared, especially usages of vocabularies in multilingual environment are analyzed.

KEYWORDS

Online Catalog System. Controlled Vocabulalry, Uncontrolled Vocabulary,Natural Language, Subject Searching, Bultilingual Environment, Classification, Multilingual Thesaurus, Category Code, Descriptor, Natural Language, Subject Heading, Hybrid System, Post－Controlled Vocabulary

* 포스코경영연구소 정보자료실장
Chief of Information Resources Center, POSCO Research Institute(POSRI)

초 록

정보검색의 성공여부는 적절한 검색어의 선정에 달려있다고 해도 과언이 아니다. 특히 CD-ROM이나 온라인목록시스템에서 주제검색시 분류표, 주제명과 같은 통제어와 자연어 등 비통제어의 비교우위 문제는 아직도 논란이 되고 있다. 본고는 두 시스템의 검색효율성을 비교하기 위하여 특히 다언어 환경하에서의 검색어 사용에 중점을 두고 다양한 형태의 검색어휘를 조사 분석하였다.

키 워 드

정보검색, 온라인목록, 통제어, 비통제어, 자연어, 다언어 환경, 분류표, 주제명, 디스크립터, 다언어 디소러스

I . Introduction

Effective uses of vocabularies in online catalog can achieve optimum retrieval for users. So, the question of the effectiveness of controlled versus uncontrolled vocabulary in information retrieval has been raised in many literatures. A debate continues in the Library and Information Science professions over the relative merit, appropriateness, and efficiency of uncontrolled vocabulary subject access in online catalog systems. Actually many research results indicate that using keyword searches for uncontrolled vocabulary subject access is not primarily an option of last resort. Users of the online catalog used to combine uncontrolled vocabulary subject searching with controlled vocabulary subject searching. But the success of user's subject search, whether it is a controlled vocabulary search, uncontrolled vocabulary search, or combination of the two, depends on his choice of search terms. Also the technical developments that facilitate cooperation among information services in general make it increasingly possible for such cooperation to take place on an international level. Thus, the ability to handle publications in many languages and to accommodate several languages in the design of information service is becoming more important as truly multilingual systems become operational.

So the purpose of this paper is to investigate search efficiency of controlled

vocabulary versus uncontrolled vocabulary subject access in online catalog systems. Therefore, several common types of vocabularies in online catalog systems are described and compared. Especially, usages of vocabularies in multilingual online catalogs are analyzed. As a result, alternative solutions on vocabulary uses for efficient subject searching in multilingual environment are suggested.

Ⅱ. Description of Subject Vocabularies Used in Online Catalog

〈Figure 1〉 lists seven major types of subject description that can be found in online and CD－ROM databases today.

〈Figure 1〉 Array of Subject Description from Generic to Specific

Hierarchical Classification
Category codes
Subject headings
Descriptors
Faceted classification
Post－controlled vocabulary
Natural language

These types are arrayed from broad to specific, that is, typical terms or categories in types that appear higher in the list are broader, more general, than typical terms or categories in types that are lower in the list(Bates 1988, 135). The top six types of vocabulary in the list can be considered forms of controlled vocabulary, while the last, natural language, is uncontrolled. Effective use of these vocabularies requires a strategic understanding of which types of classification and indexing are involved, and taking advantage of the particular mix of vocabularies in given database to achieve optimum retrieval. So, I will describe and explain these types of subject vocabularies in databases.

