hypothesis is not valid as a common measure in evolutionary biology.

The volume closes with applications in social sciences and medicine. *I. Balderjahn* (in "Cross validation of covariance structures in one and multi group analysis: The case of ecologically concerned consumers") discusses the cross validation of covariance structures in multigroup analysis with data from ecology-oriented consumer behavior. In cognitive psychology, probands are invited to subdivide objects into classes with respect to their similarity. *T. Eckes* (in "The sorting procedure for obtaining proximity data in multivariate psychological research") stresses the necessity to analyse such experimental results with suitable multivariate methods. Finally, *H.P. Schmidt* and *C. Oberwittler* (in "Numerical taxonomy of brain tumors: A challenge to contemporary mathematical classification") discuss problems and preliminary numerical approaches to the malignancy classification of astrocytomas and mixed gliomas.

The very different contributions are clearly subdivided with respect to their contents; they give a good overview on the present state of the art so that the reader will find many stimulating ideas. The book has a detailed index.

Peter Jaenecke

Dr. P. Jaenecke
Forschungszentrum FB, Standard Elektrik Lorenz AG
Ostendstr. 3, D-7530 Pforzheim

WALNE, Peter (Ed.): **Dictionary of Archival Terminology; Dictionnaire de terminologie archivistique.** Compiled by Frank B.Evans, Francois-J. Himly, P.Walne. München, FRG: K.S.Saur 1984. 226p. ISBN 3-598-20275-X

This convenient guide to archival terminology in seven languages will interest not only archivists - its primary audience - but also information specialists, terminologists and lexicographers. The work supersedes an earlier LEXICON OF ARCHIVAL TERMINOLOGY (Elsevier, 1964). Both projects were sponsored by the International Council on Archives (ICA), which started work on the first version in 1954. The ICA Working Party responsible for this revised text labored, under UNESCO sponsorship, from 1977 until 1984. It was chaired by Peter Walne, U.K., and included Charles Kecskeméti, executive secretary of ICA. Other members were François-J.Himly and Michel Duchein, France; Eckhart G.Franz, FRG; Antonio Arago, Spain; Filip J.Dolgih, USSR; and Frank B.Evans, USA and UNESCO representative.

Each entry contains two definitions, in English and French, arranged alphabetically according to the English terms. Equivalent terms in Dutch, German, Italian, Russian and Spanish are listed subsequently, without definitions. 503 numbered terms are defined and unnumbered cross-references from synonyms to the entry terms are interspersed in alphabetical order. The cross-references and first entries are all in English, alphabetically, so no index in English is needed. However, indexes coded to entry numbers are given for each of the six other languages. Consequently users can go directly to the defining entry from any synonym for a concept, as listed in these indexes.

Normally French equivalents are available for each concept, but we are not informed about French terms that lack English equivalents. When a term has two or more senses, the equivalent may not have identical senses. Translations can then be problematical. For example, CHARGE-OUT may mean the act of recording the removal of a document from storage in an archive, or the document used to record this action. In French, there is no term for the first of these senses. Accordingly the concept is defined by a direct translation of the English definition and dashes, in brackets - "(---)" - substitute for the missing term. When French terms do exist, however - as they usually do - they are defined in French, and users must judge by comparing the definitions whether the concepts are indeed identical.

There are two French terms for the second sense of "charge-out": *fiche de déplacement* and *fantôme*. However, "fantôme" has another sense, equivalent to the English, *dummy*, meaning a card or sheet used to replace the borrowed item in its normal storage place. Comparison with terms in the other languages may reveal additional terms for the concepts that lack unequivocal terms in English or French. For example, the first sense of "charge-out" may be called, in German, *Aushebung* or *Ausleihe*; the second sense, *Leihzettel*, or *Bestellzettel*; and the idea of a "dummy" - the second sense of "fantome" - can be represented by *Stellvertreter* or *Retent*.

This example shows how terminology varies between languages and it also illustrates a problem common to all technical glossaries that rely on the alphabetical arrangement of entries. If the concepts used in archival work had been classified according to their definitions, all the terms used for them in each of the seven languages could have been listed after each concept. (See (a-d) below). Those that are equivocal, i.e. used to represent more than one archival concept, could have been marked as such, and users would more easily find the unambiguous terms for each concept. Whenever a useful concept lacks an unequivocal term, the editor or users might feel encouraged to suggest new ones that would not be ambiguous. A classified glossary, moreover, would bring related concepts together for easy comparison, thereby enabling users to grasp their logical relationships very quickly.

