

view, knowledge presentation would then mean a thesaurus-like presentation of knowledge structure, and the processing of knowledge the reshaping of knowledge structures.

UNGVARY's contribution, "The use of the thesaurus method in knowledge conveyance", runs on similar lines. Ungvary, however, confines himself primarily to the theoretic aspects of the use of the thesaurus method in knowledge "conveyance" and does not go into the problem that, for example, good use may also be made of thesaurus changes in the acquisition of knowledge. A discussion of connexions to other possibilities of applying the thesaurus principle, for example in prognosis (Dobrow) or problem solving (Bauer), would have been useful.

The wide thematic spectrum of the contributions in this volume with all the questions that are touched on, but still open is, at the same time, an appeal to promote classification research even more than has been the case up to now.

"But classification training itself should be improved and coordinated in all areas; above all, essential fields of classification science should be offered to students of all disciplines at colleges, universities and training centres as a subject which not only teaches methods of order, but also gives a very necessary overall view of the various fields of knowledge, which draws the individual out of the isolation inherent in a specialised field of knowledge, and which gives him again insight into the general view of knowledge" (Dahlberg).

In this review it was only possible to go into a few of the particularly acute problems and focal points taken from the overall theme of the volume. As already stated in the Preface, the problems presented here have, for the moment, only been laid open for discussion. But one thing is sure: this volume should be stimulating for all specialists, well beyond those who belong to the circle members of the classification society.

Gerd Bauer

Dr.rer.nat.Dr.sc.phil.G.Bauer
Am Treptower Park 50, GDR-11 93 Berlin

MANIEZ, Jacques: Les langages documentaires et classificatoires: conception, construction et utilisation dans les systemes documentaires.

(Documentary and classification languages: Their conception, construction and use in documentary systems).

Paris: Les Editions d'Organisation 1987. 291p. ISBN 2-7081-0833-6.

Jacques Maniez's book and that of G. Van Slype (Indexation Languages: Design, Construction and Use in Documentary Systems) published in the same year and by the same publisher, are a joint project dealing with documentary languages in general, and principally intended for professionals, though of possible interest to readers interested in information techniques and methods of content analysis.

The major portion of J. Maniez's book thus deals with classificatory languages. The other portion discusses documentary languages, which are presented in the light of linguistic theory rather than from a practical and descriptive point of view. The advantages and drawbacks of documentary languages are related to the properties

of natural language. J. Maniez uses his competencies as a linguist to provide an in-depth analysis of the principal relationships used in documentary languages (synonymy, hierarchy and association) as well as the operations of characterization and condensation of a document and research in automated documentation.

In the principal portion of J. Maniez's work, the first chapter gives us a definition of important terms and presents a typology of classification systems. In the second chapter, the author shows us how to use these tools and describes those which are the most well-known (Dewey Decimal and Universal Decimal Classification, Colon Classification, Library of Congress Classification and the Bliss Classification). Finally, an interesting chapter is devoted to the construction and maintenance of these classificatory languages.

As shown in the above description of these different chapters, J. Maniez's book is of interest at different levels:

(1) The practicing classificationist whose knowledge was principally acquired through practical experience will find in this book a theoretical presentation which will allow him to better understand the basic principles of the design and use of classification systems. Certain empirical rules which he discovered through trial and error will here find their justification - if not theoretical, then at least within the framework of a coherent set of practical considerations. On numerous occasions, the author's linguistic background enables him to explain and criticize certain rules in current use. For example, he presents the notion of facet by comparing it to case in a language, thus allowing for a better understanding of the concept, as well as the advantages and drawbacks of its use. J. Maniez does not hesitate to present projects which are still in a research stage, such as G. Salton's famous SMART system. The author's presentation of the foundation of automated classification techniques is in itself a lesson in pedagogy.

(2) The second level will be useful to those who, having little knowledge of classification systems, will find the essential elements needed for an understanding of this field. The main classification systems are presented in a sufficiently precise manner to allow the reader to understand their interest and their functions. Incidentally, the classificatory techniques are covered with true pedagogical talent by the author. Each classification system is presented in a highly structured form which successively describes the principles, the working instruments, the use of these instruments by the classifier and by the reader, as well as a critical conclusion. Extracts of tables and indexes as well as precise examples provide a very concrete and practical understanding of these tools.

(3) The last level, to be found in chapter III, will be extremely valuable to those who must design and set up a classification system in a library. This chapter surely sets this book apart from others that are limited only to explanation or description. Placing himself in a familiar context which corresponds to a specialized library of approximately 6000 works, J. Maniez constructs a veritable case study which allows us to determine the characteristics of two classification systems which can answer the needs of this case (the final choice is left to the head of the establishment).

A well-focussed bibliography is located at the end of each chapter. The work includes an index of terms and an index of proper names.