1. Controlled Vocabulary

(1) Classification

The oldest types of controlled subject vocabulary by far is classification scheme. A classification is a kind of controlled subject vocabulary which arranges topics according to how they are logically related than they are verbally named. The resulting arrangement is expressed as a notation, a sequence of numeric, alphabetic, or alphanumeric codes which makes it possible to arrange the classified sequence as easily as one can file subject access points in word form(Hagler 1991, 185). Even though online catalogs are becoming the main bibliographic tool, classification continues to function as a shelf location device. With the capability of online searching by class number or call number, the role of classification as a bibliographic tool has begun to re-emerge. Recognizing the potentials of online classification access, scholars and researchers in the field have begun making feasibility and retrieval effectiveness studies of classification as an online tool. One of the more interesting, is the study conducted by Markey under the joint auspices of OCLC, Forest Press, and the Council on Library Resources(Markey 1987, 37-68). The purpose of the study was to explore the effectiveness of the Dewey Decimal Classification as a searcher's tool for online subject access. Two test catalogs were prepared for the project: one offered access to subject heading, keywords in titles, series, and notes and call number; the other offered access through various of DDC in addition to the keyword approach. Two categories of user carried out searches both catalogs:regular library users who pursued their own queries, and librarian who searches on assigned topics. Test results of the DDC enhanced test catalog were revealing, particularly in terms of how many unique subject terms were generated. The study as a whole indicates that incorporating a classification scheme into the online catalog can provide avenues of subject access that are not possible through the alphabetical approach. It suggests three areas in which classification schemes could affect online searching: classification schemes as enhanced vocabulary, online subject browsing, and class number searching

(Gorman 1990, 80).

In these days despite new text retrieval packages with loud claims of novelty, and despite the availability of superhighways, the old fundamental techniques of classification and language handling are increasingly needed and no compact answer has yet emerged(Vickery 1994, 113).

(2) Subject Headings

Jewett and Cutter were instrumental in introducing the vocabulary to index libraries' bibliographic records as follows;

- to replace the use of rotated or permuted natural language words for subject retrieval
- to supplement the only controlled vocabulary subject approach then prevalent, the classified catalog

They carried over some aspects of those approaches, a fact quite evident in the oldest of the vocabularies still in use, the 'Library of Congress Subject Headings(LCSH)'. So subject headings date back to the nineteenth century and have been the traditional form of subject description used by the Library of Congress and by academic libraries for books. 'Wilson Indexes'. as well as 'Engineering Index' use this approach as well. Subject heading often contains main headings followed by one or more subdivisions. Subject headings are precoordinate, that is, all subject elements are combined into a single long heading by the indexer. The idea is to describe the whole document in that one headings or more(Bates 1988, 47).

(3) Descriptors

Post－coordinate searching caused a different of vocabulary to come into existence in the 1950s. Descriptors got their big boost after World War II, when science information specialists realized that subject headings were insufficiently detailed to describe highly specific scientific articles and reports. They experimented with a variety of a different forms of manual access, with names like ' peek－a－boo cards' and 'edge－notched cards.' Quite a number of indexing

theories were involved, and some descriptors were quite broad, while others were very specific. Individual descriptors were intended to describe a single concept with in a document, rather than the whole document—hence the phrase ' concept indexing.' Concept indexing is usually more specific, with 5—25 terms assigned to individual documents. Descriptor systems apply far more terms as a rule than subject heading systems do, identifying numerous individual concepts, rather than describing the whole document in one or two headings(Bates 1988, 47).

(4) Category Codes

Category Code is the label that have some classificatory features and some index term features and can therefore be considered mixed type. BIOSIS 'Concept Codes' shown in (Figure 2) is example of this type. BIOSIS Concept Codes are a mixed type, because,on the one hand, the concepts are carefully defined and given code numbers, rather in the manner of classification categories, but on the other hand, while grouped logically, the concepts are not embedded in a highly defined structure in the manner of most classification schemes.

〈Figure 2〉 Concept Codes from BIOSIS

These are two of the ten codes within the broader area of phytopathology or plant disease. Note the extensive scope notes and carefully worked out strategy recommendations. Though these "concepts" are formed like index terms. they also have classificatory features in that their coverage is made rigorously mutually exclusive and jointly exhaustive.

CC54512 NONPARASITIC DISEASES

CODE FREQUENCIES Major(7760) Minor(2040)

CODE APPLICATIONS This code retrieves studies on plant diseases caused by non-biological factors, such as weather conditions and pollution.

