5 Subject Analysis

IRENE L. TRAVIS
Planning Research Corp.

RAYA FIDEL
University of Maryland

INTRODUCTION

Scope and Coverage

This review surveys recent developments reported in the literature of subject analysis. Of particular interest are those developments attributable to the continuously accelerating automation and communications technologies. The article highlights the impact of automation on subject analysis, including the automation of intellectual and clerical processes that are the components of subject retrieval systems.

The review covers the literature of subject analysis from 1977, the date of the last *ARIST* chapter on this topic (LISTON & HOWDER), through 1981. Because so much is published in this area, the sources covered here consist mainly of the research literature. (No attempt is made, for instance, to cover textbooks comprehensively.) Availability of the material is of special concern. Although considerable effort has been made to review the international scene, all covered material is in English or is available in English translation.

The amount of literature is so great that criteria for exclusion had to be developed. Citation indexing as well as specific indexing languages and classification schemes are not discussed. Exceptions are made for new types of languages, particularly string languages and switching languages. Also excluded is subject analysis as typically practiced in academic and public libraries. This area is reviewed annually in the summer issue of *Library Resources & Technical Services* (*LRTS*). Because we consider classification to be an integral part of all aspects of subject analysis and index language construction, it is treated as such in this review and not as a separate subject.

The authors thank the College of Library and Information Services, University of Maryland, College Park, and PRC Government Information Systems for their financial support. We also thank our bibliographic assistant, Katherine Wern, and Hans Wellisch for the use of his bibliographic materials.

Organization

The chapter is organized into four major sections. The first covers terminological control. It is followed by the closely related areas of thesaurus construction and index language development. Indexing theory and practice are discussed next, and finally, there is a brief section on new fields of application for subject analysis techniques.

Other Reviews and Bibliographies

In addition to the annual reviews in *LRTS*, there are other current sources for references to new research. The bibliographic section in each issue of *International Classification* deserves special mention since it supplements the coverage of this article in carrying English language abstracts of European work.

We also call attention to two comprehensive retrospective bibliographies, one recently published and another to appear in 1982. The first (WELLISCH, 1980a) is a monumental effort covering indexing and abstracting through 1976. The second, a three-part work being sponsored by *International Classification*, will cover the literature of classification generally (not just in its library applications) from 1950 to 1980.

There have also been several noteworthy review articles. WELLISCH (1980b) summarizes quantitative characteristics of the literature cited in his bibliographical survey. Other reviews include that by DAHLBERG (1977) on classification and that by LANCASTER (1977) on vocabulary control. Svenonius (SVENONIUS, 1981; SVENONIUS & SCHMIERER) has published two assessments of the state of the art of subject analysis. She pleads for more theoretical research and for more attention to cost effectiveness, which we agree is much neglected.

TERMINOLOGICAL CONTROL

The automation of bibliographic retrieval systems, especially the facility for interactive search provided by online systems, has an important effect on terminological control. The additional search capabilities available in computerized systems shed new light on the importance of terminological control for effective retrieval.

To discuss this issue, we introduce some definitions of our own. Databases that are indexed by descriptors assigned from a thesaurus have "controlled vocabulary indexing." Others have "uncontrolled vocabulary indexing." Index terms that are assigned without reference to any authority list create "natural language indexing." When a computerized database is searched for the occurrence of a certain term in one or several of the elements of the document record, it is "free-text" searched. (The term "full-text searching" is reserved for searches of the full text of the document.)

Controlled vs. Uncontrolled Vocabularies

Hard-copy retrieval tools can use controlled or natural language vocabularies for indexing but only a limited combination of both because of cost restrictions. Faced with this problem, designers and researchers have conducted tests comparing index languages in manual or batch environments. Comparisons of index languages (including free-text) are also suggested for online retrieval (e.g., DREESE; ROBERTS).

On the other hand, online systems can simultaneously accommodate different types of index languages for searching the same database since most commercially available databases can be free-text searched and many provide indexing (controlled or uncontrolled) as well. The online environment, therefore, introduces two new factors into research involved in controlled vocabulary vs. natural language vocabularies or free-text searching: 1) research can be directed to discover how these two types complement one another rather than to identify the one which performs best, and 2) testing can now be done on real-life systems since both search modes are available.

Although the literature contains statements about the relationship between controlled or uncontrolled vocabularies and precision and/or recall, the notion that each has its merits and demerits is gradually emerging. LANCASTER (1977; 1979) and RAITT discuss the importance and advantages and disadvantages of the use of controlled and uncontrolled vocabularies in searching. Raitt concludes, "successful searching depends on thinking of alternative approaches to retrieval as well as a knowledge of the vocabulary used" (p13).

Two examples of small-scale experiments to test the properties of controlled and uncontrolled vocabularies in an online environment are provided by CARROW & NUGENT and by HENZLER. Carrow and Nugent compared retrieval provided by controlled vocabulary and free-text searching of 23 user requests. They found that searching with controlled vocabulary terms resulted in significantly better recall but no difference in precision. Based on a failure analysis, they suggested improvements in the search strategy and concluded that the two search methods largely complement one another. Henzler, on the other hand, quantitatively compared free-text terms with the controlled vocabularies used in indexing in the CANCERNET system. He examined 100 thesaurus descriptors and found that only 5% were properly represented in the text when no vocabulary aids or word-proximity operators were used. On the other hand, only 70% of two samples of 100 title words could be mapped into the thesaurus without loss of information. After sketching the advantages of each search method, he concluded that both free-text and controlled vocabulary approaches, in an optimum combination, are necessary.

More research in this direction is badly needed. Although free-text searching is a valuable feature of online systems, the importance of a controlled vocabulary for searching online systems is unquestionable. Index languages based on controlled vocabularies can provide methods for improving the precision and/or recall of retrieved sets. The searching of multidisciplinary or

multilingual databases can also be facilitated only with some kind of terminological control. Finally, a well-structured thesaurus can support searching by any method, including free-text searching.

Problems in Terminological Control

The problem of terminological control is to impose standards on a dynamically growing language, which is, as KOBLITZ claims, a "socially determined phenomenon that serves people in the cognitive process: for exchanging thoughts, emotions, and wishes; and for recording and storing the acquired knowledge" (p10). In his opinion, terminological control has largely been performed without scientific guidelines and has been determined by the development of practical information activities. As a first step in enhancing scientific guidance in this area, he summarizes the process of terminological control by outlining 12 sequential steps of fundamental importance—e.g., defining objectives and identifying sources of information. In a substantial treatise, DAHLBERG (1978) considers the problems of concept definition, identification, and organization as applied to the development of a new universal classification scheme. LESKI & LESKA analyze errors in terminological control and suggest that the creation of terminological banks may be important in changing the current situation. Such banks would collect terms and their definitions, establish relationships among terms, and support semantic and morphological analysis.

Machine-aided terminological control utilizing terminological banks is a new and developing area. A terminological bank for the social sciences called Interconcept is described below. DOVBENKO & UMANSKI, moreover, show that the relationships among terms included in the definition of a term in a terminological bank and the term itself may correlate with relationships also identified in indexing languages. They conclude that coordinating the construction and maintenance of terminological banks with thesauri would aid reliable decisions about terms, relations, and definitions to be included in both constructs. The structure of multilingual terminological banks is described by RONDEAU.

Part of the problem involved in term standardization is the choice of the "best" term to represent a concept. KAPLUN & AZGALDOV propose an interesting qualitative method to support such decisions. Based on the goals and the environment of the system, the properties characteristic of term quality are hierarchically arranged in a "property tree," using a qualimetric approach. The authors give examples of how to determine numerical values for properties, but it is not clear whether such detailed analysis for each term would be feasible in operating systems.

Using as an example the terminology of information science, CLEVERDON points to some disadvantages of strict terminological control for retrieval purposes. He suggests that the following typology could be useful in terminological control:

- Terms that have an accepted meaning;
- Accepted terms that have an imprecise meaning;
- Terms that are still competing for general acceptance;
- Terms that are introduced by researchers; and
- Terms that are so general that they will probably never have an accepted definition.

With or without a well-accepted, integrated theoretical basis, terminological control is widely used. Problems reported by practitioners relate most often to the terminology of scientific disciplines. Since the terminology of the social sciences causes special difficulties, there have been several studies in that area. GREAVES and JULIAN looked at the field of education. The latter examined the descriptors assigned to a sample of articles using two index languages, while the former compared the representations of 50 concepts in four thesauri.

Interconcept, in contrast, is a system and a program initiated by Unesco (United Nations Educational, Scientific, and Cultural Organization) to provide actual conceptual control and analysis in all the social sciences. Begun in the late 1970s, the system is intended to serve as a terminological bank with various products. The project is still in its initial stage. VASARHELYI outlines the problems that the project is supposed to solve and the solutions under study. Some results of research into specific aspects of the project have been published. The criteria for term selection and for identification of relevant definitions for the bank are delineated by DAHLBERG (1981a). Examining three terminological paradigms—normative, analytic, and synthetic—RIGGS (1979) observes that the decision to establish the terminological bank indicates the usefulness and feasibility of synthetic glossaries. RIGGS (1981) further supports the decision to create the Interconcept glossary by examining lexical, terminological, and thesaural paradigms.

The relationship of biological nomenclature to classification was the subject of the 1980 meeting of the Aslib Biological and Agricultural Science group. In papers presented in this meeting, PARKER examined the problems involved in using a standard as a guide to preferred names of living organisms, especially in the light of the ever-continuing developments in taxonomy, while NORRIS, in describing the use of MEDLINE (Medical Literature Analysis and Retrieval System Online) for searching biological literature, mentioned the loss of specificity that results from terminological control.

Finally, GÖDERT analyzes and characterizes mathematical terminology. The author points out that everyday terms are used extensively with different meanings and that the use of general terms (e.g., theory) cannot be avoided. Quasi-synonyms are also more abundant in this field than one might expect. He uses a classification scheme as an aid in terminological control.

The existence of free-text searching as an available option has not been integrated into most of the discussions about terminological control. Some of the problems of terminological control may seem less severe when this method of searching is kept in mind. For example, in following typologies such as the one developed by CLEVERDON, we might identify categories of terms that should not be standardized. We might decide, for example, that

new terms generated by research workers should not be controlled at all until they are generally accepted. These terms can still serve as free-text or natural language access points. It seems that the effect of automated search procedures on the design of controlled vocabularies is not yet fully recognized or implemented.

INDEX LANGUAGES

This section covers thesaurus construction and maintenance, string languages, and recent research into vocabulary structures to facilitate integrated searching of multiple databases and multilingual databases.

Thesaurus Construction and Maintenance

The notion of what constitutes a thesaurus varies. The publication of new guidelines and manuals for thesaurus construction such as those by AUSTIN & WATERS, BRITISH STANDARDS INSTITUTION (1979a; 1979b), and TOWNLEY & GEE has been followed by discussions of this topic. SOMERS compares various definitions of a thesaurus as well as guidelines for thesaurus construction, while ROBERTS critically reviews some of these perceptions of what qualifies as a thesaurus, which he claims often are reflected in guidelines and standards for thesaurus construction rather than in unambiguous definitions. A study of 20 thesauri, sponsored by the Commission of the European Communities concluded that the previous guidelines were poor in facilitating standardization, lacked quantitative data to guide decisions (such as the ratio of precoordination), and did not provide clear instructions for problems such as the notation of semantic relations (VAN SLYPE). To review the literature on thesauri, we adopt the following broad definition: a thesaurus is any terminological control device that is used for indexing and/or retrieval.

Theoretical developments in thesaurus construction. Several attempts have been made in the past five years to develop general and/or theoretical approaches to the construction of thesauri. These attempts have taken two types of approach. The first relates the process of thesaurus construction to a general framework. The second approach constructs formal models of the process of thesaurus construction.

In some instances the framework used is a generalized approach to index languages. In an attempt to develop a general theory of index languages, BHATTACHARYYA (1979a) defines two structural dimensions for subject classifications: 1) the semantic and 2) the elementary. The first is the intrinsic dimension of the subject's meaning, and the second is the artificially imposed dimension of the constituent elements of the subject's name. He systematically generalizes fundamentals of different index languages. SOKOLOV (1979), on the other hand, defines three types of index language, which together form a compound language for integrated information retrieval systems. The first is designed for file retrieval, the second for document retrieval, and the third for retrieval of data (e.g., facts, quotations). He

analyzes the characteristics of each language as derived from its functions. Finally, KOROLEV (1977a) places the "information language" within a typology of automated systems for text processing. The typology is based on two characteristics: 1) the extent of development of language devices used to resolve semantic problems, and 2) the extent of the use of predefined algorithms and of the subsequent automation of the text-handling processes.