Richard Bouche

Prof.Dr.R.Bouche, Lab.d'Informatique Documentaire, ISIDT
Universite Claude Bernard - Lyon
43, Boulevard du 11 Novembre 1918, F-69622 Villeurbanne,
Cedex, France.

AITCHISON, Jean; GILCHRIST, Alan: **Thesaurus Construction. A Practical Manual.** 2nd ed. London: Aslib 1987. 175p., 41 diagrams. ISBN 0-85142-197-0

The requirements necessary for the construction and maintenance of a thesaurus are discussed in thirteen sections wherein the relevant mental and technical measures concerned are gone into. The frame for these reflections is formed by a lucid survey of the specific weaknesses and strengths of natural language and index languages. From these, the necessity of an index language is derived in as far as a reasonable degree of search accuracy is to be attained using such an information system. Today, the thesaurus as the vocabulary of such an index language is a widely used variant. In the very first diagram it is shown in which specific way both kinds of language effectively complement each other. This, at the same time, answers the cardinal question of choice between both kinds of language, a question which is always posed today when a new information system is developed. This lucid comparison not only deserves the first place among all the illustrations given, the reader should also keep it in mind in the further course of these reflections.

When, further on in the text, the variants of the management of texts without the use of a thesaurus and of indexing are discussed, these seem, at times, to be based on the publications of those research groups that are active in this field, and their often far too great optimism and too lenient evaluation of the deficiencies in such methods show through again and again. This is particularly true of the different variants of statistical procedures; the deficiencies in these methods, when they boil down to the mere counting of and searching for words instead of for concepts, were recently exposed in the studies of E.Liddy, S.Bonzi, J.Katzer, and E.Oddy (J.ASIS 38(1987)p.255-261). In other places, too, e.g. when discussing the abandoning of all indexing, the reader is not always reminded of the negative consequences of doing so.

Section J, the last section, which rounds off the frame, describes details of procedure in constructing a thesaurus, illustrated by examples taken from the field "Catering". This particular chapter can be recommended to anyone who undertakes the construction of an information system as it will give him an impression of the manner and magnitude of the task he has taken upon himself. This example not only gives an exemplary picture of the single steps to be taken, it is also exemplary in its choice in as much as it gives a model for those prerequisites, demands, and possible ways to a solution such as are likely to be met with in many other cases. The great practical experience of the authors is also reflected in the many well-chosen, convincing examples used in other parts of the text. For the majority of the

measures discussed, the authors not only take an objective look at the advantages, but also at the specific disadvantages. The reader himself must decide what weight he will give to these conclusions in order to come to his own decision.

The great uncertainty prevalent in the field of concept analysis is naturally reflected in a practical manual of the kind under review. For example, semantic factoring, on the one hand, is rejected in the literature quoted here, on the other hand, however, this operation is an essential element in any analytic-synthetic approach, i.e. an approach which - with a correspondingly high expenditure - no doubt achieves the highest degree of effectiveness.

In the manual, the morphological form of analysis dominates, i.e. an analysis oriented to the surface structure of a linguistic mode of expression rather than to contents and meaning. This means that the component parts of a concept are made explicit only when they have been isolated in the linguistic formulation encountered in the text to be indexed. Here it becomes apparent that little experience had been gathered in the use of this approach when it was introduced in many places for practical application and that there are a great number of impediments to taking the appropriate and logical steps quickly in an information system when the disadvantages of an approach become evident. It would be of great advantage to designers of information systems if such deficiencies could be spotted at an early stage.

The literally superficial character of many traditional approaches is even set down in the standards (and is still awaiting revision there!). This is also reflected in the recommendations quoted in the book: that one must pay attention to the natural language word form in making descriptors. In these rules, much importance is attached to whether a conceptual feature is expressed as an adjective or not, although it is left to chance or personal taste whether one says "oxidative splitting" or "splitting oxidation" or "oxidation with simultaneous splitting". Whether a descriptor counts as a "compound term" is judged only according to the number of single words contained in it and not according to its conceptual complexity. Commendably, the authors point out that it is often imperative to turn one's back on standards.

Furthermore it should also be mentioned that a thesaurus as a vocabulary is always only one part of an index language. Here, and in the reviewer's opinion, many a gap and inadequacy found even in using the best thesaurus is to be ascribed to the lack of a grammar, the natural second component in any language which is rich in expression. Any thesaurus will - in the long run - be overtaxed if it is to take on the tasks, too, which are really those of an index language grammar.

The present uncertainty in information science is also reflected in the fact that the supposed "lawful inverse relation of precision and recall" is quoted again although in the meantime a great number of counterexamples have become known in which loss of precision was not the inevitable consequence when recall was increased, and vice versa. On the contrary, total recall and, at the same time, total precision could be achieved.

Another unresolved contradiction frequently encountered is reflected in the book: Loss in recall is said to be