EXAMPLES—STUDIES ON • Pollution • weather • plant nutrient deficiencies or toxicities • toxic effects of pesticides on plants • soil waterlogging • genetic and developmental disorders

STRATEGY RECOMMENDATIONS

• For studies on genetic disorders of plants, use this code with appropriate keywords and the Plant Genetics code CC03504.

- For studies of the adverse effects of radiation on plants, use this code with appropriate keywords and the Plant Physiology, Biochenistry and Biophysics Radiation Effects code CC51516.
- For studies on plant diseases caused by climatic factors, use this code with appropriate keywords and the Bioclimatology and Biomenteorology code CC07504.
- For studies of physiological factors relevant to plant disease development, use this code with appropriate keywords and relevant Plant Physiology, Biochemistry and Biophysics codes. See Directory, page B－15.
- For studies on the effects of environmental pollutants on plants, use relevant Phytopathology codes and the Air, Water and Soil Pollution code CC37015.

CC54514 PARASITISM AND RESISTANCE

CODE FREQUENCIES Major(7950) Minor(570)

CODE APPLICATIONS This code retrieves studies on plant host susceptibility and resistance to diseases.

EXAMPLES－STUDIES ON • Breeding for resistance • genetic, physiological and morphological factors relating to plant resistance

STRATEGY RECOMMENDATIONS

- For studies relating plant genetics to disease resistance, use this code with appropriate keywords and the Plant Genctics and Cytogenetics code CC03504.
- For studies on plant resistance to entomological pests, use this code with appropriate keywords and relevant Economic Entomology codes. See Directory, page B－7.
- For studies on plant resistance to fungi, algae, bacteria, and other agents, use this code with appropriate keywords and relevant phytopathology codes. See Directory, page B－14.

2. Post－controlled Vocabulary

A disadvantage of typical controlled vocabulary lists, however, is that they are rather inflexible with respect to new terms and topic areas when they appear in a rapidly changing discipline. New vocabulary cannot be included until it has been evaluated into a new edition of the thesaurus. The earliest system developed to search large bodies of legal text(at the University of Pittsburgh Health Law Center) used a kind of thesaurus to aid the search process. This was merely a compilation of words with similar meanings, resembling 'Roget's Thesaurus' more than the thesaurus structure commonly used in information retrieval. Even without any significant degree

could be an extremely valuable searching aid;words with similar meanings are potentially substitutable in a search, and such a tool relieves individual searchers from having to think of all the words that might express a particular idea. Investing in the construction of such a searching aid allows significant economies in a system in which large numbers of searchers are performed. This simple type of thesaurus in a kind of controlled vocabulary, with the control applied at output rather than input. It is a post-controlled vocabulary(Lancaster 1986, 168). A post-controlled vocabulary has some of the advantages of natural language and some of those of controlled vocabulary. Typically, with a post-controlled vocabulary, indexing is not limited to an established list. Rather, natural language terminology is permitted as it appears in new incoming documents, but indexers then do something to that vocabulary to assist searchers. For example, they may create lists of closely related terms, or 'hedges,' so that the searcher covers a topic well without having to think up a dozen other related terms to ensure coverage. The BIOSIS 'Keywords' are an excellent examples of a post-controlled vocabulary. These terms are drawn from the natural language of the titles of documents indexed in BIOSIS. Frequently appearing terms are listed as 'Keywords' in the Master Index of the'Search Guide-BIOSIS Previews Edition.' The presence of these terms in this index will give the searcher ideas, but any natural language term, whether appearing in the Master Index or not, may appear in document titles and consequently constitute a searchable keyword. Controlled terms for a given topic('Concept Codes' and 'Biosystematic Codes') as well as term frequencies are listed next to the keyword so that the searcher may decide which is best to use in particular instances. In addition, BIOSIS indexers add some'controlled keywords' to ensure complete coverage. Because the post-control done by the indexer adds some consistency, subject searching on this vocabulary can be expected to be more effective, as a rule, than pure natural language searching(Bates 1988, 46).