The role and nature of the thesaurus have also been related to theoretical frameworks borrowed from other areas. The most attractive are general systems theory and linguistics. FOSKETT, for instance, outlines some practical implications of the systems approach to classification. Other researchers have linked thesaurus construction to linguistics. HENRIKSEN, for instance, using the Universal Decimal Classification (UDC) as an example, suggests that index languages could be analyzed as semiotic systems (i.e., those that deal with functions of signs and symbols in both constructed and natural languages) according to the criteria of structural linguistics. LEICHIK, on the other hand, analyzes the features of a classification language as a scientific and technical sublanguage.

Formal models of thesauri have various purposes. A mathematical model can provide a formal description or classification of entities. CHKHENKELI, for example, describes several types of index languages. Applying elements of set theory, he defines sets of lexical units, formalizes the relationships among these units, and analyzes the characteristics of these relations. He then classifies languages according to parameters such as: 1) the nature of the lexical units, 2) the allowances for paradigmatic relations, and 3) the existence of a grammar and of algorithms to select descriptors from among all natural language terms included in the system.

Formal models can also be used to develop theories. Simply put, one can "map" entities and relations from the subject matter examined to establish a formal construct. Various relations can then be deduced within this construct, which may then be "mapped" back to the issue under study. Several such attempts in thesaurus construction are described in the eastern European literature (e.g., RADECKI); however, most manage only the first step.

A more pragmatic use of mathematical models is to develop measures to be used in the actual construction process. GORKOVA & SHISHOVA (1979a; 1979b) developed a simulation model to determine structural connections among thesaurus descriptors. The definitions of the connections are based on the theory of electric circuits and on the statistical frequency of term co-occurrence in documents. The authors process the subject heading list of the abstract journal *Informatics* and present the resulting tree of connections.

New methods. New methods and approaches to the individual steps in thesaurus construction are also described. Term selection is discussed by KIM & KIM as an issue in knowledge organization. They challenge the prevailing assumption that the structure of knowledge based on a consensus of experts in a field is different from that expressed in the literature. In an experiment they compared a list of terms from the social sciences generated by a committee of experts with a list of terms selected on the basis of frequency of occurrence in articles sampled from five major journals. Their

finding that the two sets were not significantly different led them to criticize some features of existing guidelines for thesaurus construction—e.g., they do not provide specific and operational instructions, and some of their recommendations may not result in improved retrieval.

BUSCH and ANDRUKOVICH & KOROLEV are also investigating term selection. Busch calls for more integration of terms derived from searchers' experience. Andrukovich and Korolev, using a computerized thesaurus, studied the relationship between frequency of occurrence of words in text and lexicogrammatical properties (e.g., number of meanings, number of synonyms). They conclude that frequency does correlate with lexicogrammatical complexity; therefore, it can serve as a basis for term selection; thus, low-frequency words are easily matched against databases. They should not be included in the thesaurus; instead they should be free-text searched.

Effect of automation. Automation affects thesaurus construction in two interrelated ways: 1) the computer can be used to mechanize functions such as editing, correcting, and generating indexes to thesauri; 2) automation can affect the design criteria for thesauri. In the first category KAZLAUSKAS & HOLT describe the application of a minicomputer in constructing the thesaurus for the National Information Center for Special Education Material (NICSEM), while DEXTRE & CLARKE suggest that the system developed for the *ROOT* thesaurus for the British Standards Institution (BSI) could be used to maintain and construct other thesauri. DEVADASON & BALASUBRAMANIAN describe a system that was developed experimentally at the Documentation Research and Training Center of the Indian Statistical Institute. The system automatically generates a thesaurus from subject headings that are structured according to facet analysis. The merits of graphic displays in a thesaurus are described by ROLLING.

Turning to changes in design criteria, we first mention again the work of ANDRUKOVICH & KOROLEV, which is clearly applicable. Also interesting are suggestions by JONES and by WALL. Jones reviews the problems associated with the use of compound words in thesauri (i.e., words or phrases that were formed by putting words or parts of words together) and the solutions suggested by various studies and systems. He argues that methods for handling compound terms in an online search can be independent of the treatment of compounds in the index language. Wall proposes a system that would identify the hierarchical relationships of new terms by distinguishing overlapping terms from other relations. The system is dynamic, and the structure could be modified for each term added. He concludes that implementation of such a system would facilitate an online thesaurus that might be complex in its internal structure but that would provide complete tracking across, up, or down the thesaurus.

DOSZKOCS demonstrates the feasibility of online associative search tools in a system at the National Library of Medicine (NLM). The Associative Interactive Dictionary (AID) displays a ranked list of terms that are statistically associated with terms from the searcher's query. The term associations are derived from the search output by a measure based on the difference between observed and expected term frequencies in retrieved sets. Users may add all or some of the suggested terms to their search.

Despite these interesting results, the effect of automated search procedures on thesaurus design is largely unexplored. This area should be important for research in the near future.

String Languages

The past five years have seen a continued interest in the development of techniques for constructing precoordinated subject headings as opposed to post-coordinated unit index terms. System developers are trying to use simple human coding to allow computers to produce reasonable index entries generated from a single string of text, and there are ongoing efforts to discover optimal word orders to enhance intelligibility and searchability. Although they vary in sophistication and complexity, these systems are all designed for printed indexes. There has been little investigation of their utility in online retrieval. FARRADANE (1977a) reminds us that string languages have yet to be truly tested as general retrieval tools.

These systems can be characterized by the degree of specification and intellectual analysis that are required to construct the computer input or "string," as it has come to be called. Although no system offers any new insight into how to select "the subject" of a document, some require considerable analysis of that subject once it is identified in order to code the string [e.g., PRECIS (Preserved Context Indexing System)]. At the other end of the scale is unenhanced keyword permutation, as exhibited by KWIC (keyword in context) or KWOC (keyword out of context) indexes, which are of little theoretical interest.

The systems that are based on well-developed theories of word order and heading structure are: 1) PRECIS, developed at the British National Bibliography; 2) POPSI (Postulate-based Permuted Subject Indexing), developed at the Documentation Research and Training Center in Bangalore, India; and 3) Farradane's relational indexing system. Craven's NEPHIS (Nested Phrase Indexing System) and LIPHIS (Linked Phrase Indexing System) require less intellectual analysis by the indexer. (These systems are discussed in detail in the section on Specific String Languages below.) SVENONIUS (1978) provides an historical perspective on string languages by reviewing the problems in categorizing concepts, which were recognized by J. Kaiser nearly a century ago and are still troublesome today.

At the almost purely pragmatic end of the scale, we have MISCHO (1977; 1979; 1980) whose systems at Iowa State University, Ames permute enriched titles and Library of Congress (LC) Subject Headings. An interesting development is the new system at the Institute for Scientific Information (ISI) (VLADUTZ & GARFIELD), which is attempting to apply automated indexing techniques to the coding of strings, making the technique fully automated, rather than using the computer only at the clerical level to generate multiple entries from humanly coded entries.

Specific string languages. Taking the reported systems individually, the least well known in the United States is probably POPSI, developed by Bhattacharyya. POPSI is a classification-based system, incorporating faceting

in term order and flagging. It is very similar to PRECIS but puts more stress on controlled vocabulary and consistency. BHATTACHARYYA (1979b) gives a very full exposition of the system in the Indian journal *Library Science with a Slant to Documentation.* Of the two extensive comparisons of PRECIS and POPSI, the one by RAJAN & GUHA is worthwhile but unfortunately now out of date. Many of the features of PRECIS these authors criticized have now been changed. The other article (MAHAPATRA) is more superficial and contains some serious errors in the examples used.

PRECIS is the most widely discussed system. Two books about it have been published in the period covered by this review. The first, the proceedings of a conference on PRECIS at the University of Maryland in 1976 (WELLISCH, 1977) contains good general articles about PRECIS, but the applications articles are now out of date. Fortunately, good alternative sources exist. ROBINSON, for example, has recently surveyed Canadian developments, which accounted for most of the systems reported at the conference, while BAKEWELL provides a broad overview of PRECIS use. The failure of LC to give PRECIS serious consideration as an alternative to its Subject Headings is lamented by DYKSTRA.

The other book is a text on PRECIS by RICHMOND. It appears to be considerably more useful as a teaching tool than the PRECIS manual. Other general introductions to PRECIS are provided by AUSTIN & DIGGER and by WEINTRAUB.

Articles that describe the use of PRECIS in languages other than English or the translation of English PRECIS strings into other languages have appeared less frequently in the past five years than in the previous five, at least in the sources that we reviewed. Multilingual use is the subject of two articles by SØRENSEN (1977a; 1977b), but the work reported was done before the period of this review. Verdier (VERDIER; VERDIER & AUSTIN) discusses translation of PRECIS strings from English to French. The lack of reference to linguistics by string language developers, criticized by MICHELL, is particularly evident here.

NEPHIS and LIPHIS are described in several articles by CRAVEN (1977b; 1978a; 1978b). LIPHIS, based on a network model, can handle a broader range of grammatical structures in the input string than NEPHIS, which handles only tree structures. Both systems provide indexers with a relatively simple way to take title-like phrases that have not been structured in some special manner and to code them to produce acceptable permuted entries for printed catalogs or other applications. It is the minimal rules for ordering the original string that distinguishes NEPHIS and LIPHIS from PRECIS or POPSI. Craven has also designed a system to incorporate vocabulary control, using the mapping of natural language phrases to a controlled list to produce cross references in the index when the index is generated (CRAVEN, 1978b).

Some special-purpose languages can be considered string languages. One example is provided by ANDERSON (1979; 1980) in an experimental system for the Modern Language Association (MLA). Indexers assign facets derived from literature and linguistics to terms in a string. The strings are then used to form alphabetical subject entries.

String language research. There has also been an increased amount

on the properties of printed indexes in general and of string languages in particular has been carried out by Keen in his EPSILON (Evaluation of Printed Subject Indexes by Laboratory Investigation) project. He reports the structure of his experiments (KEEN, 1977a) and some conclusions as to the amount of rephrasing that users must do to read different types of printed entries (KEEN, 1977b). Similar experiments were done by JAMIESON, who also compared the effects of connectives and word order on comprehension and processing speed. Both projects raise questions about how people actually scan printed indexes. Do they hunt for keywords, read the full entries, or partially read the entries until they reject them as nonrelevant? Both authors stress that answers to these questions are badly needed as a basis for index design. Such research is also pertinent to the design of displays of search results in online systems. Many systems provide some means for preliminary screening of retrieval, and such screens can be read as printed pages. These aspects of user behavior are very difficult to research but are fundamental to index design.

A study that specifically looked at features associated with POPSI but that is of wider interest is reported by RAGHAVAN & IYER. They found that the inclusion of a contextual subject field in the string was often confusing to users, that users had difficulty interpreting strings with more than seven components, and that entry of concretes and differences ("speciators") in inverse order also resulted in misinterpretations of the string—e.g., adjective-noun order is read more easily than noun-adjective.

FARRADANE (1977a) also compares four systems—PRECIS, POPSI, NEPHIS, and an adaptation of his relational indexing system—though in a less rigorous way. His brief characterizations of their differences are of considerable interest. The *String Indexing Series* from the University of Western Ontario (AUSTIN & VERDIER; CRAVEN, 1977a; FARRADANE, 1977b) also provides brief descriptions and numerous examples of indexing in PRECIS, NEPHIS, and Farradane's system. All three authors were asked to index the same set of articles from the *Journal of the American Society for Information Science* (*JASIS*), but Svenonius, the series editor, feels that there is insufficient control over the input strings to allow tight comparison of the results. In the series introduction that accompanies each volume she notes that one must accept that the language influences the selection of indexable concepts as well as the actual heading structures.

String languages have been an active and productive area of research and development (R&D). Their future role, however, remains in question so long as there is no clear understanding of how they can contribute to retrieval in an online environment.

Language Compatibility for Integrated Retrieval Systems

The application of computer technology to information storage and retrieval systems has made large scale integrated retrieval systems possible. Databases that cover various areas within the same discipline, that were created in different languages, or that include various types of media (e.g.,

major subproblem is the mechanism for integrating subject retrieval. Thus, much effort has been invested in compatibility research, particularly in Europe. The proposed Integrated Thesaurus of the Social Sciences that is being developed by the Unesco Division for the International Development of the Social Sciences (LITOUKHIN) provides impetus to research in language compatibility. Sager (SAGER & MCNAUGHT; SAGER ET AL.) has recently published guidelines for comparison and compatibility among thesauri in the social sciences, while DAHLBERG (1981b) suggests that various types of conversion tables (which she calls compatibility matrices) may be used as tools for compatibility research.

