Book Reviews

UNGVARY, Rudolf (Ed.): **OSZTAURUSZ. Periodika osztalyozási rendszerés tezaurusz.** ((A Universal "class-thesaurus" for Hungarian Periodical Literature). Budapest: Országos Széchényi Könyvtár 1990. 131p.

A new combined universal classification system and thesaurus came to light in Hungary, aimed at providing subject access to periodical publications. It is existing as a printed version (which is the object of this review) and as an online thesaurus as well, for the national information system, containing the Hungarian Periodical Literature.

The editor of this comprehensive work is Rudolf Ungvary, whose great knowledge and insight into the problems of classification and indexing has clearly contributed to the successful outcome of this tedious project. The OSZTAURUSZ has been under preparation since 1984 at the Hungarian National Library (Országos Széchényi Könyvtár) and the collection of terms was based on data available in the national database for periodical literature and the indexing vocabulary of the National Technical Information Centre and Library (OMIKK). The online thesaurus is running on CDS/ISIS software and its Micro version.

This new retrieval tool, named as OSZTAURUSZ in Hungarian, is an amalgamated combination of a universal classification system *and* a thesaurus. It provides multiple access points at two conceptual levels:
- descriptors (formulateld in natural language form)
- classification codes/notation (two alphabetical signs, combined with numerical ones).

In practical terms, this means that beside each descriptor the corresponding classification code (notation) is indicated. Thus, in spite of its universal scope, the system is simple; it contains relatively few classes, but at the same time, it is flexible enough to provide *specified* subject access to periodical literature, e.g., 1-5 classification codes can be used for one document. The degree of specificity in OSZTAURUSZ could be examined from many different points of view, but since it is designed to be used as a metathesaurus, its present size and structure is suited to its purpose. This standardized vocabulary system is aimed at serving as the basis for the development of other specialized thesauri, as well as at functioning as switching instrument between these thesauri.

The combined Hungarian "ClassThesaurus" (some prefer the term "Classaurus") contains 194 classes, 209 geographical names and some elements of facets. It has a hierarchical construction with a maximum of four levels which indicate the usual types of relationships between the terms such as generic, part-whole, causal, topical. The notation does not always reflect the hierarchical relationships, it is therefore of a rather enumerative nature. There is probably a good reason why the Hungarian double consonants are not treated as single phonetical signs but are composed of two separate characters in the alphabetical list of terms. The fact that "csillagászat" is listed prior to "cukrászat" might cause confusion for many Hungarians in browsing and searching cases.

Irene Wormell

Dr.I.Wormell, Prof., Royal School of Librarianship, Copenhagen, Denmark

Indexing Tradition and Innovation. Proceedings of the 22nd Annual Conference of the American Society of Indexers, June 27, 1990. Chicago, IL: American Society of Indexers 1991. 123p., ISBN 0-936547-13-8.

This book is a compilation of the papers read at the 1990 annual conference of the American Society of Indexers. "Indexing tradition" is the dominating theme among the contributions, with primary attention being given to experiences from the practice. Only one contribution, namely the final one, in the book deals with "innovation": it describes an attempt at an innovation in the field of automatic thesaurus compilation and indexing. The contributions preceding it present a vivid picture of the requirements imposed on indexing in actual practice outside the experimental lab. This is particularly true of the two quite voluminous contributions by Ann P. Bishop, Elizabeth D.Liddy and Barbara Settel. Closely interrelated as they are, they are discussed jointly below. The initial contribution deals with indexing standards and in so doing makes apparent the problems inherent in such standards.

Jessica MILSTEAD: *Indexing Standards.* In order to justify the use of standards in the information field, the authoress starts out with several examples from the field of crafts, where the usefulness of standards is particularly evident. In the information field, too, a few convincing examples are given to demonstrate that standardization makes good sense here. Next, a list of the most important British and American standards organizations, also including the International Standards Organization at Geneva, as well as a description of the standards-making process are given.

Five groups of standards of importance to the indexer are recognized, namely standards for:
- *Contents and structure of bibliographic records.*
- *Standard numbers and codes for books, serials, technical reports, book industry, serial issue, and article identifyers.*
- *Electronic information exchange: reducing the number of specific languages, record formats, etc.*
- *Romanization and transliteration: systems for converting writing systems and alphabets.*

- Miscellaneous, e.g. proof corrections, information on fiche headers and book spines.