A common complaint against a classified format is that it requires two steps in searching, to go from the index to the numbered record. However, only English users of this dictionary can go directly to the entry terms without using the index. The users of all six other languages have to consult an index first in order to find a concept's entry.

Moreover, because of the arbitrariness of any alphabetical arrangement, users easily miss logically related concepts. In the example just given, users will be reminded under the CHARGE-OUT entry to "see also PRODUCTION TICKET", but they will not be directed to the entry for DUMMY. We discover that a "production ticket" is a document signed by users when requesting a loan (or charge-out). Four related concepts that are involved here are displayed in the following:

(a) a document signed by the user requesting an item: PRODUCTION TICKET, BULLETIN DE DEMANDE, *BESTELLZETTEL*, -SCHEIN



https://doi.org/10.5771/0943-7444-1987-2-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

(b) a document recording a loan:
*CHARGE-OUT*, *FANTOME*, *BESTELLZETTEL*, LEIHZETTEL

(c) surrogate for a borrowed item used to show that it was removed:
DUMMY, *FANTOME*, STELLVERTRETER, RETENT

(d) the action of recording the removal of an item from storage:
*CHARGE-OUT*, AUSHEBUNG, AUSLEIHE, and no French equivalent.

To discover such conceptual clusters, and to detect the weaknesses of established terminology - note the equivocal terms in Italics - in an alphabetical array requires more work than would be needed if a classified conceptual glossary were used. (In passing, we might note that although the book is very carefully edited, it does contain some mistakes. For example, the index entry for *Bestellschein/-zettel 74, 277* has several errors. *Bestellschein* is found only in entry no.73 (not 74 or 277) and *Bestellzettel*, will be found at entries 73 and 377 not 74, or 277.

In lexicographic practice the purpose of an entry is to describe a word and explain its uses. This goal correctly determines the standard dictionary format: each entry word is followed by a set of sense definitions. The same format is imitated here: for example, the entry for ARCHIVES identifies three senses of the word: (1) a set of records; (2) an institution responsible for managing such records; and (3) a building in which archives (first sense) are handled.

This format is not well suited to meet the terminological goals of the ICA. These goals are indicated in the "Introduction ' which states that the Working Party has "drawn up definitions..." that "include the essential elements in varying national legally enacted definitions" (p.7). The goal was not to find out what a set of words mean. Rather, the terminological goal was to identify clearly each of the concepts needed by archivists (as given in a definition) and to list the available terms for that concept, in seven languages. The ICA goal was, clearly enough, to establish standards and to improve communication.

The alphabetical arrangement of word entries fails to perform this function (only a classified arrangement of concept records can do the job well). Moreover, the lexicographic entry format suggests that the goal was to define words rather than to name concepts. When a word has several meanings - as in the example of ARCHIVES, no doubt the core concept of this project - users are given no concrete help in finding unambiguous terms. If every concept (e.g. the 3 senses of *archives* had a separate term entry, at least then users would become more conscious of the need for additional unequivocal terms. Note that the single word, *archives*, is used as three terms in archival work. The editors mark them as *archives (1)*, *archives (2)*, and *archvies (3)*.

This expedient does facilitate the writing of unequivocal interdependent definitions, i.e. definitions in which technical terms are used ("entailed") and so marked that readers can easily find their definitions. For example, the definition of TICKET reads: "a *document* issued by *archives (2)* granting permission to a user to consult *records (1)/archives (1)* during a specified period", p.137. Unfortunately, this technique compels users to remember the different senses of "archives" and "records" by their numbers.

An alternative approach which the ICA Working Party might have followed involves finding unequivocal terms for each of the concepts needed in archival work. They might be borrowed as loan words from another language. In the case of "archives", for example, there appear to be terms in Dutch for each sense: i.e. *Archiefbescheiden* for the first; *Archiefdienst* for the second; and *Archiefgebouw* for the third. English phrases could easily be composed for the same distinctions, e.g. "archival collection", "archival agency", and "archival depository". (The last of these terms is actually listed as American usage). Unfortunately the use of sense numbers in place of unequivocal terms may not only confuse users -it apparently confused the editors: for example, in the entry for ARCHIVE we read, under sense (2), "An individual *item* forming a part of *archives (2)*...", Clearly "archives (1)" was intended.

When writing interdependent definitions it is, of course, important to mark all of the entailed terms, something that is normally done in this work -but not always. For example, the definition for REGISTRY PRINCIPLE reads: "The principle that *archives (1)* of a single *provenance* should retain the arrangement established by the creating agency...". *Arrangement* is not marked even though this term has a defining entry which says that it is a process of organizing records according to one of two approved archival "principles". Records ought not, we are told, be arranged according to the PRINCIPLE OF PERTINENCE, which prescribes the classification of materials according to their subject content.