Subject integration can be achieved in different ways. WERSIG reminds us of three approaches to compatibility among index languages: 1) use of a single index language in all systems; 2) establishment of a concordance among the index languages used; or 3) development of a switching language to allow movement from one index language to another. A similar typology of language compatibility mechanisms is suggested by VILENSKAYA (1977; 1980), who defines: 1) a "world language"—a universal language used on a global level of communication; 2) an "intermediary language"—a language that is supposed to transfer the contents of documents expressed in terms of any index language to another without loss of information; and 3) a "switching language"—a language that deals with files or even with whole information systems through which the transition from one system to another is performed. Each mechanism can be used either to integrate existing systems or to create new integrated systems.

Examples of a single index language for an integrated or multilingual database include the system described by KOLLIN & KURANZ, who converted access terms to merge three databases in management-related areas. The problems involved in creating a single index language for several databases are also delineated by KURBAKOV & BOLDOV. They suggest a multistep process that distinguishes between the classificatory structure of an index language and its vocabulary. Multilingual thesauri that represent a single index language are used by several specialized information centers in Europe, and experiences with some of them are described (GIERTZ; MUSSO & ZANGRANDO; NEUMANN-DUSCHA & ULENBERG).

Dictionaries and concordances, on the other hand, are used mainly to translate thesaurus terms from one index language to another. MORTON and RONDEAU describe a structure for multilingual concordances, while PIGUR gives a linguistic basis for selecting equivalent descriptors from different natural languages. Some experimental and operating systems are described in the literature. CANISIUS describes a system for translating abstracts and descriptors used at the Bundesanstalt für Strassenwessen in Cologne. The terms are translated with a dictionary that assigns a code to each term. SEMTURS suggests the use of a multilingual thesaurus as a dictionary in generating inflections to support automated free-text searching of multilingual databases. A more specific technique, based on English and French texts and thesauri, is described by FIELD.

Turning to the third category, switching languages, we find that most of the proposed mechanisms do not qualify, strictly speaking, as languages. Therefore, we use the less restrictive term, switching mechanism. Switching

mechanisms can be built pragmatically from an analysis of the existing vocabularies to be integrated, or they can be based on a general construct.

There are several examples of switching mechanisms based on the conceptual and terminological information contained in existing vocabularies. The Vocabulary Switching System (VSS) developed by NIEHOFF ET AL. (1979; 1980) is an experimental automated subject switching mechanism for searching multiple databases in a single natural language. It was developed from existing energy-related vocabularies, and it consists of several key switching logics that are combined with the vocabularies. In a preliminary study for the Integrated Thesaurus of the Social Sciences, MEYRIAT also analyzed 60 existing information languages to identify for each the scope and depth of coverage of 41 subject fields. A more general mechanism for analyzing existing vocabularies is explained by ANTOPOLSKII ET AL. They define the concept of "lexical intersections" (a set of lexical units from different thesauri found by means of formal and/or semantic identification) for use in the development of a switching mechanism. In an experiment they examined the efficiency of various correspondence criteria for retrieving lexical intersections.

In contrast to these somewhat pragmatic approaches, NEELAMEGHAN (1979a) maintains that switching mechanisms are easier to develop when the index languages have a common knowledge structure and framework for representation of subjects. It is not surprising, therefore, that elements from classification theory, especially facet analysis, are proposed to enhance switching mechanisms. A system that was developed as a switching language is the Broad System of Ordering (BSO). COATES ET AL. (1979) describe its development and application. They explain, using Ranganathan's terms, that the diversity of the different index languages is in the language and terminology plane and that switching is feasible in the thought and idea plane. As a universal classification system, BSO can serve as a concept-holding device (a common reference for clustering related terms) to enhance switching. In a critique of BSO, SOERGEL concludes that despite its announced purpose, BSO cannot serve as a switching mechanism. Such mechanisms, he maintains, cannot use a very broad scheme but require an enormous language. This conclusion is supported by a study carried out by DAHLBERG (1980). She examined BSO as a possible basis for the Integrated Thesaurus of the Social Sciences, which AITCHISON maintains can be used as a switching language itself. Dahlberg concludes that the shallowness provided by the BSO precludes the possibility of its serving in this capacity. A more general approach is put forth by SOKOLOV (1977), who suggests a general faceted framework for newly created thesauri to ensure compatibility in a general system. He and his associates used this framework to construct compatible thesauri for the subjects of navigation and seaports (SOKOLOV ET AL.).

Among existing systems, PRECIS is also suggested as a switching mechanism. SØRENSEN (1977b) examines the correlation between the role operators used in PRECIS and certain "deep cases" (such as location and agent), which are now regarded as linguistic universals, to conclude that PRECIS could be adapted for various European languages. VERDIER & AUSTIN describe a research program set up by the British Library to inves-

tigate the switching of PRECIS input strings from one language to another.

Language compatibility is clearly a major focal point of present research, especially in Europe. The Unesco integrated thesaurus project, in particular, may produce results of general interest and applicability.

INDEXING THEORY AND PRACTICE

A complete and commonly accepted theory of indexing, once created, would be a central theoretical construct in information science. The importance of such a theory is recognized by many researchers, and several theories of indexing have been developed. A synthesis of compatible theories may prove to be the first step in establishing a comprehensive indexing theory. In 1977 Borko described indexing theories formulated by Jonker, Heilprin, Landry, and Salton. He concluded that although some of them needed to be validated, the "remarkable degree of congruence among these theories. . .is indicative that a comprehensive theory of indexing may not be far in the future" (BORKO, p365). Unfortunately, no systematic attempt has been made to carry out this particular suggestion.

Recent Research

By contrast, there has been a considerable coalescence among American and British researchers interested in the application of probability theory to indexing and retrieval. MARON & KUHNS first proposed their theory of probabilistic indexing in 1960. In the late 1970s and early 1980s this approach is being heavily discussed. Of particular interest is Maron's recent work with other researchers to derive general models that incorporate the various approaches of several schools of indexing theorists now active. W. S. Cooper, Maron's colleague at Berkeley, recently published a seminal work on utility-theoretic indexing, discussed below (COOPER, 1978). Maron and Cooper subsequently worked together to incorporate their theories into a single, more general model (COOPER & MARON), and we understand that there is now an effort to coordinate these results with those of British researchers who also use probabilistic approaches, specifically ROBERTSON. The general flavor of the writings of these authors is conveyed in a highly readable issue of the *Drexel Library Quarterly* (MARON, 1978a) that contains articles by Maron, Cooper, Wilson, van Rijsbergen, Harter, Robertson, and Kuhns.

Probabilistic approaches to indexing are closely tied to work on the use of relevance feedback in improving information retrieval system performance. The British experiments with probabilistic approaches center, not on initial indexing, but on weighting terms on the basis of relevance feedback from searchers. Since much of this work was recently reviewed for *ARIST* by MCGILL & HUITFELDT, we do not cover it extensively here. However, a recent collection edited by ODDY ET AL. is noteworthy. Of particular interest is an article by ROBERTSON ET AL., who experimented with Harter's probabilistic indexing techniques for term weighting, thus

providing another link among this group of researchers. In studying feedback, these researchers are, of course, joined by Salton, who, in a paper that also helps pull together recent experimental and theoretical work, compares the results of several weighting techniques (SALTON ET AL.). He finds that weighting schemes that incorporate relevance feedback are superior to weighting schemes that are based on inverse collection term frequencies alone, even if relevance information is minimal or estimated. His experiments confirm results reported by SPARCK JONES (1979a; 1979c) and also tally with those obtained by Robertson et al. in the study mentioned above. This area of research is one of the few in which we can see definite progress and cross examination of results among researchers. The research group is sizable and includes more workers than can be mentioned here.

An axiomatic theory of indexing is being created by FUGMANN (1979; 1980). He acknowledges the influence of other theoreticians (prominently Ranganathan, the Indian School, Rush, and Landry), and states that the explanatory value of the theory has been tested since it was first proposed in 1972. In his 1979 paper (p13), he lists five axioms that he claims are suitable "to describe and explain all currently known phenomena in information supply." They are: 1) the axiom of definability, 2) the axiom of order, 3) the axiom of the sufficient degree of order, 4) the axiom of predictability, and 5) the axiom of fidelity. He demonstrates the applicability of this theory to various aspects of indexing. In analyzing index languages in his 1979 paper, Fugmann concludes that only highly syntactical and rigorously controlled languages are accurate enough to support continually changing information demands. When Fugmann's theory is more fully developed, it may be useful for describing and explaining indexing phenomena.

LANCASTER (1980) takes a very different view. He believes that what he perceives as the present trend toward free-text searching as a substitute for indexing will continue. He speculates that controlled vocabulary indexing will be displaced entirely.

In the traditional document-oriented indexing approach, the first step in the indexing process is to determine what the document is "about." HUTCHINS discusses the literature on "aboutness" (with which he is not, unfortunately, completely familiar). Based on an expansion of the "theme/rheme" notions of functional sentence perspective (FSP) (according to which the "theme" expresses what the sentence as a whole is "about"), he concludes that if users want just one or two documents on a topic, the "aboutness" should be determined from the early passages of the text. FSP is discussed below in the section on Automated Indexing and Abstracting. SPARCK JONES (1979b), on the other hand, takes a skeptical view of the relationship between meaning representation in text and "aboutness." In a review of the artificial intelligence (AI) literature on the representation of meaning, she concludes that meaning representation and indexing are different in kind not just in degree.

The Subject Access Project (SYRACUSE UNIVERSITY) also used document-oriented techniques but in an automated environment. The MARC (Machine Readable Cataloging) record indexing of a test collection of books was augmented by enriching the indexing with natural language index terms from back-of-the-book indexes or tables of contents (SETTEL). The retrieval

tests showed that searches done with the enriched indexing yielded more documents and were shorter than the searches conducted on the MARC record indexing alone. No significant difference in precision was measured. ATHERTON recommends the enriched indexing for online subject catalogs.

User-oriented indexing approaches are more feasible in interactive systems. TAGUE tested the value of adding natural language terms from previous queries to the indexing of documents that had proven relevant to those queries. The first phase included an analysis of the relationships between titles and user-supplied terms. She concludes that it is impractical to allow users to add natural language index terms to the indexing of documents in an operating online retrieval system because of problems such as lack of error control.

In the theoretical paper mentioned above (which received the *JASIS* Best Paper Award in 1978), COOPER (1978) also takes a user-oriented approach. He recommends that indexers use "gedanken" or "thought" experiments to estimate the utility of each index term to the system's users. A term should be assigned to a document only if, in the opinion of the indexer, the benefits outweigh the costs. Some of his assumptions are questioned by WILSON.

While the practice of indexing is being rapidly influenced by automation, the theory of indexing is evolving more slowly. The relatively short experience with operating automated retrieval systems (whose design is often based on the principles of manual systems) cannot indicate all the differences between indexing for manual and automated searching, although the results of work by Salton and others would certainly indicate that they may be substantial. For example, it is not clear whether indexer consistency is important for retrieval performance, although it is being measured (e.g., DIODATO; HENZLER). In searching online retrieval systems, indexing is used also as feedback to improve query formulation. A searcher can display the index terms assigned to relevant citations to find the common descriptors. Knowing that indexing serves this additional function, we may wish to re-examine the issue of indexer consistency and its importance. Which types of consistency, for example, are important for retrieval, and which are irrelevant? With more experience in automated retrieval, answers to questions such as these will be incorporated into our theoretical approaches.

Automated Indexing and Abstracting

Although there are a number of active researchers in automated indexing and abstracting, little new material is appearing in the journal literature. Most of the operational systems discussed below were functioning before 1977. What is new is the involvement of some of the major database suppliers [Chemical Abstracts Service (CAS), BioSciences Information Service (BIOSIS) and ISI] in automated indexing research and a renewed interest in methods for assigning documents to broad classes using statistical techniques that operate on small amounts of text. Also notable is the growing group of researchers who are interested in applying the principles of FSP to automated indexing and abstracting, as described below.

The literature in this area is exceptionally frustrating to review. First, there is the problem of what is missing. Much automated indexing activity is probably not in the journal literature, either because it is so cheap, simple, and obvious that no one thinks it worthwhile to report or because the systems are proprietary. A survey aimed at eliciting the use of simple phrase-extraction techniques by organizations that process large amounts of text might find such devices being used rather commonly. In addition, the major journals that are publishing work in this area need to upgrade their editorial practices, particularly in monitoring system descriptions for intelligibility and in editing contributions from authors whose first language is not English.

