BOOK REVIEWS

WERSIG, Gernot: Thesaurus-Leitfaden. **Eine Einführung in** das Thesaurus-**Prinzip in Theorie und Praxis.** (Thesaurus Guideline. An introduction into the thesaurus principle in theory and practice). München–New York: Verlag Dokumentation Saur KG 1978. 346 p. = DGD-Schriftenreihe Bd. 8, ISBN 3-7940-3628-X.

This book is directed at persons involved in the planning or implementation of thesaurus development and maintenance. Chapter 1 reviews briefly the author's views of the conceptual basis of indexing and various types of index languages. The relationship of thesauri to other types of index languages on the one hand and to other types of vocabularies (e.g., dictionaries) on the other is explained. Chapter 2 discusses the role of a thesaurus as a tool for terminological and conceptual control. Chapter 3, Thesaurus Structure, deals primarily with the structure and form of individual terms used in the thesaurus and the relationships between these terms. Chapter 4 discusses the parts of a thesaurus and the various options for their design. Chapter 5 details the process of thesaurus development and maintenance. Chapter 6 gives a typology of thesauri and reproduces sample pages from a thesaurus of each type.

This book is useful as a source of complementary information for those involved in the development of a thesaurus, particularly a German language thesaurus, but as a guide to be used alone it falls short. On the positive side is the wealth of examples illustrating problems of conceptual relationships and of word forms (particularly in German) and the rules given in standards, used in some thesauri, or suggested by the author to solve these problems. The problem illustrations are useful even if one does not agree with the solutions given. Likewise, there are many examples of thesaurus pages illustrating various possibilities of thesaurus arrangement. In Chapter 5 there is a useful section on the overall planning of thesaurus development which emphasizes the need to specify clearly for each phase the inputs needed and the outputs produced. The general planning methodology should be followed even if one does not agree with the process sequence suggested in the remainder of the chapter.

The shortcomings of this book can all be traced to insufficient depth in the analysis of the problems treated. There are four major points of criticism:

(1) The very definition of thesaurus itself is questionable. The thesaurus of the (New York Times) Information Bank is considered a thesaurus; the structurally very similar Library of Congress Subject Headings are not. The criteria that are used to set off a thesaurus from other index languages discussed in Chapter 1 may not be the most fundamental for characterizing different types of index languages. Moreover, it is questionable whether it is useful to limit a book such as this to a particular type of index language when the problems discussed are to a large degree common to all types of index language.

(2) There is little or no reflection about the functions of an index language/thesaurus in the total context of an information storage and retrieval system, particularly in indexing and search request formulation, nor is there a reference where such discussion could be found. On the function of a thesaurus in general, we learn: "The most important function of a thesaurus is to achieve fulfillment of the requirements of the information system in which it is used or intended to be used." To this reviewer fulfillment of the information needs of those to be served by the system is the ultimate function. But even with that substitution the statement does nothing to enlighten the reader. The chapter on thesaurus functions is devoted almost exclusively to the means by which the objective of a thesaurus within an ISAR system is achieved, namely, terminological control, conceptual control, and orientation of the thesaurus user in the terminology and concepts of the area covered by the thesaurus.

(3) The discussion of the elements of the thesaurus vocabulary, namely, descriptors and non-descriptors, and of the relation between them consistently confounds the plane of terms (German: Benennungen) and the plane of concepts to the detriment of both the quality and the clarity of the rules. Failure to distinguish explicitly between preferred terms (terms selected for the unique and unambiguous designation of a concept) and descriptors selected from among these preferred terms leads to a rather convoluted treatment of the "use combination of" cross-reference (p. 116/117). A closely related point is the deficient treatment of semantic factoring which is part of the section on terminological control rather than part of the section on conceptual relationships where it belongs. Significantly the book talks about semantic factoring of terms (Benennungen) and great term length is given among the reasons why semantic factoring should be considered in a specific instance. However, the book does clearly state that factoring must be based on semantic, and not morphological, considerations. Unfortunately, some of the book's own rules do not follow this principle. There is another defect in the treatment of semantic factoring. It is considered only as a means for reducing the number of descriptors; i.e., it is recommended only if it is decided that a very compound concept should not be used as descriptor and that a combination of two or more existing (and less compound) descriptors should be used instead. However, semantic factoring and the related principle of facet analysis are very important tools in clarifying the structure of a subject field and detecting relationships between concepts. Therefore, all compound concepts should be factored semantically, whether they are used as descriptors or not. In fact, the decision about descriptor selections should be made after semantic factoring has been applied and not the reverse.

(4) The suggested process for thesaurus development consists roughly of the following steps: (1) Collection of terms. (2) Elimination of duplicates, elimination of terms not be included in this thesaurus, first detection and consolidation of synonyms (only cases that can be easily decided). (3) Thorough examination of all terms, detection and consolidation of synonyms and quasi-synonyms (difficult cases), decisions about homonyms and polysemous terms, selection of descriptors. (4) De-

https://doi.org/10.5771/0943-7444-1980-2-90 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

velopment of the conceptual structure. This is backwards. The conceptual structure and the synonym-homonym structure should be developed hand in hand: while terms are arranged first in a coarse and then in a fine classified order, synonyms and quasi-synonyms are brought together. For meaningful decisions on descriptor selection a classified order is essential.

