obliged to conform to the usages laid down in a standardized vocabulary promulgated by an official agency.

Although, as a linguist, Sager describes the point of view of terminologists and term planning in a detached way, he seems to accept the premise that, ideally speaking, there should be only one (preferred) term for each concept designated in a special language. When describing the rules followed in scientific nomenclature, Sager writes that "names should be univocal and unique but simple and concise." (p. 293) However, he admits that this is possible only when "all users agree on concepts and their terms . . . standardization of designation can only begin when conflicting theories are resolved." "Since knowledge is constantly evolving . . ." however, Sager concludes that this is a rare condition. (p. 330)

Nevertheless, Sager describes with apparent approval the methods used by the BSI in which glossaries *prescribe* for their users a "preferred term" that is presented as an "entry term" for the definition which follows. Also included in an entry may be "alternative" and "deprecated" terms: an example is FEATURE CARD (preferred); ASPECT CARD, and TERM CARD (alternative); and DESCRIPTOR CARD (deprecated) – taken from BS 5408 (1976). More acceptable in the social sciences, by contrast, would be a *descriptive* approach that simply identifies the terms in use (with information about their users) and does not seek, overtly, to influence usage. To sustain this descriptive stance, it is possible in a classified glossary to abandon the use of "entry terms" by listing all the terms in use *after,* rather than *in front of,* their definitions.

Admittedly it is easier to accept the prescriptive norms of terminology (by contrast with the descriptive method) when attention is focused on the fields of technology and natural science, as they are in this book. Nevertheless, specialists in the social and information sciences are interested in the development of their own special languages even though they cannot reach the levels of terminological rigor achieved in the "harder" subject fields.

A major obstacle to the formation of special languages in the "softer" sciences arises from the difficulties encountered by creative scholars when they attempt to validate a claim that they have discovered or created a "new" concept. Although the validation of such claims in the "hard" sciences may not be automatic, it is certainly easier than in the social and information sciences, in part because existing concepts are both more tangible and also better defined and named. The point is that if an author cannot win acceptance of a claim for conceptual innovation are presumtuous and ego-gratifying, even though they cannot themselves cite earlier works in which the supposedly new concept had been defined and named.

The elaborate discussion by Sager of the linguistic forms and processes used to name new concepts begs this prior question which every author must face: is this indeed a new concept and, if so, will my efforts to name it lead to acceptance or baffling frustrations?

The uses of a glossary in this connection deserve careful attention. Sager writes (p. 335) that glossaries "can greatly simplify communication among specialists and ensure unambiguous and therefore more economical and effective communication." Glossaries that follow the British Standard are always classified: ". . . they are ordered by concepts so that related terms are grouped together." Yet in their own glossary of "traditional forms," a "glossary" is described as ". . . a list of terms with explanations and/or definitions." (p. 162) Clearly Sager thinks of a glossary as an alphabetized dictionary restricted to a single subject field or special language. Such glossaries cannot help authors establish the newness of new concepts. Only a classified glossary can do that, provided it is widely accepted as comprehensive among users working in its subject field, and provided the logical place for a concept can be found in the scheme, even though it lacks a "term" to be defined. The fact that BS glossaries are actually classified means that the kind of tool which could potentially be used to provide this fundamental service to writers is already available – yet its use for this purpose is not examined in this book.

These considerations bring us back to the emphasis placed by Sager in his Preface on the ability of special languages to provide *new words* to designate the *new concepts* generated by scientific and technological progress. Such progress is, indeed, a continuing and even accelerating phenomenon – thus the emergence of new concepts that need to be named has become an ever-increasing flood. Until the writer's need for help in making the case for novelty, and thereby legitimating the subsequent process of naming, is recognized, the core problem involved in the efficient generation and stabilization of special languages has escaped attention.

The problems of text production are complementary to those of text interpretation. The practitioners who create special languages are, for the most part, engaged in text *production.* Information scientists and linguists, by contrast, focus on problems of text *interpretation* – even though, as writers about their own subject field they are themselves also engaged in text production. *English Special Languages* gives us an important and useful analysis of how to interpret special languages *after* they have taken shape. It provides, regretably, little help for those who are interested in the complementary processes: how to *create* special language.

Fred W. Riggs
University of Hawaii
Honolulu, Hawaii 96822, U.S.A.