Dictionary-based automated indexing systems. A large class of automated indexing systems, including almost all those used in a production environment, use a dictionary to assign grammatical categories to words (e.g., noun, verb, or categories tailored to the application). These categories are then used to carry out syntactic analyses that result in phrase extraction. Systems of this type that are under development in the USSR are reported by LEONTEVA & VISHNYAKOVA, SHINGAREVA, and TKACH. As a group, they are characterized by more complex grammatical analysis than is usual in operational U.S. systems. Shingareva, for example, sorts terms into frames with categories such as "initial product." KOROLEV (1977b) confirms that statistical techniques for automated indexing are seldom used, which seems surprising, since KNOWLES characterizes the application of statistics to linguistics in the USSR as extensive. A review article in English that covers recent Russian work from Russian language sources is badly needed.

VLEDUTS-STOKOLOV is carrying out a related program of research in the United States at BIOSIS to automate their indexing. The approach uses dictionaries to analyze text terms into facets. Fortunately, BIOSIS can use existing files of taxonomies and chemical names. Vleduts-Stokolov estimates that a dictionary of only 13,000 additional terms will have to be specially constructed.

An older experimental system developed at the National Institutes of Health (NIH) in the United States also involves the assignment of terms to role categories. This system, reported by DUNHAM ET AL., translates diagnostic reports into the Systematized Nomenclature of Pathology (SNOP) for human review. SNOP is a faceted classification scheme (although its creators do not call it that). This automated indexing system was heavily documented before the publication of the cited article in medically oriented information science journals.

Systems of a similar type but involving less syntactic analysis have been developed for the International Nuclear Information System (INIS) in Vienna (BARNES ET AL.) and for the Defense Technical Information Center (DTIC) and the Smithsonian Science Information Exchange (SSIE) in the United States, by KLINGBIEL and HERSEY ET AL., respectively. None of these systems was begun after 1977. The DTIC and SSIE systems, however, have been used daily as production systems. [SSIE was recently absorbed by the National Technical Information Service (NTIS).]

In eastern Europe, work reported by JANOŠ (1979a) is also linguistically oriented. Janoš is attempting to apply the principles of FSP to automated indexing—that is, the "theme/rheme" or "topic/comment"

analysis mentioned above. As JANOŠ (1979b, p20-22) explains, FSP divides sentences into that which is spoken about, develops the preceding context, and is bound to it, and that which is not bound to the preceding context and "pushes" the sentence forward—the "theme" and the "rheme." Eastern European languages and others contain grammatical clues to this structure, although English relies heavily on word order and is therefore more difficult to analyze automatically from this point of view. Most indexing and abstracting researchers, however, are not interested in applying these notions at the sentence level but rather at the text level to analyze "functional text perspective." Under this approach, text is analyzed to find sentences that state, for example, the formulation of the initial hypothesis or the relationship of an experiment to previous work. MAEDA ET AL. in Japan are also using this approach to isolate significant phrases from text. Janoš's research does not seem to have progressed very far, while the article describing the Japanese system is difficult to follow, and the application of the theme/rheme ideas to indexing is vague. A logical but apparently isolated effort to automate string analysis for string language indexing is under way at ISI (VLADUTZ & GARFIELD), as mentioned earlier.

Statistically-based automated indexing systems. Experimental activity in the United States and Great Britain with statistically based automated indexing systems has continued. HARTER has produced a very readable and frank tutorial paper on statistical techniques in automated indexing for a theme issue of the *Drexel Library Quarterly* that also includes a tutorial on automatic classification by VAN RIJSBERGEN (1978). Both articles are highly recommended for readers who are seeking an introduction to these techniques.

WEINBERG carried out an ambitious research project in statistical automated indexing. In a meticulous study she demonstrates that humanly assigned index terms do not fall into any one frequency class. They are low-frequency, high-frequency, and medium-frequency terms in text, assuming that they occur at all. She also provides a detailed account of the problems of compiling term frequency statistics from photocomposition tapes. HELBICH and PAO also explored statistical selection of index terms, as did BRINER, using a model based on Shannon's equations for information content and channel capacity.

Salton and his associates at Cornell have also continued their work, which was discussed in the section on Indexing Theory and Practice, Recent Research. One indexing-oriented paper is a model-based analysis of the effects of adding terms from a controlled vocabulary to supplement terms from the text (YU ET AL.). In this paper, as in most of his recent ones, Salton stresses that medium-frequency terms are the best index terms. This statement seems to contradict Weinberg's findings, but Salton is not claiming to duplicate human indexing. In fact, Salton's SMART (System for the Mechanical Analysis and Retrieval of Text) system is so different from the conventional indexing and abstracting systems (used in Weinberg's study) that there is little carryover. SMART does not select a few terms or phrases to represent a document in a printed index or database. Instead it carries almost all text words with their frequencies in storage, creating a vector to represent the document. Requests are then treated in the same way, and retrieval is implemented by measuring

the distance between the request vector and the document vectors. This approach allows retrieval by the overall patterns of word usage. Although Weinberg's dissertation is interesting, it does not prove that statistically based indexing, even with its well-known limitations, cannot be useful under the proper circumstances. It does prove that it will not behave like human indexing if term extraction only is used. It also does not address the question of using statistical techniques to assign headings to documents from controlled vocabularies, as the researchers described below are attempting.

Two other systems, still statistically based but somewhat apart from the term- or phrase-extraction experiments described above, are the assignment indexing system being developed at CAS by HAMILL & ZAMORA (1978; 1980) and the one being built for the U.S. Energy Information Agency (EIA) by CAHN & HERR at Lawrence Berkeley Labs (LBL). Both attempt to assign documents to broad categories according to the statistical correlation of keywords with the categories, as determined by analysis of a "teaching set" of humanly indexed documents; this technique was pioneered by MARON (1961). The CAS system, based on work by KAR & WHITE at Ohio State University, uses only terms from titles, while the LBL system uses index terms from other databases.

Automated abstracting. Ideas from FSP also figure prominently in current work in automated abstracting. Almost all reported automated abstracting studies, many of whose algorithms are not yet programmed, attempt to identify different functional parts of abstracts (e.g., conclusions and methodology) through the use of key phrases. Papers in this area include those by MAEDA (1981a; 1981b), JANOŠ (1978; 1979b), and PAICE. None of these authors attempts to show that this approach improves abstracting quality beyond that achieved by previous methods (e.g., RUSH ET AL.).

Given the present state of the art, are there cost-beneficial applications of automated indexing? As mentioned above, several systems do operate in production environments, although this fact alone does not prove that they are cost effective. Moreover, we know at what level an automated indexing system will perform. Reports over the years have been very consistent. Almost any algorithm will produce accuracy around 60%, and an enhanced set of procedures will produce up to 80% accuracy in terms of acceptable headings. The amount of underassignment of headings is not frequently discussed. The issue that really has not received serious attention is the conditions under which that range of performance is acceptable.

Possible application situations occur either when indexer productivity must be increased or when human indexing has never been feasible, while free-text searching is also inadequate because of storage or time constraints or because it retrieves an unmanageable amount of material. We need much more experience and better cost figures on the use of automated indexing to preprocess text for human review, and we need a better notion of when automated indexing can reasonably replace or augment free-text searching. We also need to know more about the cumulative effects of a 20% error rate for unreviewed automated indexing, whether and how that level is tolerable, and for what types of applications. Now that text in machine-readable form is

becoming widely available, it is the answers to these questions rather than the indexing technologies themselves that are holding back rapid proliferation of the simpler automated indexing systems.

NEW APPLICATION FIELDS

In the early stages of preparing this review, we had planned to cover extensively subject analysis as applied to materials other than those usually kept by libraries. As somewhat expected, however, the literature in this area is too thin, too scattered, and, in many cases, too unsophisticated to be of great interest. However, some items deserve brief consideration.

Archives seem to be an area that presents challenging problems of subject retrieval, but the literature is not inspiring. Some idea of the general state of automation in archives, including some discussion of subject access problems, appears in the proceedings of a recent symposium co-sponsored by the American Society for Information Science (ASIS) and the University of Maryland (MCCRANK). Of more direct interest is the dissertation, also done at Maryland, by LYTLE (1979), whose findings have appeared in *American Archivist* (LYTLE, 1980a; 1980b). He compares retrieval using the provenance method traditionally used in archives with retrieval using a thesaurus.

Closely related to the problem of archives subject access is that of access to current organizational data, of which archives are, so to speak, the fossilized remains. SWANSON & CULNAN surveyed the literature pertaining to the implementation of document control systems in industry. While the article is worth reading, we feel that in this case, as with low-level automated indexing techniques, a literature survey gives little impression of the true state of the art. Most internal document control systems, such as project libraries, are not sophisticated, and people do not bother to write them up. What is needed is a survey of businesses themselves. One system that takes an integrated approach to document control in a business setting is AMANDA (Automated Management of Document Access), reported by SCHWARTZ ET AL. CARROLL has experimented with automated indexing in a word processing environment. We believe that business document control has tremendous potential for subject retrieval system designers, but a lot of selling needs to be done to the data processing professionals who control these systems. CULNAN (1980; 1981) also discusses the application of information control to the automated office, and HAINES gives a readable overview of the problems of correspondence control.

An application area for sophisticated subject control techniques is the description of numeric data. MURDOCK (1978; 1980) states that interest in identifying and indexing statistical data in texts seemed to have peaked by 1977 when its expense became obvious. Recently, however, there has been increased interest in the control of government-collected data and databases. Information managers and others in several federal and state agencies have undertaken data-indexing projects—e.g., the EIA's FEDEX (Federal Energy Data Index) system (FORD & BROWN) and Data Resources Directory (TRAVIS). The Paperwork Reduction Act of 1980 also mandates a system to describe statistical data. The growth of the use of online numeric databases generally should continue to spark interest in indexing and describing data. The development of data dictionaries is closely related to these efforts. Some references to the literature in this area were included in the *ARIST* review of database management systems (DBMSs) in 1979 (HUFFENBERGER & WIGINGTON, p171).

CONCLUSION

With the advent of online interactive systems, the boundaries between indexing and searching are increasingly blurred. Indeed, as attempts to measure indexing effectiveness in the previous 20 years clearly show, they are not separable, and online interaction reinforces the merger. It is not surprising, therefore, that the areas of research that are most active now are closely tied to this interdependency. Under this rubric we include many issues in terminological control, index language compatibility, and term weighting with relevance feedback. However, the full effect of this interdependency on research patterns and on the way we think about retrieval problems is still unfolding.

Another clear pattern over the past 20 or 30 years has been the shift from thinking of different retrieval techniques as opposing systems to considering them as complementary. Repeatedly we have allowed ourselves to fall into the trap of asking "which is better," when we should have been asking, "how can we combine these techniques to improve performance?" For the first half of this century the issue was classification vs. indexing; then it became citation indexing vs. subject indexing; most recently, it is free-text searching vs. controlled vocabularies. To this list of subject retrieval modes we should also add automated indexing in its various forms, which is still seen as an opposing rather than complementary mode. For archives and records management, there is also provenance arrangement. The controversies have, of course, helped to elucidate the properties of the systems, but it is their interaction that we must understand. System automation is a catalyst. Not only is it the necessary vehicle for implementing many of these techniques individually, it is also the necessary ingredient for implementing them simultaneously on one file to provide the capability to cycle or to intermix them.

Finally there is the question of new application areas such as those discussed above. Subject analysis is a mature field with a large body of theory and practice gained over more than a century of systematic study. The very existence of this body of knowledge is largely unknown to the general public. As information resources become more and more unmanageable in all contexts, the need for this knowledge is apparent to anyone who is familiar with it. The difficulty is to persuade data processing professionals, who often reinvent the wheel, to pay more attention to this area. Quite apart from the survival of librarians as professionals or of library schools as institutions, it would be a terrible waste of knowledge and skill if the practitioners and body of literature in this field were passed by in the information age. A review article cannot solve this problem, but it can remind us of our responsibilities to share this knowledge.

BIBLIOGRAPHY

AITCHISON, JEAN. 1981. Integration of Thesauri in the Social Sciences. International Classification (Germany). 1981; 8(2):75-85. ISSN: 0340-0050.

ANDERSON, JAMES D. 1979. Contextual Indexing and Faceted Classification for Databases in the Humanities. In: Tally, Roy D.; Dueltgen, Ronald R., eds. Information Choices and Policies: Proceedings of the American Society for Information Science (ASIS) 42nd Annual Meeting: Volume 16; 1979 October 14-18; Minneapolis, MN. White Plains, NY: Knowledge Industry Publications, Inc.; 1979. 194-201. ISSN: 0044-7870; ISBN: 0-914236-47-4; LC: 64-8303.