A number of controlled experiments have consistently shown that natural language searching can outperform controlled vocabulary. The natural language

systems used in these experiments had only minimal searching aids. A post－controlled vocabulary system can offer all the advantages of natural language with many of the attributes of the pre－controlled vocabulary. In principle, there is no reason why such tools could not be made multilingual, for use in international networks(Lancaster 1986, 169).

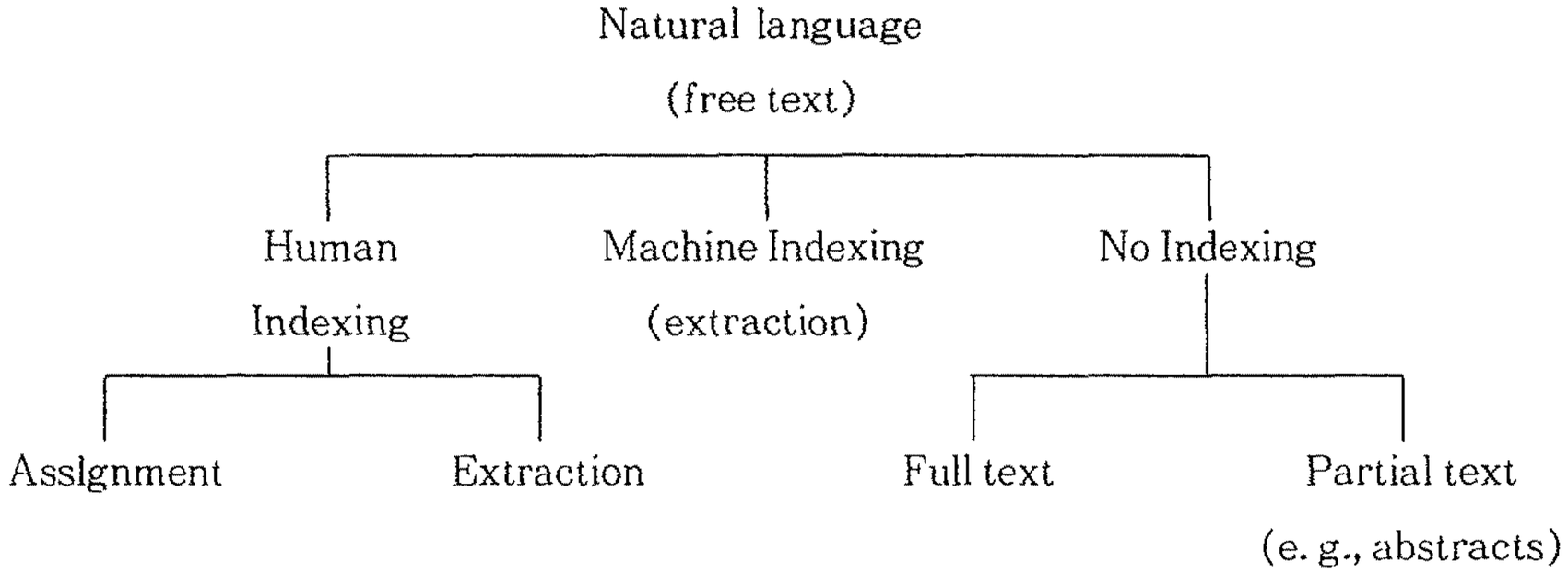
3. Uncontrolled Vocabulary

(1) Natural language

Uncontrolled, or natural language, words and terms are distinguished from words of a controlled vocabulary. The former can be of considerable value as subject access points when they come from such sources as the titling of a document, a cataloger's note on its record, or an abstract. Such a word, thought of by a searcher or used by an indexer as subject access point, is often called a keyword or a derived keyword, that is, from the document of its record rather than from a controlled vocabulary(Hagler 1991, 179).