OPITZ, Otto: Numerische Taxonomie. (Numerical Taxonomy). (In German) Stuttgart–New York: Gustav Fischer Verlag 1980. 191 p., DM 16,80 = Grundwissen der Ökonomik: Betriebswirtschaftslehre, UTB Nr. 918. ISBN 3-437-40079-7

This is an introductory textbook on numerical taxonomy in its wide sense embracing different problems and mathematical techniques from multivariate analysis, exploratory data analysis and cluster analysis. The author emphasizes on three main topics: classification of objects (i.e. the construction of homogeneous groups of objects), representation of objects (as points in some multivariate space), identification of objects (extraction of representative features explaining a given classification or representation). In each case the starting point is a set of objects whose properties are described by a set

of variables or by some indices of pairwise similarity resp. dissimilarity. These notions are clarified in Section 2 which not only discusses nominal, ordinal and quantitative variables, but also treats on variables whose alternatives are ordered hierarchically or bear some lattice structure. It is shown how several measures of distance may be aggregated to a "global" distance. Section 3 presents some clustering methods for forming partitions, coverings, hierarchies or quasi-hierarchies of classes using several criteria for measuring the homogeneity of classes or evaluating the goodness-of-fit of a classification. In Section 4 a Euclidean representation of objects is found by the usual methods of principal component analysis or by nonmetric multidimensional scaling. – For the identification of objects (Section 5) an optimal weighting of (quantitative) variables is found by discriminant analysis or by regression and canonical correlation analysis. Identification with qualitative or mixed data is handled by calculating some distance index for each variable and linearly aggregating these indexes to a global index d such that the partition to be explained is a minimum-distance partition generated by d (system of linear inequalities) resp. such that d is a monotone function of δ (= distance induced by the given representation; Kruskal – like gradient algorithm). These methods are new. – Section 6 informs on existing computer programs. – The text is written for students of economics, its style is informal and illustrative. Because all formulas and algorithms are given in their exact mathematical form the reader should have some prior mathematical or statistical knowledge (the Section 1.3 on "mathematical foundations" seems to be insufficient). However the methods are only heuristically motivated, no proofs or probabilistic arguments are given. Most algorithms are illustrated by a numerical example (5 objects).

H.H. Bock
Inst. f. Statistik u. Wirtschaftsmathematik der
RWTH Aachen, Wüllnerstr. 3, 5100 Aachen

BRITISH STANDARDS INSTITUTION: **BSI ROOT Thesaurus.** Part 1: Subject display; Part 2: Alphabetical list. Hemel Hempstead, Herts.1981. 620+676 p., £ 155.-

According to the Foreword, the ROOT Thesaurus can be regarded as the product of both, the in-house thesaurus of the British Standards Institution and the institution's contribution to the ISONET thesaurus started in 1974 by a working group of the International Standardisation Organisation. A French version of this work is still held on computer file only, the English one was printed recently as the ROOT Thesaurus in two heavy volumes off the magnetic tapes.

Once the British became famous for the finest car on earth, called Rolls Royce. For my opinion the ROOT Thesaurus is the Rolls Royce in thesaurus making, and again a product of the good advice of Mrs. Jean Aitchison, our model-setter in this field!

Although nowhere in the introduction an explanation is given why the name ROOT was chosen, (an acronym?) the design on the cover page explains it by showing big stem roots of named technical areas emerging out of a center and extending into smaller and smaller roots of less comprehensive technical fields and their subfields. Usually a classification hierarchy is depicted by a tree; the root idea comprises in addition the possibility that sections of this thesaurus can be used in a given field as a starting point and compatible instrument for the development of separate thesauri at other centers.

What are the excelling features of the new thesaurus breed?

ROOT consists of a subject display embracing 24 main classes with a one-letter notation covering mostly technical areas in which standards have been developed. These are subdivided by altogether 139 subclasses with a two and sometimes a three letter notation (capital letters). ROOT contains altogether some 11 800 descriptors and 5500 lead-in terms. The arrangement in the subject display is in faceted order; characteristics of division are added in brackets. In a few cases, such characteristics have become class descriptors themselves, but in general this kind of structuring was avoided. Recognition of hierarchy is facilitated by the typography with a boldface type in different sizes for the first three levels. Wherever necessary, additional information is added to a descriptor such as synonyms, related terms in alternative hierarchy, and broader and narrower terms in alternative hierarchy. A specialty is also the indication of synthesized terms by a certain symbol which should warn not to use the descriptor following but the combination of terms as indicated.

The socalled alphabetical list contains in bold-face print all descriptors, their notations and the descriptors of the next hierarchical levels as well as the non-descriptors. At their respective entries an arrow points to the descriptor to use.

Yes, one uses a new the symbolization, namely the internationally known mathematical symbols which are easy to learn and easy to write but not easy to type.

Indeed with all of this we are having a new thesaurus model and a fine one too. Is it a perfect model? The user will soon answer this question. And what will the theoricians have to say? They might observe that the conceptual structure of the fields as indicated by the characteristics of division does not always comply with the hierarchy, e.g. if a descriptor denotes a process it is sometimes treated as if it were a field with its subdivisions including objects, materials, systems etc. Also regrettably the elaboration of a recurring array of facets was not aimed at. The notation depicts the hierarchy, however, in cases of concept combination (syntheses) the notation precoordinates the otherwise differing facets. There is no rule for expression of syntax in cases of compound terms or term combinations. Thus this product is meant rather to serve as a tool for coordinate indexing than as one to express complex subjects in a predictive and reconstructable way. However, since no other symbols are used with the ROOT notation than capital letters and a period after three such letters, there may still be a chance, at some later date, to develop a syntax and its symbols for an improved condensation, organization and retrieval of information.

Ingetraut Dahlberg