

than two pages, a very useful list of the "most salient points on vocabulary control that this book has attempted to illustrate" (p. 123), cross-referred to the appropriate chapter(s).

The critical comments on individual points should not detract from the intrinsic value of the book. Based on his vast experience the author provides a useful overview of the structure and display of vocabularies, of methods for their construction, and of experiments that were intended to clarify the role of the vocabulary in the performance of reference storage and retrieval systems. Although this review has disputed a number of points, the author presents tenets that are widely accepted. In this sense the critical comments in this review reflect really a controversy in the field. In summary, the book is a significant contribution to the literature of information science.

Dagobert Soergel

SOERGEL, Dagobert: *Indexing Languages and Thesauri: Construction & Maintenance*. Los Angeles: Melville 1974. XXXIX, 632 p. ISBN 0-471-81047-9, A Wiley-Becker & Hayes Series Book.

This volume deals with the characteristics and construction of controlled vocabularies. It is very complete and, by and large, extremely accurate. The contents are divided into four major areas: the structure of indexing languages, methods by which such vocabularies are arranged and presented, methods by which they may be constructed and maintained, and the use of thesauri as the basis of cooperation among information services. A novel feature of this book is that it presents the material at various clearly defined levels. A reader who wants only a general understanding of indexing languages need read only designated sections of the work. Other sections are marked as "technical", "special" or "advanced". These need be read only by those who want a deeper understanding of the subject or who wish to extract information relating to a special problem area. The way the volume is structured, then, makes it more suitable for use as a handbook — a volume to consult when we need to find out about a particular aspect of vocabulary control — than as a textbook or as a series of chapters to be read consecutively. Soergel, however, would like to think of it as both a handbook and a textbook.

Viewed as a handbook, the work is excellent. I find myself in complete agreement with much that the author says. There is a great deal of common sense here, and the author strips away the unnecessary mystique that surrounds much of the other writing in this area. He is insistent, and rightly so, that an effective controlled vocabulary must be built around the special needs of the user group it is to serve. Consequently, the maker of a controlled vocabulary must learn as much as he can about the characteristics of this user group, especially the types of requests they are likely to make to the system. I support these sentiments fully. Elsewhere I have said that "user warrant" is even more important than "bibliographic warrant" in the construction of an indexing language.

Soergel is a careful writer. In particular, he is careful to define all the terms that he uses. Some may consider him

too careful, that there is too much definition, and that some of this is hair splitting. I do feel that terms should be carefully defined but I also feel that the author goes overboard on occasions. Sometimes I find myself thinking "all this precise definition is fine, but let's get on with the discussion". He also introduces new terms for familiar concepts when he feels that the "old" terms are inappropriate. For example, he prefers the terms "precombination" and "postcombination" to "precoordinate" and "postcoordinate". Again, I find myself mostly in agreement with his terminology, but I am not always certain that the new terms he introduces are an improvement on the ones they replace. Sometimes his choice of terminology is unfortunate I feel (e. g., "quasi-synonym" used for "near synonym").

Soergel's book is mostly well arranged, although on occasions he introduces terms that he has not yet defined. For example, on page 6 he introduces the terms "precombination" and "postcombination" long before these terms have been explained. It is, of course, difficult to maintain an optimum sequence in a work of this type and minor blemishes of this kind can be forgiven. More annoying is the fact that the proofreading of the text leaves a lot to be desired. For example, on page 20 the word "in" appears twice in place of the correct "ion" and on page 22 "lightning" is listed as a synonym of "illumination". While such errors should be obvious to the reader, it is unfortunate that an author who is so careful in his definitions should allow typographical errors of this kind to creep in.

There are some other defects that I would like to point out. One of these is the tendency of the author to make sweeping, authoritative assertions without in any way justifying them. For example, he says categorically that "the higher the degree of mechanization of an ISAR system, the greater the need for a good thesaurus that indicates conceptual relationships". I am not at all sure that this is true. At least, I cannot accept a statement of this kind without some justification being given. But such justification is lacking in the text. Let me quote one more example. The author states that, in determining the appropriate level of exhaustivity of indexing, and specificity of vocabulary, important factors to be considered include amount of time available to do a search and expected frequency of search requests. Why are these important? It is not at all obvious, at least to me.

Another criticism I have relates to the incomplete treatment accorded to certain topics. On page 9, for example, Soergel lists three "criteria for the evaluation of a thesaurus", namely degree of conceptual completeness, degree of terminological completeness, and quality of the display. Although he is well aware of the importance of specificity of the vocabulary, Soergel makes no mention here of this evaluation criterion, which is an extremely important one. Such omissions are dangerous in a handbook that is not necessarily read in toto.

Very occasionally the text is inaccurate. His statement on page 56, for instance, that roles and links cannot be used with peek-a-boo cards is just not true. It is difficult but it can be done.

I have deliberately looked for defects in this book and I have pointed these out when I have found them. They are, however, minor blemishes in what is otherwise an

excellent work. I thoroughly recommend it as a reference tool to students, teachers and practitioners.