ANDERSON, JAMES D. 1980. Structure in Database Indexing. The Indexer (England). 1980 April; 12(1): 3-13. ISSN: 0019-4131.

ANDRUKOVICH, P. F.; KOROLEV, E. I. 1977. The Statistical and Lexicogrammatical Properties of Words. Automatic Documentation and Mathematical Linguistics. 1977; 11(2): 1-11. ISSN: 0005-1055.

ANTOPOLSKII, A. B.; SALII, A. D.; FEDOSIMOV, V. I. 1979. Automatic Retrieval Methods for Lexical Intersections of Information Retrieval Thesauruses. Automatic Documentation and Mathematical Linguistics. 1979; 13(1): 85-91. ISSN: 0005-1055.

ATHERTON, PAULINE. 1978. Books Are for Use—Evaluation of MARC Records in Online Subject Retrieval Systems. 17-20. See reference: BRENNER, EVERETT H., comp.

AUSTIN, DEREK; DIGGER, JEREMY A. 1977. PRECIS: The Preserved Context Index System. Library Resources & Technical Services. 1977 Winter; 21(1): 13-30. ISSN: 0024-2527.

AUSTIN, DEREK; VERDIER, VERONICA. 1977. PRECIS: Introduction and Indexing. London, Ontario, Canada: University of Western Ontario, School of Library and Information Science; 1977. 1 volume (unpaged). (String Indexing Series, E. Svenonius, ed.).

AUSTIN, DEREK; WATERS, J. 1980. UNISIST Guidelines for the Establishment and Development of Multilingual Thesauri. Revised text. Paris, France: Unesco; 1980. (PGI/80/WS/12).

BAKEWELL, K. G. B. 1979. Indexers' Reactions to PRECIS. Journal of Documentation (England). 1979 September; 35(3): 164-178. ISSN: 0022-0418.

BARNES, C. I.; COSTANTINI, L.; PERSCHKE, S. 1978. Automatic Indexing Using the SLC-II System. Information Processing & Management. 1978; 14(2): 107-119. ISSN: 0306-4573.

BASER, KENNETH H.; COHEN, STANLEY M.; DAYTON, DAVID L.; WATKINS, PENNELL B. 1978. On-Line Indexing Experiment at Chemical Abstracts Service: Algorithmic Generation of Articulated Index Entries from Natural Language Phrases. Journal of Chemical Information and Computer Sciences. 1978 February; 18(1): 18-25. ISSN: 0095-2338.

BHATTACHARYYA, G. 1979a. Fundamentals of Subject Indexing Languages. 83-89. See reference: NEELAMEGHAN, A., ed.

BHATTACHARYYA, G. 1979b. POPSI: Its Fundamentals and Procedure Based on a General Theory of Subject Indexing Languages. Library Science with a Slant to Documentation (India). 1979 March; 16(1): 1-34. ISSN: 0024-2543.

BIVINS, KATHLEEN T. 1980. An Approach to an Atomic Theory of Indexing. In: Taylor, Peter J., ed. New Trends in Documentation and Information: Proceedings of the International Federation for Documentation (FID) 39th Congress; 1978 September 25-28; Edinburgh, Scotland. London, England: Aslib; 1980. 55-60. (FID Publication no. 566). ISBN: 0-85142-128-8.

BORKO, HAROLD. 1977. Toward a Theory of Indexing. Information Processing & Management. 1977; 13(6): 355-365. ISSN: 0306-4573.

BRENNER, EVERETT H., comp. 1978. The Information Age in Perspective: Proceedings of the American Society for Information Science (ASIS) 41st Annual Meeting: Volume 15; 1978 November 13-17; New York, NY. White Plains, NY: Knowledge Industry Publications, Inc.; 1978. 381p. ISSN: 0044-7870; ISBN: 0-914236-22-9; LC: 64-8303.

BRINER, L. L. 1978. A Mathematical Theory of Indexing. 55-58. See reference: BRENNER, EVERETT H., comp.

BRITISH STANDARDS INSTITUTION. 1979a. Guidelines for the Establishment and Development of Monolingual Thesauri. London, England: British Standards Institution; 1979. (BS 5723; 1979). ISBN: 0-580-10740-x. Available from: British Standards Institution, Sales Department, 101 Pentenville Road, London N1 9ND.

BRITISH STANDARDS INSTITUTION. 1979b. Guidelines for the Establishment and Development of Multilingual Thesauri. London, England: British Standards Institution; 1979. (BS 5723: 1979). Available from: British Standards Institution, Sales Department, 101 Pentenville Road, London N1 9ND.

BUSCH, JOSEPH. 1978. A Method for Evaluating the Multiple Relations between Subject Descriptors—Related Terms in the *Thesaurus of Engineering and Scientific Terms:* A Pilot Study. 59-62. See reference: BRENNER, EVERETT H., comp.

CAHN, DAVID F.; HERR, JESSIE J. 1978. Automatic Document Classification Based on Expert Human Decisions. 63-65. See reference: BRENNER, EVERETT H., comp.

CANISIUS, P. 1977. Automatic Partial Translation in a Multilingual Information System. 259-269. See reference: COMMISSION OF THE EUROPEAN COMMUNITIES, eds. Volume 1.

CARROLL, JOHN M. 1981. Towards a Marriage of Word Processing and Document Content Analysis. In: Lunin, Lois F.; Henderson, Madeline; Wooster, Harold, eds. The Information Community—An Alliance for Progress: Proceedings of the American Society for Information Science (ASIS) 44th Annual Meeting: Volume 18; 1981 October 25-30; Washington, DC. White Plains, NY: Knowledge Industry Publications, Inc.; 1981. 136-138. ISSN: 0044-7870; ISBN: 0-914236-85-7; LC: 64-8303.

CARROW, DEBORAH; NUGENT, JOAN. 1977. Comparison of Free-Text and Index Search Abilities in an Operating Information System. In: Fry, Bernard M.; Shepherd, Clayton A., comps. Information Management in the 1980s: Proceedings of the American Society for Information Science (ASIS) 40th Annual Meeting: Volume 14, Part 2: Full Papers; 1977 September 26-October 1; Chicago, IL. White Plains, NY: Knowledge Industry Publications, Inc.; 1977. 2-e8; Microfiche; 24X reduction. ISSN: 0044-7870; ISBN: 0-914236-12-1; LC: 64-8303.

CHKHENKELI, T. I. 1977. The Formal Representation of Descriptor Information Retrieval Languages. Automatic Documentation and Mathematical Linguistics. 1977; 11(2): 115-119. ISSN: 0005-1050.

CLEVERDON, CYRIL W. 1979. Terminological Control. 32-39. See reference: GORKOVA, V. I.; BAKULINA, I. I., eds.

COATES, ERIC; LLOYD, GEOFFREY; SIMANDL, DUSAN. 1978. BSO—Broad System of Ordering: Schedule and Index. 3rd revision. The Hague, The Netherlands: International Federation for Documentation (FID); 1978. 102p. (FID Publication no. 564). ISBN: 92-66-00564-9.

COATES, ERIC; LLOYD, GEOFFREY; SIMANDL, DUSAN. 1979. The BSO Manual: The Development, Rationale and Use of the Broad System of Ordering. The Hague, The Netherlands: International Federation of Documentation (FID); 1979. 157p. (FID Publication no. 580). ISBN: 92-66-00580-0.

COMMISSION OF THE EUROPEAN COMMUNITIES, eds. 1977. Overcoming the Language Barrier: Proceedings of the 3rd European Congress on Information Systems and Networks; 1977 May 3-6; Luxembourg. Munich, Germany: Verlag Dokumentation; 1977. 2 volumes. ISBN: 3-7940-5184-x.

COOPER, WILLIAM S. 1978. Indexing Documents by Gedanken Experimentation. Journal of the American Society for Information Science. 1978 May; 29(3): 107-119. ISSN: 0002-8231.

COOPER, WILLIAM S. 1979. Utility-Theoretic Indexing: A Note on Wilson's Note. Journal of the American Society for Information Science. 1979 May; 30(3): 170-172. ISSN: 0002-8231.

COOPER, WILLIAM S.; MARON, M. E. 1978. Foundations of Probabilistic and Utility-Theoretic Indexing. Journal of the Association for Computing Machinery. 1978 January; 25(1): 67-80. ISSN: 0004-5411.

CRAVEN, TIMOTHY C. 1977a. NEPHIS: Introduction and Indexing. London, Ontario, Canada: University of Western Ontario, School of Library and Information Science; 1977. 1 volume (unpaged). (String Indexing Series; E. Svenonius, ed.)

CRAVEN, TIMOTHY C. 1977b. NEPHIS: A Nested-Phrase Indexing System. Journal of the American Society for Information Science. 1977 March; 28(2): 107-114. ISSN: 0002-8231.

CRAVEN, TIMOTHY C. 1978a. Linked Phrase Indexing. Information Processing & Management. 1978; 14(6): 469-476. ISSN: 0306-4573.

CRAVEN, TIMOTHY C. 1978b. A NEPHIS Thesaurus for Computer Generation of Permuted Cross References. 79-82. See reference: BRENNER, EVERETT H., comp.

CRAVEN, TIMOTHY C. 1980. A General Concept Network Structure for the Production of Index Displays. Information Processing & Management. 1980; 16(6): 269-276. ISSN: 0306-4573.

CULNAN, MARY J. 1980. Document Processing in the Automated Office: Implication for MIS Research and Education. In: McLean, Ephraim, ed. Proceedings of the 1st International Conference on Information Systems; 1980 December 8-10; Philadelphia, PA. 165-173. (Sponsored by: Research and Education in Information Systems, The Society for Management Information Systems). OCLC: 7062585.

CULNAN, MARY J. 1981. Information Science and the Automated Office: Challenges and Opportunities. In: Lunin, Lois F.; Henderson, Madeline; Wooster, Harold, eds. The Information Community—An Alliance for Progress: Proceedings of the American Society for Information Science (ASIS) 44th Annual Meeting: Volume 18; 1981 October 25-30; Washington, DC. White Plains, NY: Knowledge Industry Publications, Inc.; 1981. 139-142. ISSN: 0044-7870; ISBN: 0-914236-85-7; LC: 64-8303.

DAHLBERG, INGETRAUT. 1977. Major Developments in Classification. In: Voigt, Melvin J.; Harris, Michael H., eds. Advances in Librarianship: Volume 7. New York, NY: Academic Press; 1977. 41-103. ISBN: 0-12-785007-4.

DAHLBERG, INGETRAUT. 1978. Ontical Structures and Universal Classification. Bangalore, India: Sarada Ranganathan Endowment for Library Science; 1978. 64p. (Sarada Ranganathan Lectures, no. 11). Available from: the author, Woogstr. 36a, D-6000, Frankfurt 50, Germany.

DAHLBERG, INGETRAUT. 1980. The Broad System of Ordering (BSO) as a Basis for an Integrated Social Sciences Thesaurus? International Classification (Germany). 1980 July; 7(2): 66-72. ISSN: 0340-0050.

DAHLBERG, INGETRAUT. 1981a. Conceptual Definitions for INTERCONCEPT. International Classification (Germany). 1981; 8(1): 16-22. ISSN: 0340-0050.

DAHLBERG, INGETRAUT. 1981b. Towards Establishment of Compatibility between Indexing Languages. International Classification (Germany). 1981; 8(2): 86-91. ISSN: 0340-0050.

DEVADASON, F. J.; BALASUBRAMANIAN, V. 1981. Computer Generation of a Thesaurus from Structured Subject-Propositions. Information Processing & Management. 1981; 17(1): 1-11. ISSN: 0306-4573.

DEXTRE, S. G.; CLARKE, T. M. 1981. A System for Machine-Aided Thesaurus Construction. Aslib Proceedings (England). 1981 March; 33(3): 102-112. ISSN: 0001-253x.

DIODATO, VIRGIL P. 1981. Author Indexing. Special Libraries. 1981 October; 72(4): 361-369. ISSN: 0038-6723.

DOSZKOCS, TAMAS E. 1978. An Associative Interactive Dictionary (AID) for Online Bibliographic Searching. 105-109. See reference: BRENNER, EVERETT H., comp.

DOVBENKO, M. A.; UMANSKI, A. 1980. A Comparative Analysis of Information Retrieval Thesauri and Normalised Scientific and Technical Terminologies with the Aim of their Coordinated Development. International Forum on Information and Documentation (USSR). 1980 October; 5(4): 9-11. ISSN: 0304-9701.