They are discussed in detail with all their strong and weak points.

In this reviewer's opinion, however, the standards presented include some which go beyond purely formal aspects, and here some skepticism as to their usefulness is called for. For one thing, on the subjects of these standards research and development is in full swing, and for another, the heterogeneity of the requirements and possibilities presenting themselves in practice in the numerous individual cases is so great that it will hardly be possible to formulate useful, uniformly applicable rules for such matters as e.g. indexing. Where this is done nevertheless, the standards concerned do more harm than good, impeding as they do progress and easily leading the user astray. The full extent of the problem is expressed by the authoress where she writes on the one hand: "Standards are not textbooks, they are not intended for teaching purposes...", and on the other hand: "A user should be able to assume that if he or she conforms to an appropriate standard, the end product will be of appropriate quality..." (p.9). But how can the indexer assume to have done a job of appropriate quality according to the standard if the standard is to be denied the qualification of a textbook, although dealing nevertheless with the material of textbooks and doing this in its typical inflexible manner, which in the interest of the primary objective, namely compatibility, is even markedly one-sided at that? If a given standard lacks e.g. any reference to the capabilities of an indexing language grammar - which should be the subject of a textbook chapter of its own -, then the practical indexer wrongly feels entitled to dispense with such a grammar in every case.

Standards should therefore not be concerned with the subject matter taught in information science. Rather they should limit themselves to their proper function of achieving and maintaining compatibility where it is really meaningful. But as regards research and information science, compatibility leads here to stagnation and paralysis. The new is always incompatible with the traditional. When a doctrine handed down by tradition has become the object of a standard, any innovation can easily (and to superficial or uninformed listeners even convincingly) be criticized by virtue of the sole claim that it violates the standards in force.

From the chapter on standard-making processes it can be seen that standards are adopted by means of majority votes (p.4), a procedure which may prove markedly counterproductive inasmuch as the panelists concerned often fail - or are even unable or unwilling - to appreciate the meaning and value of an innovation.

Standards, it is said further, should be revised every five years. But this, in this reviewer's experience, does not always guarantee that the standards will be up to date. Here we find the words of Nobel physicist Max Planck confirmed according to which a new school of thought often will gain the upper hand only as the proponents of the old school die out rather than becoming convinced of the new manner of thinking.

Ann P. BISHOP, Elizabeth D. LIDDY, Barbara SETTEL: *Index Quality Study, Part I: Quantitative Description of Back-of-the-Book Indexes and Part II: Publisher's Survey and Qualitative Assessment.* The basic purposes of the study were

- To describe characteristics and features of recent back-of-the-book indexes;

- To assess the quality of back-of-the-book indexes, particularly in terms of subject access and ease of use;

- To describe current publishers' policies related to the production of back-of-the-book indexes.

In a detailed and instructive introduction the tasks are set forth which a good back-of-the-book index must fulfill. The large differences are explained which result from the requirements imposed in any given case, also as a function of the searching aids (printed vs. electronic indexes), of the type of the subject field, the type of the index (author and/or subject index), of the resources as to time, skill, subject knowledge and effort available or expended on the part of the indexers, of the space assigned in the book, furthermore as a function of the totality of things to be covered in the book (hence of the scope of the index), of the type of concepts to be covered (proper names or subjects), and also as a function of the costs allowed by the publisher or author. No absolute valid rules can be given, only general recommendations.

More than 400 fairly recent books from the following fields: humanities, fine arts, social sciences as well as science and technology, were examined according to the following quality criteria:

- Index scope, length, and density; number and type of headings

- Number of "see also" references and subheadings; precoordination

- Number of "see" references; locators

- Number of explanatory expressions

- Number of multiple entries

- Arrangement of indexes, headings, subheadings, and locators, physical appearance

- Introduction to the index.

Recommendations are given as to how indexes should be drawn up so as to be optimally adapted to the constellation of requirements and factualities encountered in any given case.

The main effort involved lies in correlating the language of users with that of the index and in the mapping of a hierarchy of concepts underlying in particular the "see also" references. A sound subject knowledge, also of the given special field and of its terminology, are required for this, as well as an appreciative understanding of the situation of the user, who often will have read the book only superficially or not at all. Such knowledge is likewise required for a systematic arrangement of the index, only rarely though it is encountered so far.