F. W. Lancaster

GOPINATH, M. A.: *Classification Research (India): 1968–1973*. Bangalore: Documentation Research and Training Centre 1974. 78 p. = FID/CR Report No. 14; FID Publ. no. 405.

This review of recent thought about classification from the Indian School is interesting and thought-provoking. If space permitted, a lengthy critique would ensue because there are fundamental ideas here which are quite controversial. It is very strongly recommended that readers start with the glossary (p. 60–62), even though it is not complete, because some of the words do not carry meanings used in standard English.

Three universes have been postulated by the late S. R. Ranganathan and A. Neelamegham: the universe of entities, the universe of ideas and the universe of subjects. An idea is generated when a knower (human) “knows” (recognizes) an entity. The systematized account of a body of ideas makes a subject. “The totality of all ideas preserved by the civilization at a particular point in time constitutes the universe of knowledge” (p. 9). The inclusion of the phrase “at a particular point in time” conjures up Zeno’s paradox of the moving arrow. There would have to be an infinity of such universes between points (and no next point) and no two universes should be exactly alike no matter how small the duration of time between points. This definition of the universe of knowledge would make any kind of subject analysis a retrospective process, which, perhaps, is what it is. The question then becomes: “To what extent should one abandon the pretense of currency?”

Where there is an assumption that the entity upon which an idea is based is tangible, presumably all civilizations existing at that specific point in time would have generated the same idea. One may suggest, however, that when an entity is intangible – God, courage, redness – there can still be an idea but not necessarily the same idea for each civilization. Some civilizations accept the existence of entities which others would not countenance (devils, unicorns, pathogenic bacteria). A large part of knowledge, even accepting the Ranganathan-Neelamegham definition, consists of belief. From isostasy to plate tectonics, for instance, is moving from one idea (or paradigm, if you wish) to another for explanation of the same collection of entities. The path of “knowledge” is strewn with the wreckage of ideas that have been superseded, but they are still part of the totality of the universe of knowledge viewed longitudinally rather than in cross-section. (From a given point in time, one may look both ways). The reigning paradigm is accepted because it best “saves the appearances”, but its ephemeral character over time suggests that in the long view it could be called “belief” since it is only “true” for a limited period. Thus “knowledge” defined (p. 60) as “the totality of ideas conserved by human beings” has to include those ideas which are beliefs. “Idea” is not defined, but presumably the Oxford definition of “archetype” or “pattern” suffices. The standard definition of knowledge specifically excludes belief.

Another definition problem occurs with the word “sub-

ject”. This is defined as “an organised or systematised account of an idea or body of ideas whose extension and intension are likely to fall coherently within the intellectual competence and field of inevitable specialization of a normal person.” (p. 60. One would interpose “highly educated” between “normal” and “person”!) Library and information scientists normally deal with a subject *literature* – the writings *about* a subject – rather than with the subject itself, which is organized, systematized and defined by its adherents. Physicists organize their data; historians organize their data, and so on. The library or information scientist organizes the *literature* of physics or of history. Some few people combine careers, contributing both to the literature in the subject itself and to the literature about the subject literature, as with the work of Derek de Solla Price in history of science and in the nature of scientific literature.

In similar vein, a diagram (p. 8) has been drawn to show the relationship between systems of subject analysis, classification, subject indexing, subject heading, etc. The core is given in three parts: “analysis of subject into component ideas”, “assembling the component ideas co-extensively”, and “symbolisation or naming the subject.” The implication is that the person doing the classification, indexing or analysis does all these things, but he does not. Did the subject analyst name cybernetics? or psycholinguistics? or any other subject? Of course not. He took over the organization, system and names from the literature written by specialists in these subjects. Would the specialists pay any attention to *his* names? This failure to distinguish between a subject and its literature has been a very common one. Robert A. Fairthorne, in particular, has taken great pains to point out the difference (cf. *Annual Review of Information Science and Technology*, v. 4, chapter 3, 1969). Classification applied to natural history (taxonomy) and classification applied to information transfer are two different activities, though there are some principles in common which may or may not be utilized.

Another problem is caused by a linguistic oddity. We are informed in an early section that the “universe of entities or *knowees* consists of all knowable entities” (p. 9, *Italics mine*). This sentence would be clear without the “or knowees”. Misunderstanding arises because one would expect “knowee” to mean “one who knows” – following the Norman French part of the English language, as found in *fiancé*, *fiancée*, *divorcé*, *divorcée*, *grantor*, *grantee*, *guarantor*, *guarantee* (the latter not to be confused with its homonym from the Old French, *guarantie* = *guaranty*, *guarantee*). “Entity” (Oxford: “a thing’s existence as opp. to its qualities or relations; thing that has real existence”; Webster: “a thing that has reality and distinctness of being either in fact or for thought”) is a perfectly good word, especially in the Webster definition. The invention “knowee” confuses unnecessarily. The same may be said for the introduction of the term “speciator” (p. 22–24).

Finally, it is interesting to see a report of the progress of classification research which describes a system without classes. One can understand the switch from the formality of “classes” to the aggregativeness of “subjects”, given the present uncertainties caused by growth and splitting-off and/or merging of parts of traditional disciplines. Presumably the whole body of theory in the Indian