These basic flaws alone would make the book unsuitable as a guideline for thesaurus development. The book contains numerous other errors and bad advice. We will mention here just a few. (1) It is stated that thesauri with a classified main part (as opposed to those using an alphabetical main part) are more or less constrained to monohierarchy. Not so. It is correct that the arrangement in the classified main part can express only a monohierarchical structure (double listing of full entries would not be economical). However, all hierarchical relationships that are deemed useful can be expressed through BT/NT cross-references just as in any other arrangement of the main part. (2) It is claimed (page 158) that there is no known thesaurus which shows multi-word terms under the component words in the main part. There may not be a thesaurus that does so with the same degree of completeness as a KWIC or KWOC index, but there are a number of thesauri that have many such entries in an alphabetical main part, for example, Thesauro-facet, the Preliminary Thesaurus for Documentation on Socio-cultural Problems of Developing Countries, and Library of Congress Subject Headings. (3) In the list of indicators for cross-references and other data elements given in main part entries (page 164 through 166) the column labeled SOE contains numerous mistakes. (4) In part 6 various sample thesauri are described but not really analyzed. In the description of the EURATOM thesaurus the reader is misled through not quite accurate translation: "Accepted terms", which in the EURATOM system are descriptors in the generally accepted sense, is translated to the German equivalent of "accepted non-descriptors". The mistake comes about because EURATOM distinguishes two types of descriptors, "Key words" and "Accepted terms".

The style of the book is shoddy in places, and sentences are replete with unnecessary words such as "in principle" (German: grundsätzlich), another indication that the book is written in a somewhat superficial manner. The author himself says as much in the preface and only claims that "a certain usefulness should not be ruled out".

For the knowledgeable and discerning reader who can separate the wheat from the chaff the book certainly has its uses, but the novice is apt to be led astray. Whoever wants to use this book should pay particular attention to Chapter 7, which is quoted here in its entirety: "Of course one can do everything altogether differently, provided one does it right." Dagobert Soergel

SERRAI, Alfredo: **Del catalogo alfabetico per soggetti. Semantica del rapporto indicale.** (On the alphabetical subject catalogue. Semantics of indexing.) Roma: Bulzoni 1979, 163 p.

In pursuance of the objective of building up theoretical foundations for bibliographic procedures and for librarianship practices, Alfredo Serrai is now facing the most complex and difficult part of a cataloguer's work, namely subject-indexing.

Having discarded as naive the possibility of setting up a body of rules for this operation, the conditions in which it is carried out are investigated. The kernel of the semantic indexing process is the defining of meaning, which, in the light of the modern language and communication theories, is viewed in its relational nature. Subject indexing implies a reduction of the global meaning of the document – which is anything but obvious – and its translation by means of a linguistic formula aiming at serving as a key-link between the searcher and the document. Subject indexing and information retrieval are both made more complex by the absence of a univocal relationship between the concept and its translation into language. The choice of the subject is based on intensional conceptual interpretations and transformations which, through the index, are translated into formula referring to individuals or classes which are of the extensional type. The reader using the catalogue follows the same path in the opposite direction.

After having analysed the operating conditions, the author considers the various subject-indexing stages, and stresses some important distinctions that the librarian should be acquainted with, and that he should comprehensively master, so as to avoid misunderstandings and equivocations: the distinction between *subject* and *object* of the document, that is between the meaning and the signified, between *subject* and *descriptor*, and finally the distinction between *subject* and *discourse*. The analysis of the indexing processes is interwoven with an investigation of the subject indexing theories that started with Cutter, a pioneer wo for many aspects has not yet been exceeded, but who has very often been misunderstood or dangerously simplified.

In Section II of the book, the author rapidly surveys – or rather dismisses – the Italian literature on subject-indexing, from Fumagalli to Revelli through the theories framed by the "Vatican Rules" and the printed "Subject Index" of the National Library in Florence. This is followed by the "Proposal", made by the Author for attempting to solve the problem of the scattering effect of the subject catalogue as well as of the ensuing loss of information, and to anchor subject-indexing to a sound reference framework. Cutter had clearly perceived this need, as proved by his stressing the importance of a network made of downward and upward references, even though he had not managed to put them into practice. Mr. Serrai suggests to anchor subjects to rigorously controlled structures of terms, a sort of microthesauri, bound to become the common reference maps of librarians and researchers; they are the semantic frameworks within which the manifold aspects of knowledge are orderly separated into fields of interest. At the basis there is, therefore, a precise but polyvalent organization that would provide the index with the advantages of the traditional subject-catalogue (easy access, possibility of gathering together the various aspects of a subject) and with those of the systematic catalogue (soundness of framework capable of overcoming the fragmentary nature of the subjects). The major drawback of setting inflexible, pre-established relationships among the composing elements would thus be avoided.

https://doi.org/10.5771/0943-7444-1980-2-90 https://www.inlibra.com/de/page/agb - Open Access -