From the nature of things, the uniqueness of back-of-the-book indexes rules out any direct comparisons among

them. Instead, each book examined was subjected, for the purpose of uncovering hidden index shortcomings, to a partial re-indexing limited in each case to 10 pages of text. The shortcomings uncovered consisted above all in inadequate subject analysis, resulting in important topics not being covered, not even e.g. when they re-occur elsewhere in the book. Also, the interlinking of related topics was generally inadequate. Proper names predominated over subject designations, and adequate subheadings were often lacking.

This reviewer suspects that we are confronted here with an effect of the widespread ignorance among publishers, authors and software firms of the cognitive rank of indexing. Such a result, namely, will easily come about through the application of current commercial indexing software unless the result obtained thereby is meticulously gone over once more and improved where needed. Such software orients itself exclusively to the keywords encountered in the given text. Also, the application of extratextual knowledge, necessary e.g. for expert subject analysis and for constructing the network of references, is dispensed with. - Regrettably, the ignorance signalled in the above is still, and continuously, being spread further by indexing software advertising.

As regards the linking together of dispersed, but related topics with the aid of their locators it would have been helpful, in this reviewer's opinion, if detailed recommendations for the use of the tool (classification or thesaurus) required to this end had been given.

Phyllis DICKSTEIN: *United Nations Documents - Approaches to Indexing.* As an introduction to the problems inherent in the documentation of the voluminous and thematically most heterogeneous literature available at the United Nations, the tasks and objectives of the UN are outlined first. Heterogeneity prevails here also in the manner of linguistic expression, not only because of the variety of national languages, but also because of the great differences developing in the course of time for one and the same subject even in one and the same language. It must be possible to relate all documents on the same subject or on closely related subjects to one another as a basis for the decision process.

Exhaustivity, specificity and consistency in the selection and presentation of subjects as well as formal compatibility must, in indexing, always be assured or be mutually harmonized. As basis for the indexing process a suitably detailed thesaurus is employed, together with a strict body of rules for its application. On the maintenance of the thesaurus, particularly on its expansion as necessary, useful rules are given which are also of general interest. A description is given of the compromise between single descriptor indexing and precoordination which had to be struck because of the necessity both of a printed and of an electronic index. - From this contribution, useful recommendations of proven worth in practice may be obtained for any type of parliamentary documentation.

Diane ULLIUS: *What Does un Editor Want?* The expectations which editors entertain with respect to their inde-

xers vary widely. Much remains unsaid which should have been made explicitly clear before the start of the indexing work. The best thing for editors and indexers is to regard their work as a joint venture and to communicate interactively, as it were, on the best path toward the common goal.

The authoress is an editor herself and reports from her experiences. She distinguishes between various types of editors: the fearful, the deluded, the irresponsible, the monstrous, and the good. With the latter, good cooperation in an interactive way, as it were, is possible. With him, above all, nature and extent of the work are jointly estimated and agreed upon *prior to* the start of the work. This paper grants further insights into the work of the editor and his worries, and it helps indexers and editors to find a good path based on partnership.

Amy J. WARNER: *An Automatic Method for Vocabulary Control.* This paper deals with a research project in an early stage of development which so far employs only low-level linguistic techniques and is limited to the processing of phrases encountered in the document titles. An unspecified number of such phrases is selected from document titles with the aid of keywords and is subjected to parsing. Each word of a phrase is subjected to morphological analysis to obtain word stems.

Because of the project description given in this paper, the hopes of its author concerning the future prospects of this approach cannot yet be shared by this reviewer.

This Proceedings volume constitutes an essential enrichment of the available teaching material on indexing. This is true particularly of the contributions from everyday indexing practice and of the recommendations derived therefrom. The latter are submitted in a dense concentration and with professional competence. They are most instructive to all those undertaking to start an indexing project, to revise it or even to automate it by means of algorithmic methods. The reading of this volume can be recommended to all of them. Particularly the authors of the purely algorithmic projects will then realize how far they are still away from being able to hold forth, with a clear conscience, hopes of useful automated indexing procedures or even to offer such procedures.

Robert Fugmann

Dr.Robert Fugmann, Alte Poststr.13, D-6270 Idstein.