〈Figure 3〉 illustrates different ways of implementing natural language systems in information retrieval. In a natural language retrieval system, the subject matter of documents and of information needs is represented by an open－ended vocabulary of words and phrases commonly used in the field. As shown in the diagram, a natural language system can be based on human index

〈Figure 3〉 Various possibilities in implementing natural language systems



ing, machine indexing, or no indexing at all. When human index with natural language, they are likely to do so by extracting words and phrases from the text of documents. This is indexing by extraction or derivative indexing. The assignment of additional terms not appearing in the text also could occur. Natural language indexing by computer which means machine indexing is always extraction indexing. Although it is possible to program a computer to assign terms to documents, the limited set of terms that the computer could be programmed for would constitute a controlled vocabulary(Lancaster 1986, 160).

The use of natural language of subject indexing is sometimes thought to have originated in the keyword-in-context(KWIC) and keyword-out-of-context(KWOC) indices which inaugurated the automated bibliographic file just after the second World War. Using the computer, indexers not only restored the KWIC/KWOC type of index but expanded the potential of natural language access points in subject searching. If a database mounted by a database vendor is not provided with a controlled vocabulary index, natural language searching is not a choice but a necessity.

(2) Automated Natural Language Indexing

In online searching, it is now considered very useful, even indispensable, to be able to search for natural language words at least in titles, and subtitles. Indexing these is no longer the labor-intensive work it once was, provided each word does not to have to be separately considered for its indexability value by a human indexer. Automated indexing depends on the availability in machine-readable form of the database to be indexed. This was once an expensive proposition requiring the keying of existing typed and printed documents anew. Programs written to make the necessary indexing decision for every word in a database(whether of titles, of abstracts, or even of full text) range from the simple-minded to the extremely sophisticated; the cost of running them varies accordingly. At the simple end, the computer may be asked to count the number of bytes per word and exclude all very short ones from the index to eliminate prepositions, articles, conjunctions, etc. Compiling a stop-

list to identify those words of natural language text which the program is not to make accessible in an index involves a little more judgement. It is a one－time cost to the indexing operation. In theory, automated indexing using a stop－list makes accessible only those words which are highly significant for retrieval purposes. To be effective, the stop－list must take into account the topics of particular material being indexed and its users' searching habits. This may be fruitful in limited subject areas with precise terminology; for interdisciplinary searching by generalists, using a stop－list is not an effective indexing techniques. More sophisticated automated indexing programs represent an attempt to identify intelligent human indexing and searching decisions as algorithms, by passing the need for a human to make the same decision over and over in the face of different documents and requests. The basis of decision is often the frequency of use of a word or word－group in the natural language of the material to be indexed or searched: at a certain threshold of use in a certain context, sentence pattern,etc., the program determines that a word or group of word is important enough to index or retrieve. A programmer can even use the morphological patterns of a language to determine what truncation will best leave only the meaningful roots of words to be searched. Investigating thought/language patterns with a view to automated translation and indexing is at the forefront of information research studies. The most advanced computers able to modify a search while it is in progression the basis of complex feedback information hold out considerable hope that what is still theory may someday be turned into practice on an operational scale in limited subject areas where the terminology and the information seeking patterns of users are predictable. Whether this will constitute artificial intelligence is still debatable (Hagler 1991, 181).

4. Hybrid Systems

The term 'hybrid' refers information retrieval systems operating on a combination of controlled terms and natural language. These include systems in

which both sets of terms are assigned by human indexers and systems in which a database can be searched on a combination of humanly assigned controlled terms and words occurring in titles and abstracts. Since the early days of computer-based systems in the 1950s, there has been evident trend toward simplification in information retrieval. It has become more widely recognized that it is possible to operate systems effectively with a minimal level of vocabulary control or with none at all. As an illustration of this trend, let us consider an actual case history involving a large information processing agency in the United States. The Central Intelligence Agency made a deliberate move away from a highly sophisticated system based on a large, carefully controlled vocabulary(a classification scheme including various relational indicators) to one of lesser sophistication, in which three separate vocabulary components are involved in the indexing and retrieval operations as follows :