DREESE, M. J. 1979. Comparison of Document Retrieval Using UDC-Notation Given by UDC-Expert with Catchwords Taken from Title by Clerk, Stored in Computer Memory. 220-223. See reference: NEELAMEGHAN, A., ed.

DUNHAM, GEORGE S.; PACAK, MILOS G.; PRATT, ARNOLD W. 1978. Automatic Indexing of Pathology Data. Journal of the American Society for Information Science. 1978 March; 29(2): 81-90. ISSN: 0002-8231.

DYKSTRA, MARY. 1978. The Lion That Squeaked: A Plea to the Library of Congress to Adopt the PRECIS System and to Reconsider the Decision to Overhaul the LC Subject Headings. Library Journal. 1978 September; 103(15): 1570-1572. ISSN: 0000-0027.

FARRADANE, JASON. 1977a. A Comparison of Some Computer-Produced Permuted Alphabetical Subject Indexes. International Classification (Germany). 1977 November; 4(2): 94-101. ISSN: 0340-0050.

FARRADANE, JASON. 1977b. Relational Indexing: Introduction and Indexing. London, Ontario, Canada: The University of Western Ontario, School of Library and Information Science; 1977. (String Indexing Series, E. Svenonius, ed).

FIELD, B. J. 1977. Automatic Indexing for Multilingual Systems. 469-492. See reference: COMMISSION OF THE EUROPEAN COMMUNITIES, eds. Volume 1.

FORD, SHELLEY G.; BROWN, EULALIE W. 1981. Federal Energy Data Index on BRS. Database. 1981 March; 4(1): 28-35. ISSN: 0162-4105.

FOSKETT, DOUGLAS J. 1980. Systems Theory and Its Relevance to International Classification. (Germany). 1980 April; 7(1): 2-5. ISSN: 0340-0050.

FRASER, WALTER J. 1978. Literary, User and Logical Warrants as Indexing Constraints. 130-132. See ref.: BRENNER, EVERETT H., comp.

FUGMANN, ROBERT. 1979. Towards a Theory of Information Supply and Indexing: Treatise VI on Retrieval System Theory. International Classification (Germany). 1979 April; 6(1): 3-15. ISSN: 0340-0050.

FUGMANN, ROBERT. 1980. On the Practice of Indexing and Its Theoretical Foundations. International Classification (Germany). 1980 April; 7(1): 13-20. ISSN: 0340-0050.

FUGMANN, ROBERT; DENK, WALTER; NICKELSEN, INGEBORG. 1980. Variations in the Order-Creating Power of Interactive Retrieval Systems; Treatise VIII on Retrieval System Theory. International Classification (Germany). 1980 July; 7(2): 73-78. ISSN: 0340-0050.

FUGMANN, ROBERT; WINTER, JAKOB H. 1979. Reserve Retrieval: Toward Analogy Inferences by Mechanized Classification; Treatise VII on Retrieval System Theory. International Classification (Germany). 1979 July; 6(2): 85-91. ISSN: 0340-0050.

GIERTZ, L. M. 1979. The SfB-System: Case Study on an Information System for Global Cooperation in the Field of Buildings Construction. 349-355. See reference: NEELAMEGHAN, A., ed.

GÖDERT, WINFRIED. 1980. Subject Headings for Mathematical Literature. Journal of Documentation (England). 1980 March; 36(1): 11-23. ISSN: 0022-0418.

GORKOVA, V. I.; BAKULINA, I. I., eds. 1979. Essential Problems in Terminology for Informatics and Documentation. Moscow, USSR: All-Union Institute of Scientific and Technical Information; 1979. 114p. (FID Publication no. 569). ISSN: 0203-6444.

GORKOVA, V. I.; SHISHOVA, L. A. 1979a. Criteria for Evaluating Structural Links between Concepts in Classification Systems. Automatic Documentation and Mathematical Linguistics. 1979; 13(1): 22-35. ISSN: 0005-1055.

GORKOVA, V. I.; SHISHOVA, L. A. 1979b. Models for Structuring Terminological Systems. 38-87. See reference: GORKOVA, V. I.; BAKULINA, I. I., eds.

GREAVES, MONICA A. 1978. A Comparative Study of Verbal Subject Data in Information Retrieval Systems in Education. Education Libraries Bulletin. 1978 Summer; 21(2): 1-19. ISSN: 0013-1407.

HAINES, MICHAEL. 1979. Company Correspondence: An Important Information Resource. Aslib Proceedings (England). 1979 August; 31(8): 401-411. ISSN: 0001-253x.

HAMILL, KAREN A.; ZAMORA, ANTONIO. 1978. An Automatic Document Classification System Using Pattern Recognition Techniques. 152-155. See reference: BRENNER, EVERETT H., comp. 152-155.

HAMILL, KAREN A.; ZAMORA, ANTONIO. 1980. The Use of Titles for Automatic Document Classification. Journal of the American Society for Information Science. 1980 November; 31(6): 396-402. ISSN: 0002-8231.

HARTER, STEPHEN P. 1978. Statistical Approaches to Automatic Indexing. Drexel Library Quarterly. 1978 April; 14(2): 57-74. ISSN: 0012-6160.

HELBICH, JAN. 1979. Experimental Comparison of Nine Statistical Procedures for Measuring Selective Power of Single Words. 246-254. See reference: NEELAMEGHAN, A., ed.

HENDRIX, FRANCES; WILSON, STEPHEN. 1981. The Provision of a Subject Index at Preston Polytechnic Library and Learning Resources Service: Use of an Adaptation of PRECIS at Preston Polytechnic. Program (England). 1981 April; 15(2): 73-90. ISSN: 0033-0337.

HENRIKSEN, TOR. 1979. A Linguistic Description Model for Indexing Languages. 3-8. See reference: NEELAMEGHAN, A., ed.

HENZLER, ROLF G. 1978. Free or Controlled Vocabularies: Some Statistical User-Oriented Evaluations of Biomedical Information Systems. International Classification (Germany). 1978 March; 5(1): 21-26. ISSN: 0340-0050.

HERSEY, D. F.; FOSTER, W. R.; PAYNE, W. H.; HUNT, B. L. 1979. Toward a Machine-Assisted Indexing Vocabulary for International Use: Recent Developments at the Smithsonian Science Information Exchange. 255-260. See reference: NEELAMEGHAN, A., ed.

HUFFENBERGER, MICHAEL A.; WIGINGTON, RONALD L. 1979. Database Management Systems. In: Williams, Martha E., ed. Annual Review of Information Science and Technology: Volume 14. White Plains, NY: Knowledge Industry Publications, Inc.; 1979. 153-190. ISSN: 0066-4200; ISBN: 0-914236-44-X; LC: 66-25096.

HUTCHINS, W. J. 1978. The Concept of "Aboutness" in Subject Indexing. Aslib Proceedings (England). 1978 May; 30(5): 172-181. ISSN: 0001-253x.

JACHOWICZ, R. L. 1979. Application of Classification as Basis for the Formulation of Thesaurus. 356-362. See reference: NEELAMEGHAN, A., ed.

JAMIESON, ALEXIS JEAN. 1981. Comprehension of Entries in Printed Indexes. College Park, MD: University of Maryland; 1981. (Ph.D. dissertation). 249p. Available from: University Microfilms, Ann Arbor, MI.

JANOŠ, JIŘÍ. 1978. Results of an Experiment in Automatic Extracting and Problems in the Use of Condensed Texts in Automated Information Systems. International Forum on Information and Documentation (USSR). 1978 January; 3(1): 13-17. ISSN: 0304-9701.

JANOŠ, JIŘÍ. 1979a. Information Systems with Automated Indexing Based on the Analysis of Texts. 296-303. See reference: NEELAMEGHAN, A., ed.

JANOŠ, JIŘÍ. 1979b. Theory of Functional Sentence Perspective and Its Application for the Purposes of Automatic Extracting. Information Processing & Management. 1979; 15(1): 19-25. ISSN: 0306-4573.

JONES, KEVIN P. 1981. Problems Associated with the Use of Compound Words in Thesauri, with Special Reference to BS 5723: 1979. Journal of Documentation (England). 1981 June; 37(2): 53-68. ISSN: 0022-0418.

JULIAN, C. A. 1978. A Comparative Study of Subject Approaches in C.I.J.E. and Education Index, with Reference to Three Areas in Education. Sheffield, England: Sheffield University; 1978. 127p. ERIC: ED 188581.

KAPLUN, L. M.; AZGALDOV, G. G. 1979. Evaluation of Terms and Terminological Systems: A Qualimetric Approach. 88-103. See reference: GORKOVA, V. I.; BAKULINA, I. I., eds.

KAR, GAUTAM; WHITE, LEE J. 1978. A Distance Measure for Automatic Document Classification by Sequential Analysis. Information Processing

KAZLAUSKAS, EDWARD JOHN; HOLT, THOMAS D. 1980. The Application of a Minicomputer to Thesaurus Construction. Journal of the American Society for Information Science. 1980 September; 31(5): 363-368. ISSN: 0002-8231.

KEEN, E. MICHAEL. 1977a. On the Generation and Searching of Entries in Printed Subject Indexes. Journal of Documentation (England). 1977 March; 33(1): 15-45. ISSN: 0022-0418.

KEEN, E. MICHAEL. 1977b. On the Processing of Printed Subject Index Entries During Searching. Journal of Documentation (England). 1977 December; 33(4): 266-276. ISSN: 0022-0418.

KIM, CHAI; KIM, SOON D. 1977. Consensus vs. Frequency: An Empirical Investigation of the Theories for Identifying Descriptors in Designing Retrieval Thesauri. Information Processing & Management. 1977; 13(4): 253-258. ISSN: 0306-4573.

KLINGBIEL, PAUL H. 1977. Machine-Aided Indexing in Practice: An Encounter with Automatic Indexing of the Third Kind. Alexandria, VA: Defense Documentation Center; 1977. 11p. ERIC: ED 156141.

KNOWLES, FRANCIS E. 1979. Recent Soviet Work on Computer Techniques for Representing Natural Language Meaning. In: MacCafferty, Maxine; Gray, Kathleen, eds. The Analysis of Meaning: Proceedings of a Conference Held by the Aslib Informatics Group and the BCS Information Retrieval Specialist Group; 1979 March 26-28; Oxford, England. London, England: Aslib; 1979. 70-84. (Informatics 5). ISBN: 0-85142-125-3.

KOBLITZ, J. 1979. Terminological Development in the Documental Information Field and Its Respective Scientific Discipline. 9-18. See reference: GORKOVA, V. I.; BAKULINA, I. I., eds.

KOLLIN, RICHARD; KURANZ, JOHN. 1980. A Linear Business Literature Index: Merger of the Information Access Corporation and the Management Contents Databases. In: Benenfeld, Alan R.; Kazlauskas, Edward John, eds. Communicating Information: Proceedings of the American Society for Information Science (ASIS) 43rd Annual Meeting: Volume 17; 1980 October 5-10; Anaheim, CA. White Plains, NY: Knowledge Industry Publications, Inc.; 1980. 169-171. ISSN: 0044-7870; ISBN: 0-914236-73-3; LC: 64-8303.

KOROLEV, E. I. 1977a. The Language Apparatus of Automated Information Systems: Elements of a Typology. Automatic Documentation and Mathematical Linguistics. 1977; 11(3): 62-69. ISSN: 0005-1055.

KOROLEV, E. I. 1977b. The Use of the Distributive Statistical Method in the Language Apparatus of Automated Information Systems. Automatic Documentation and Mathematical Linguistics. 1977; 11(1): 31-37. ISSN: 0005-1055.

KURBAKOV, K. I.; BOLDOV, V. G. 1979. Problem of Compatibility of Information Classification Systems and Some Probable Methods of Its Solution. 150-156. See reference: NEELAMEGHAN, A., ed.

LANCASTER, F. WILFRID. 1977. Vocabulary Control in Information Retrieval Systems. In: Voigt, Melvin J.; Harris, Michael H., eds. Advances in Librarianship: Volume 7. New York, NY: Academic Press; 1977. 1-40. ISBN: 0-12-785007-4.

LANCASTER, F. WILFRID. 1979. Vocabulary Control for On/line, Interactive Retrieval Systems: Requirements and Possible Approaches. 40-52. See reference: NEELAMEGHAN, A., ed.