- A small controlled vocabulary of approximately 250 broad subject codes
- A list of codes representing geographic areas(area codes)
- Keywords or phrases occurring in the titles or texts of documents

Keywords are uncontrolled in this system. the indexer may assign any word of phrases he wishes. The searcher is supplied with various searching aids, including a complete list of keywords with an indication of the number of times each one have been used, and lists showing the number of times each keyword has co-occurred with each subject code and vice versa. The result is crude classification of keywords that is nevertheless extremely useful in searching. This is an interesting example of one system operating effectively with a minimum of vocabulary control. In fact, the system works on a combination of small controlled vocabulary of broad descriptors used in association with a very large uncontrolled vocabulary of natural language expressions occurring in documents, particularly document titles. Many of the databases that are now accessible through online networks can be searched on combinations of controlled terms and keywords or phrases occurring in titles and abstracts, the latter permitting greater specificity. This combination is still powerful.

Ⅲ. Controlled Vocabulary versus Uncontrolled Vocabulary in Subject Searching

1. Advantages and Disadvantages in Subject Searching

Natural language systems offer one distinct advantage over controlled vocabulary systems. Because they use unrestricted vocabulary, they tend to allow greater specificity in retrieval. In fact, the more specific an information system need, the more likely it is that the natural language system will produce better results than the controlled language system. Controlled vocabularies have advantages, too. A Controlled vocabulary serves three principal functions: It tends to reduce semantic ambiguities, to promote consistency in the representation of subject matter, and to facilitate the conduct of comprehensive searches. The first function is achieved by distinguishing among the various meanings of homographs, the second through the control of synonymous expressions, and the third by some superimposed structure that links semantically related terms.

In the implementation of retrieval systems, a possible trade－off exists between input and output costs of effort. Controlled vocabulary systems impose cost and effort at the time of input, whereas natural language systems shift this cost and effort to output. An experienced searcher develop a strategy that will compensate for the lack of vocabulary control at input. In essence, he is using the strategy to achieve the same ends that a controlled vocabulary would be designed to achieve.

The function of the controlled vocabulary, and the probably the most important, is to facilitate the conduct of comprehensive searches. Also another alternative is for the system to display the thesaurus entry online and allow the searcher to select all necessary terms from the display without having to enter them at the keyboard. But controlled vocabulary system allows only the generic search, whereas natural language permits more specific search and has greater

flexibility. Finally, a very crude natural language system, one with no searching aids, has been assumed. In reality, a natural language system may help the searcher with various devices and aids. Some favorable assumptions have been made about the controlled vocabulary system, too. More-or-less perfect indexing has been assumed, whereas human indexing rarely approaches perfection.

2. Comparison of Controlled Vocabulary versus Uncontrolled Vocabulary

Still a debate continues in the library and information science professions over the relative merits, appropriateness, and efficiency of natural language subject access in online catalog systems. Recently, Peters and Kurth compared controlled and uncontrolled vocabulary subject searching in an academic library online catalog(Peters & Kurth 1991, 201-211). An analysis of transaction logs from an academic library online catalog describes instances in which users have tried both controlled and uncontrolled vocabulary subject access during the same search session. User persistence was greater during controlled vocabulary search legs, but search output was greater during uncontrolled vocabulary search legs. The results of this study indicate that uncontrolled vocabulary is not an primarily an option of last resort. Users combined uncontrolled vocabulary subject searches with controlled vocabulary subject searches. If the debate over controlled versus uncontrolled vocabulary subject access is to be grounded in actual user behavior, studies examining the ways that users test the limits of system functionality can improve future online catalog design and instruction.

Ⅳ. Subject Searching in Multilingual Online Catalog

1. Overview of Multilingual Systems

The technological developments that facilitate cooperation among information services in general make it increasingly possible for such cooperation to take

place on an international level. One impediment to effective global information transfer, however, is the large number of national language in use. So, the ability to handle publications in many languages and to accommodate several languages in the design of information services is becoming more important as truly international systems become operational.

Multilingual processing in documentation activities can be considered at several levels, including access to bibliographic records of foreign language literature and intellectual access to the content of actual documents. Significant improvements in intellectual access have occurred over the last 30 years. More material is translated than even before; translation is facilitated by improved glossaries and other aids; and national and International translation centers maintain records of translations, announce these translations, and offer copies for sale.

Multilingual access to bibliographic records can be broken down to at least three levels:translation of titles, translation of index terms, and translation of abstracts. Several indexing and abstracting services include titles in the original language and in translated form, but databases containing index terms in more than one language are rare. The French data base PASCAL does have some subfiles in which keywords appear in English and French, allowing searchers to be performed in either language. The Canadian Work place Automation Research Center has developed an online bilingual catalog of its specialized collections. Although it is not a general public catalog, the bilingual system (MINISIS) is accessible to local and remote users as a database. French language items are cataloged in French and English language items in English with descriptors assigned in the language of the item. In the case of bilingual items, separate records are prepared with bilingual descriptors (Rolland－Thomas & Mercure 1989, 146). The Canadian version of DOBIS is an integrated library management system which has been developed by the National Library of Canada and Canada Institute for Scientific and Technical Information. DOBIS is bilingual. Users can converse in either French or English. The system is built a

Canadian MARC which permits differentiation of the language of the document and of the subject headings by its special control field 17 and its common subfield $1(Rolland-Thomas & Mercure 1989, 147).

2. Use of Multilingual Thesaurus

In general, however, a database indexed by terms selected from language A can only be searched by terms in language B if A's terms are humanly translated into B's, possibly through a multilingual glossary or thesaurus. The compilation multilingual thesauri has been made easier by the appearance of the 'UNISIST Guidelines for the establishment and Development of Multilingual Thesauri' (UNISIST 1980) and development of computer programs to manipulate such thesauri. Whichever method is used, the major problem faced is finding equivalences among terms in the various languages. Some terms may translate exactly from one language to another, whereas others will not. In some cases. a term in one language needs not to be translated into two or more terms in a second language.

In a multilingual community, however, it may be more expedient to abandon multilingualism in favor of the adoption of a single language for indexing and searching. But bilingual or multilingual vocabularies can improve this situation. If a bilingual theaurus exists in machine-readable form, a national center can index and search in its own language and still prepare input that is completely compatible with the needs of the parent system. The thesaurus of International Nuclear Information System, for example, exits in French and Russian, as well as in English.

3. Classification Scheme as a Multilingual Pointer

For international operations, a strong case can be made for the adoption of a 'neutral' language. Classification schemes, such as the UDC or DDC, can satisfy this requirement since the class numbers in such a vocabulary are not depen-

dent on language. Specially, UDC has great advantages; namely, it exits in several languages and is used in many countries. Nevertheless, its use within the United States, still the World's largest producer of scientific and technical information, is insignificant. Lloyd (Lloyd 1989, 204－208) recognizing that information services in an international network would be reluctant to abandon vocabularies already in use, once urged that the UDC be adopted as a standard switching language with which each existing vocabulary could be reconciled. Use of the UDC for this purpose has not found favor, partly because of wide dissatisfaction with somewhat uncontrolled procedures by which this tool is updated. In fact, the development of the Broad System of Ordering can be considered a deliberate rejection of the UDC as an international switching device.

Classificaition schemes can provide multilingual pointers to the subject matter of documents. Of course, they do not provide a detailed level of access to the content of these documents. Future online catalog will have characteristics similar to bibilographic databases and will call for sophisticated processes including 'switiching languages' and multilingual thesauri. Another way to provide subject access to the searcher would be provide a built－in automation translation module. Multilingual thesauri are by definition confined to special areas of the universe of knowledge whereas OPACS thesauri should help the general with a wide variety of information needs.