LANCASTER, F. WILFRID. 1980. Trends in Subject Indexing from 1957 to 2000. In: Taylor, Peter J., ed. New Trends in Documentation and Information: Proceedings of the International Federation for Documentation (FID) 39th Congress; 1978 September 25-28; Edinburgh, Scotland. London, England: Aslib; 1980. 223-233. (FID Publication no. 566). ISBN: 0-85142-128-8.

LEICHIK, V. M. 1979. Classification as a Type of Scientific or Technical Sublanguage. Automatic Documentation and Mathematical Linguistics. 1979; 13(4): 67-77. ISSN: 0005-1055.

LEONTEVA, N. N.; VISHNYAKOVA, S. M. 1977. An Experiment in Automatic Indexing with Semantic Compression. Automatic Documentation and Mathematical Linguistics. 1977; 11(3): 26-35. ISSN: 0005-1055.

LESKI, K.; LESKA, M. 1979. An Analysis of Errors in Terminology and a Concept for Terminological Control. 19-31. See reference: GORKOVA, V. I.; BAKULINA, I. I., eds.

LISTON, DAVID M., JR.; HOWDER, MURRAY L. 1977. Subject Analysis. In: Williams, Martha E., ed. Annual Review of Information Science and Technology: Volume 12. White Plains, NY: Knowledge Industry Publications, Inc.; 1977. 81-118. ISSN: 0066-4200; ISBN: 0-914236-11-2; LC: 66-25096.

LITOUKHIN, J. 1980. Toward an Integrated Thesaurus of the Social Sciences. International Classification (Germany). 1980 July; 7(2): 56-59. ISSN: 0340-0050.

LYTLE, RICHARD H. 1979. Subject Retrieval in Archives: A Comparison of the Provenance and Content Indexing Methods. College Park, MD: University of Maryland; 1979. 206p. (Ph.D. dissertation). UMI: 79-25746. Available from: University Microfilms, Ann Arbor, MI.

LYTLE, RICHARD H. 1980a. Intellectual Access to Archives: I. Provenance and Content Indexing Methods of Subject Retrieval. American Archivist. 1980 Winter; 43(1): 64-75. ISSN: 0360-9081.

LYTLE, RICHARD H. 1980b. Intellectual Access to Archives: II. Report of an Experiment Comparing Provenance and Content Indexing Methods of Subject Retrieval. American Archivist. 1980 Spring; 43(2): 191-206. ISSN: 0360-9081.

MAEDA, TAKASHI. 1981a. An Approach toward Functional Text Structure Analysis of Scientific and Technical Documents. Information Processing & Management. 1981; 17(6): 329-340. ISSN: 0306-4573.

MAEDA, TAKASHI. 1981b. A Formal Treatment of Document Information Systems. Information Processing & Management. 1981; 17(6): 319-329. ISSN: 0306-4573.

MAEDA, TAKASHI; MOMOUCHI, YOSHIO; SAWAMURA, HAJIME. 1980. An Automatic Method for Extracting Significant Phrases in Scientific or Technical Documents. Information Processing & Management. 1980; 16(3): 119-127. ISSN: 0306-4573.

MAHAPATRA, M. 1978. Syntactical Difference between POPSI and PRECIS. Libri (Denmark). 1978 September; 28(3): 235-245. ISSN: 0024-2667.

MARON, M. E. 1961. Automatic Indexing: An Experimental Inquiry. Journal of the Association for Computing Machinery. 1961 July; 8(3): 404-417. ISSN: 0004-5411.

MARON, M. E., ed. 1978a. Theory and Foundation of Information Retrieval. Drexel Library Quarterly. 1978 April; 14(2): 1-107. (Entire issue on title topic). ISSN: 0012-6160.

MARON, M. E. 1978b. Theory and Foundation of Information Retrieval: Some Introductory Remarks. Drexel Library Quarterly. 1978 April;

MARON, M. E.; KUHNS, J. L. 1960. On Relevance, Probabilistic Indexing and Information Retrieval. Journal of the Association for Computing Machinery. 1960 July; 7(3): 216-244. ISSN: 0004-5411.

MCCRANK, LAWRENCE J., ed. 1981. Automating the Archives: Issues and Problems in Computer Applications: Proceedings of a Conference; 1980 April 25-26; College Park, MD. White Plains, NY: Knowledge Industry Publications, Inc.; 1981. 363p. (Professional Librarian Series). ISBN: 0-914236-95-4.

MCGILL, MICHAEL J.; HUITFELDT, JENNIFER. 1979. Experimental Techniques of Information Retrieval. In: Williams, Martha E., ed. Annual Review of Information Science and Technology: Volume 14. White Plains, NY: Knowledge Industry Publications, Inc. for the American Society for Information Science; 1979. 93-127. ISSN: 0066-4200; ISBN: 0-914236-44-x; LC: 66-25096.

MEYRIAT, JEAN. 1980. Social Science Information Languages: A Comparative Analysis. International Classification (Germany). 1980 July; 7(2): 60-65. ISSN: 0340-0050.

MICHELL, GILLIAN. 1979. The Natural Language Foundations of Indexing Language Relations. Canadian Journal of Information Science. 1979 May; 4: 99-104. ISSN: 0380-9218.

MISCHO, WILLIAM H. 1977. A Computer-Produced Newspaper Index. Journal of Library Automation. 1977; 10(1): 41-52. ISSN: 0022-2240.

MISCHO, WILLIAM H. 1979. Expanded Subject Access to Reference Collection Materials. Journal of Library Automation. 1979; 12(4): 338-354. ISSN: 0022-2240.

MISCHO, WILLIAM H. 1980. Expanded Subject Access to Library Collections Using Computer-Assisted Indexing Techniques. In: Benenfeld, Alan R.; Kazlauskas, Edward John, eds. Communicating Information: Proceedings of the American Society for Information Science (ASIS) 43rd Annual Meeting: Volume 17; 1980 October 5-10; Anaheim, CA. White Plains, NY: Knowledge Industry Publications, Inc.; 1980. 155-157. ISSN: 0044-7870; ISBN: 0-914236-73-3; LC: 64-8303.

MORTON, SARA E. 1978. Multilingual Terminology Banks Here and Abroad; Online Terminology Dissemination as an Aid to Translation. 242-244. See reference: BRENNER, EVERETT H., comp.

MURDOCK, JOHN W. 1978. Current Knowledge on Numerical Data Indexing and Possible Future Developments. Rockville, MD: Informatics, Inc.; 1978. 66p. NTIS: PB 279 924.

MURDOCK, JOHN W. 1980. Numerical Data Indexing. Journal of Chemical Information and Computer Sciences. 1980 August; 20(3): 132-136. ISSN: 0095-2338.

MUSSO, V.; ZANGRANDO, R. 1977. The FIAT Thesaurus: An Experiment by a Multinational Company. 57-70. See reference: COMMISSION OF THE EUROPEAN COMMUNITIES, eds. Volume 2.

NEELAMEGHAN, A. 1979a. Absolute Syntax and Structure of an Indexing and Switching Language. 165-176. See reference: NEELAMEGHAN, A., ed.

NEELAMEGHAN, A., ed. 1979b. Ordering Systems for Global Information Networks: Proceedings of the 3rd International Study Conference on Classification Research; 1975 January 6-11; Bombay, India. Bangalore, India: International Federation for Documentation, Committee on Classification Research (FID/CR); Sarada Ranganathan Endowment for Library Science; 1979. 511p. (Sarada Ranganathan Endowment for Library Science Services, no. 13; FID Publication no. 553).

NEUMANN-DUSCHA, I.; ULENBERG, G. 1977. The Thesaurus of the European Brewery Convention. 439-462. See reference: COMMISSION OF THE EUROPEAN COMMUNITIES, eds. Volume 1.

NIEHOFF, ROBERT T.; KWASNY, STAN C.; WESSELLS, MICHAEL B. 1979. Overcoming the Database Vocabulary Barrier—A Solution. Online. 1979 October; 3(4): 43-54. ISSN: 0340-1545.

NIEHOFF, ROBERT T.; KWASNY, STAN C.; WESSELLS, MICHAEL B. 1980. The Design and Evaluation of a Vocabulary Switching System for Use in Multi-Base Search Environments. Columbus, OH: Battelle; 1980. 82p. NTIS: PB 80-154479.

NORRIS, C. 1981. MeSH—The Subject Heading Approach. Aslib Proceedings (England). 1981 April; 33(4): 153-159. ISSN: 0001-253x.

ODDY, ROBERT; ROBERTSON, STEPHEN E.; VAN RIJSBERGEN, C. J.; WILLIAMS, P. W., eds. 1981. Information Retrieval Research. London, England: Butterworths; 1981. 416p. ISBN: 0-408-10775-8.

PAICE, C. D. 1981. The Automation of Literature Abstracts: An Approach Based on the Identification of Self-Indicating Phrases. In: Oddy, R. N.; Robertson, S. E.; van Rijsbergen, C. J.; Williams, P. W., eds. Information Retrieval Research. London, England: Butterworths; 1981. 172-191. ISBN: 0-408-10775-8.

PAO, MIRANDA LEE. 1978. Automated Text Analysis Based on Transition Phenomena of Word Occurrences. Journal of the American Society for Information Science. 1978 May; 29(3): 121-124. ISSN: 0002-8231.

PARKER, A. H. 1981. Problems of Standardization in the CAB Database: Biological Nomenclature with Special Reference to Entomology. Aslib Proceedings (England). 1981 April; 33(4): 149-152. ISSN: 0001-253x.

PIGUR, V. A. 1979. The Linguistic Basis for the Construction of Multilingual Thesauruses, Selecting Equivalents. Automatic Documentation and Mathematical Linguistics. 1979; 13(6): 19-31. ISSN: 0005-1055.

RADECKI, TADEUSZ. 1976. Mathematical Model of Information Retrieval System Based on the Concept of Fuzzy Thesaurus. Information Processing & Management. 1976; 12(5): 313-318. ISSN: 0306-4573.

RAGHAVAN, K. S.; IYER, HEMALATA. 1978. Structuring of Compound and Complex Subjects in Social Sciences: A Users' Survey. International Classification (Germany). 1978 March; 5(1): 8-14. ISSN: 0340-0500.

RAITT, DAVID I. 1980. Aspects of Searching via On-Line Systems Using Controlled and Uncontrolled Vocabularies. International Association of Technological University Libraries (IATUL) Proceedings (The Netherlands). 1980; 12: 3-21. ISSN: 0018-8476.

RAJAN, T. N.; GUHA, B. 1979. A Comparative Study of Subject Heading Structuring According to POPSI and PRECIS. 369-381. See reference: NEELAMEGHAN, A., ed.

RICHMOND, PHYLLIS A. 1981. Introduction to PRECIS for North American Usage. Littleton, CO: Libraries Unlimited; 1981. 321p. ISBN: 0-87287-240-8; LC: 80-25977.

RIGGS, FRED W. 1979. A New Paradigm for Social Science Terminology. International Classification (Germany). 1979 November; 6(3): 150-158. ISSN: 0340-0050.

RIGGS, FRED W. 1981. Interconcept Report: A New Paradigm for Solving the Terminology Problems of the Social Sciences. Paris, France: Unesco; 1981. 49p. (Reports and Papers in the Social Sciences, SS/CH 47). ISBN: 92-3-101856-6.

ROBERTS, N. 1979. Some Thesaural Problems. Bibliotheekgids (Belgium).

ROBERTSON, STEPHEN E. 1977. The Probability Ranking Principle in IR. Journal of Documentation (England). 1977; 33(4): 294-304. ISSN: 0022-0418.

ROBERTSON, STEPHEN E.; VAN RIJSBERGEN, C. J.; PORTER, M. F. 1981. Probabilistic Models of Indexing and Searching. In: Oddy, R. N.; Robertson, S. E.; van Rijsbergen, C. J.; Williams, P. W., eds. Information Retrieval Research. London, England: Butterworths; 1981. 35-57. ISBN: 0-408-10775-8.

ROBINSON, C. DEREK. 1979. PRECIS Canada: Achievements and Prospects. Canadian Journal of Information Science. 1979 May; 4: 105-114. ISSN: 0380-9218.

ROLLING, LOLL N. 1979. Computer Management of Multi-Lingual Thesauri. 382-388. See reference: NEELAMEGHAN, A., ed.

RONDEAU, G. 1979. Bilingual and Multilingual Terminological Banks. 104-115. See reference: GORKOVA, V. I.; BAKULINA, I. I., eds.

RUSH, JAMES E.; SALVADOR, R.; ZAMORA, ANTONIO. 1971. Automatic Abstracting and Indexing—II: Production of Indicative Abstracts by Application of Contextual Inference and Syntactic Coherence Criteria. Journal of the American Society for Information Science. 1971 July-August; 22(4): 260-274. ISSN: 0002-8231.