V. Conclusions

It seems certain that natural language will become the norm in information retrieval and the use of conventional controlled vocabularies will decline. There are numerous reasons for this, including the escalating costs of human intellectual processing, the rapidly declining costs of computer storage, the increasing amount of text accessible in machine－readable form, and gradual reduction in dependence on the skilled intermediary in online searching. Manwhile, it seems reasonable to estimate that the development of postcontrolled vocabularies

offers considerable promise for improving the effectiveness and cost-effectiveness of online searching.

But still many studies show users of the online catalog or CD-ROM used to combine uncontrolled subject searches with controlled subject searches. In these days despite new text retrieval packages with loud claims of novelty and despite the availability of superhighways, the old fundemental techniques of classification and controlled vocabulary handling are needed and no compact answer has yet emerged. Furthermore controlled vocabulary subject access is more efficient in multilingual environment. Classification schemes such as DDC can provide multilingual pointers to the subject matter of document. Because they are international language which can be identified any place in the world regardless of their language.

〈References〉

- Bates, Marcia J., "How to Use Controlled Vocabularies More Effecively in Online Searching," *Online*, 14, 1988, pp. 45~55.
- Boyce, Bert R. and John P. Mclain, "Entry Point Depth and Online Search Using a Controlled Vocabulary," *JASIS*, 40(4), 1989, pp. 273~276.
- Byrne, Alex and Mary Micco, Improving OPAC Subject Access, *College & Research Libraries*, 1988, pp. 434~441.
- DeHart Florence E., "Subject Enhancement and OPACS: Planning Ahead," *Technical Services Quarterly*, 7(4), 1992, pp. 35~52.
- Gerhan, David R., "LCSH in vivo: Subject Searching Performance and Strategy in the OPAC Era," *Journal of Academic Librarianship*, 15(2), 1992, pp. 83~89.
- Gorman, Michael, *Technical Services Today and Tommorrow*, Englwood, Colo: Libraries Unlimited, 1990, pp. 74~197.
- Hagler, Donald, *Bibliographic Record and Information Technology*, 2nd ed., Chicago: American Library Association, 1991, 185p.
- Lancaster, F.W., *Vocabulary Control for Information Retrieval*, Arlington: IRP, 1986, pp. 159~174.
- Langridge, D. W., *Classification: Its Kinds, Elements, Systems and Applicaition*, London: Bowker-Saur, 1992, pp. 62~72.

- Larson, Ray R., "The Decline of Subject Searching: Long－Term Trends and Patterns of Index Use in an Online Catalog," *JASIS*, 42(3), 1991, pp. 195～215.
- Lloyd, G.A., "The UDC in Its International Aspects," *Aslib Proceedings*, 21, 1989, pp. 204～208.
- Markey, Karen, "Searching and Browsing the Dewey Decimal Classification in an Online Catalogs," *Cataloging and Classification Quarterly*, 7, 1988, pp. 37－68.
- Murray, Philip, "Documentation Goes Digital," *Byte*, 1993, pp. 121～129.
- Peters, Thomas A. and Martin Kurth, "Controlled and Uncontrolled Vocabulary Suvject Searching in an Academic Library Online Catalog," *Information Technology and Libraries*, 1991, pp. 201～211.
- Rolland－Thomas, Paule and Gerald Mercure, "Subject Access in a Bilingual Online Catalog," *Cataloginging and Classification Quarterly*, 10, 1989, pp. 141～163
- Slater, Ron, "Authority Control in a Bilingual OPAC:MultiLIS at Laurentian," *LRTS*, 35(4), 1992, pp. 422～457.
- Tyckoson, David ed., *Enhancing Access to Information: Designing Catalogs for the 21st Century*, New York: Haworth Press, 1991, 243p.
- *UNISIST Guidelines for the Establishments and Developments of Multilingual Thesauri*, Rev. ed., Paris : UNESCO(PG1－8－/WS12), 1980.
- Vickery, Brian C. ed., *Fifty Years of Information Progress*, London: ASLIB, 1994, pp. 85～118.