SAGER, J. C.; MCNAUGHT, J. 1981. Guidelines for the Establishment of Comparison Matrices between Thesauri in the Social Sciences. Manchester, England: University of Manchester, Institute of Science and Technology, Centre for Computational Linguistics; 1981. (Report for the Division for the International Development of Social Sciences, Unesco). (CCL/UMIST Report 81/2).

SAGER, J. C.; SOMERS, H. L.; MCNAUGHT, J. 1981. Thesaurus Integration in the Social Sciences. Part I: Comparison of Thesauri. International Classification (Germany). 1981; 8(3): 133-138. ISSN: 0340-0050.

SALTON, GERARD; WALDSTEIN, R. K. 1978. Term Relevance Weights in On-Line Information Retrieval. Information Processing & Management. 1978; 14(1): 29-35. ISSN: 0306-4573.

SALTON, GERARD; WU, H.; YU, CLEMENT T. 1981. The Measurement of Term Importance in Automatic Indexing. Journal of the American Society for Information Science. 1981 May; 32(3): 175-186. ISSN: 0002-8231.

SCHWARTZ, RONALD; FORTUNE, JOAN; HORWICH, JULIAN. 1980. AMANDA: A Computerized Document Management System. MIS Quarterly. 1980 September; 4(3): 41-49. ISSN: 0276-7783.

SEMTURS, F. 1977. Information Retrieval from Documents in Multilingual Textual Data Banks. 463-467. See reference: COMMISSION OF THE EUROPEAN COMMUNITIES, eds. Volume 1.

SETTEL, BARBARA, ed. 1977. Subject Description of Books: A Manual of Procedures for Augmenting Subject Descriptions in Library Catalogs. Syracuse, NY: Syracuse University, School of Information Studies; 1977. 68p. ERIC: ED 142241.

SHINGAREVA, E. A. 1978. The Use of Frames for Pattern Recognition in Text. Automatic Documentation and Mathematical Linguistics. 1978; 12(4): 32-45. ISSN: 0005-1055.

SOERGEL, DAGOBERT. 1979. The Broad System of Ordering—A Critique. International Forum on Information and Documentation (USSR). 1979 July; 4(3): 21-24. ISSN: 0304-9701.

SOKOLOV, A. V. 1977. A Way to Ensure Thesaurus Compatibility (A Unition and Mathematical Linguistics. 1977; 11(1): 22-30. ISSN: 0005-1055.

SOKOLOV, A. V. 1979. Compound Indexing Languages. 394-397. See reference: NEELAMEGHAN, A., ed.

SOKOLOV, A. V., MANKEVICH, I. A.; FEDOTOVA, A. P. 1978. An Experiment in the Development of Compatible Specialized Information-Retrieval Thesauri. Automated Documentation and Mathematical Linguistics. 1978; 12(4): 8-19. ISSN: 0005-1055.

SOMERS, H. L. 1981. Observations on Standards and Guidelines Concerning Thesaurus Construction. International Classification (Germany). 1981; 8(2): 69-74. ISSN: 0340-0050.

SØRENSEN, JUTTA. 1977a. Multilingual Aspects of PRECIS. In: Wellisch, Hans H., ed. The PRECIS Index System. New York, NY: H. W. Wilson; 1977. 81-97. ISBN: 0-8242-0611-8; LC: 77-1932.

SØRENSEN, JUTTA. 1977b. PRECIS as a Multilingual System. 293-321. See reference: COMMISSION OF THE EUROPEAN COMMUNITIES, eds. Volume 1.

SPARCK JONES, KAREN. 1979a. Experiments in Relevance Weighting of Search Terms. Information Processing & Management. 1979; 15(3): 133-144. ISSN: 0306-4573.

SPARCK JONES, KAREN. 1979b. Problems in the Representation of Meaning in Information Retrieval. In: MacCafferty, Maxine; Gray, Kathleen, eds. The Analysis of Meaning: Proceedings of a Conference; 1979 March 26-28; Oxford, England. London, England: Aslib; 1979. 193-201. (Informatics 5). ISBN: 0-85142-125-3.

SPARCK JONES, KAREN. 1979c. Search Term Relevance Weighting Given Little Relevance Information. Journal of Documentation (England). 1979 March; 35(1): 30-48. ISSN: 0022-0418.

SPARCK JONES, KAREN. 1980. Search Term Weighting: Some Recent Results. Journal of Information Science (England). 1980; 1: 325-332. ISSN: 0165-5615.

SVENONIUS, ELAINE F. 1978. Facet Definition: A Case Study. International Classification (Germany). 1978 November; 5(3): 134-141. ISSN: 0340-0500.

SVENONIUS, ELAINE F. 1981. Directions for Research in Indexing, Classification and Cataloging. Library Resources & Technical Services. 1981 January/March; 25(1): 88-103. ISSN: 0024-2527.

SVENONIUS, ELAINE F.; SCHMIERER, HELEN F. 1977. Current Issues in the Subject Control of Information. The Library Quarterly. 1977 June; 47(3): 326-346. ISSN: 0024-2519.

SWANSON, E. BURTON; CULNAN, MARY J. 1978. Document-Based Systems for Management Planning and Control: A Classification, Survey, and Assessment. MIS Quarterly. 1978 December; 2(4): 31-46. ISSN: 0276-7783.

SYRACUSE UNIVERSITY. SCHOOL OF INFORMATION STUDIES. SUBJECT ACCESS PROJECT. 1978. Books Are for Use: Final Report of the Subject Access Project to the Council on Library Resources: Syracuse, NY: Syracuse University, School of Information Studies; 1978. 172p. (IST-10). Available from: Syracuse University Publications Office, 113 Euclid Ave., Syracuse, NY 13210.

TAGUE, JEAN M. 1981. User-Responsive Subject Control in Bibliographic Retrieval Systems. Information Processing & Management. 1981; 17(3): 149-159. ISSN: 0306-4573.

TAYLOR. PETER J.. ed. 1980. New Trends in Documentation and Infor-

(FID) 39th Congress; 1978 September 25-28; Edinburgh, Scotland. London, England: Aslib; 1980. 519p. (FID Publication no. 566). ISBN: 0-85142-128-8.

TKACH, S. S. 1977. Automatic Translation of Natural-Language Text into an Information Retrieval Language with a Grammatical Apparatus. Automatic Documentation and Mathematical Linguistics. 1977; 11(4): 28-36. ISSN: 0005-1055.

TOWNLEY, HELEN M.; GEE, RALPH D. 1980. Thesaurus-Making: Grow Your Own Word-Stock. London, England: Andre Deutsch; 1980. 206p. ISBN: 0-233-97225-0.

TRAVIS, IRENE L. 1981. Describing and Indexing Federal Statistical Data Systems. In: Lunin, Lois F.; Henderson, Madeline; Wooster, Harold, eds. The Information Community—An Alliance for Progress: Proceedings of the American Society for Information Science (ASIS) 44th Annual Meeting: Volume 18; 1981 October 25-30; Washington, DC. White Plains, NY: Knowledge Industry Publications, Inc.; 1981. 344-346. ISSN: 0044-7870; ISBN: 0-914236-85-7; LC: 64-8303.

VAN RIJSBERGEN, C. J. 1977. A Theoretical Basis for the Use of Co-Occurrence Data in Information Retrieval. Journal of Documentation. 1977; 33(2): 106-119. ISSN: 0022-0418.

VAN RIJSBERGEN, C. J. 1978. Automatic Classification in Information Retrieval. Drexel Library Quarterly. 1978 April; 14(2): 75-89. ISSN: 0012-6160.

VAN SLYPE, G. 1977. The Qualitative and Quantitative Characteristics of Monolingual and Multilingual Thesauri. 403-415. See reference: COMMISSION OF THE EUROPEAN COMMUNITIES, eds. Volume 1.

VASARHELYI, PAL. 1980. The Relevance of INTERCONCEPT for Classification and Indexing. International Classification (Germany). 1980 April; 7(1): 6-9. ISSN: 0340-0050.

VERDIER, VERONICA. 1979. The Use of Translation Categories in the PRECIS/Translingual Project. In: MacCafferty, Maxine; Gray, Kathleen, eds. The Analysis of Meaning: Proceedings of a Conference; 1979 March 26-28; Oxford, England. London, England: Aslib; 1979. 180-192. (Informatics 5). ISBN: 0-85142-125-3.

VERDIER, VERONICA; AUSTIN, DEREK. 1977. Research into the Translingual Potential of PRECIS. 322-337. See reference: COMMISSION OF THE EUROPEAN COMMUNITIES, eds. Volume 1.

VILENSKAYA, S. K. 1977. Information Retrieval Language Compatibility (The Intermediary Language and the Switching Language). Automatic Documentation and Mathematical Linguistics. 1977; 11(2): 54-62. ISSN: 0005-1055.

VILENSKAYA, S. K. 1980. On the Compatibility of Different Information Retrieval Languages within the Integrated Information System. In: Taylor, Peter J., ed. New Trends in Documentation and Information: Proceedings of the International Federation for Documentation (FID) 39th Congress; 1978 September 25-28; Edinburgh, Scotland. London, England: Aslib; 1980. 315-325. (FID Publication no. 566). ISBN: 0-85142-128-8.

VLADUTZ, G.; GARFIELD, E. 1979. KWPSI—An Algorithmically Derived Key Words/Phrase Subject Index. In: Tally, Roy D.; Dueltgen, Ronald R., eds. Information Choices and Policies: Proceedings of the American Society for Information Science (ASIS) 42nd Annual Meeting: Volume 16; 1979 October 14-18; Minneapolis, MN. White Plains, NY: Knowledge Industry Publications, Inc.; 1979. 236-245. ISSN: 0044-7870;

VLEDUTS-STOKOLOV, NATASHA. 1981. On Automatic Support to Indexing at BIOSIS. In: Using Information: Papers contributed to the American Society for Information Science (ASIS) 10th Mid-Year Meeting; 1981 May 14-16; Durango, CO. Englewood, CO: Information Handling Services; 1981. Microfiche; fiche 1; page 12; 24X reduction. (Abstract appears in hard copy on page 1).

WALL, R. A. 1980. Intelligent Indexing and Retrieval: A Man-Machine Partnership. Information Processing & Management. 1980; 16(2): 73-90. ISSN: 0306-4573.

WEINBERG, BELLA H. 1981. Word Frequency and Automatic Indexing. New York, NY: Columbia University; 1981. 221p. (Ph.D. dissertation). Available from: the author, 2840 Shady Ave., Pittsburgh, PA 15217.

WEINTRAUB, D. KATHRYN. 1979. An Extended Review of PRECIS. Library Resources & Technical Services. 1979 Spring; 23(2): 101-115. ISSN: 0024-2527.

WELLISCH, HANS H., ed. 1977. The PRECIS Index System: Principles, Applications, and Prospects: Proceedings of the International PRECIS Workshop; 1976 October 15-17; College Park, MD. New York, NY: H. W. Wilson Co.; 1977. 211p. ISBN: 0-8242-0611-8; LC: 77-1932.

WELLISCH, HANS H. 1979. The Broad System of Ordering, or Bishop Wilkins Redivivus: A Review Article. Library Quarterly. 1979 October; 49(4): 444-452. ISSN: 0024-2519.

WELLISCH, HANS H. 1980a. Indexing and Abstracting: An International Bibliography. Santa Barbara, CA: ABC-Clio Press; 1980. 308p. ISBN: 0-87436-300-4.

WELLISCH, HANS H. 1980b. Some Vital Statistics of Abstracting and Indexing. International Classification (Germany). 1980 November; 7(3): 135-139. ISSN: 0340-0050.

WERSIG, GERNOT. 1979. Experiences in Compatibility Research in Documentary Languages. 423-430. See reference: NEELAMEGHAN, A., ed.

WILSON, PATRICK. 1979. Utility-Theoretic Indexing. Journal of the American Society for Information Science. 1979 May; 30(3): 169-170. ISSN: 0002-8231.

YU, CLEMENT T.; SALTON, GERARD. 1977. Effective Information Retrieval Using Term Accuracy. Communications of the ACM (Association for Computing Machinery). 1977 March; 20(3): 135-142. ISSN: 0588-8069.

YU, CLEMENT T.; SALTON, GERARD; SIU, M. K. 1978. Effective Automatic Indexing Using Term Addition and Deletion. Journal of the Association for Computing Machinery. 1978 April; 25(2): 210-225. ISSN: 0004-4511.