The REGISTRY PRINCIPLE, however, is endorsed, and the PRINCIPLE OF RESPECT FOR ORIGINAL ORDER is listed as an equivalent term. However, no entry for this phrase is given, but virtually the same idea is explained under the term, PRINCIPLE OF RESPECT FOR ARCHIVAL STRUCTURE. It is not clear whether these terms name subtly different concepts or whether, by mistake, we just have two definitions for the same concept. This is the kind of confusion that easily arises with alphabetized entries, but is less likely to occur in a classified system of concepts.

The entry for CLASSIFICATION, which has a "see also" reference to ARRANGEMENT, defines the term to mean "The preparation of a *filing plan/ system* or *classification scheme* for *records (1) / archives (1)* and the placing of *series* and/or *items* within such a plan/ system or scheme". (The entry also identifies "security classification" as a second meaning of "classification".) By this definition both the design of a classification scheme and the classing of items in such a scheme are treated as a single concept.

Interestingly no French equivalent for "classification" is offered, but the explanatory phrase *"planification des classements"* is listed, with a see also reference to CLASSEMENT (1). This sense of "classement" is equivalent to ARRANGEMENT - see definition above. The second sense of "classement" is what would be called FILING in English. It is defined as "The placing



https://doi.org/10.5771/0943-7444-1987-2-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

of documents in a predetermined location according to a *filing plan/system*". The definition for FILING PLAN/ SYSTEM reads, "A predetermined classification plan for the physical *arrangement*, storage and retrieval of *files (1)*...". Strangely, there is no definition for CLASSIFICATION PLAN, but the term has a see reference to FILING PLAN/SYSTEM. This produces a circular definition since FILING PLAN/SYSTEM, by inference, is defined as "A predetermined filing plan/ system for...".

There is, however, an entry for CLASSIFICATION SCHEME, which is defined as "A pattern of arrangement of *archives (1)* by *groups, series* and *items* (not US)". This definition calls attention to discrepancies between British and American usage - the US term for this concept is not specified. Interestingly, also, the entailed terms are somewhat confusing. GROUPS has no entry - but we do find definitions for ARCHIVE GROUP, RECORD GROUP, and five other kinds of "groups". ARCHIVE GROUP is explained as "The primary division in the *arrangement* of *archives (1)* at the level of the independent originating unit or agency".

ITEM names "The basic unit of *arrangement* and *escription...*" and SERIES, "*Item* or *documents* arranged in accordance with a *filing plan/system* or maintained as a unit...". SERIES has a see also reference to CLASS, which is defined as "An identifiable and self-contained subdivision of an *archive group* consisting of a number of *items* with one or more common characteristics (UK). Generally equivalent to *series*". Thus "series" and "class" appear to be used as virtual synonyms but they are defined by reference to the origin ("provenance") of documents, rather than their subject matter. If classificationists find this confusing, it is because archivists use some of their vocabulary, but with significant shifts in meaning.

Enough has now been written to give readers a feeling for the book. On the whole it is very readable and clearly laid out, a good reference tool for archivists. It has minor mistakes and discrepancies scattered about and it overlaps classification theory, using some of its terms for different notions. Above all, its quasi-lexicographic format is a real handicap: a well conceived terminological design would make its contents more intelligible and helpful to its users.

Fred W.Riggs

Prof.Dr.F.W.Riggs, Department of Political Science.
University of Hawaii, Honolulu, HI, 96822, USA.

BUTLER, Christopher: **Computers in Linguistics.** Oxford, GB: Blackwell 1985. IX,266p.,
ISBN 0-631-14266-5

"Computers in Linguistics" ist eine Einführung in die philologische Datenverarbeitung unter ganz besonderer Berücksichtigung von SNOBOL 4.

Der Autor erklärt, wie Computer im Prinzip arbeiten. Ein sehr nützliches Kapitel seines Buches gibt Auskunft über den Funktionsumfang von Softwarepaketen für die philologische Textverarbeitung: EYEBALL, OXEYE, OCP, CLOC. Auch auf Statistikpakete, wie SPSS wird hingewiesen.

Butlers Anliegen ist es, Philologen und Linguisten an die computerunterstützte Sprachforschung heranzuführen. Für Aufgaben, die mit Softwarepaketen nicht zu bewältigen sind, schlägt er als Programmiersprache SNOBOL vor, eine seit 1962 entwickelte Sprache zur Text- und Symbolverarbeitung. Sie unterscheidet sich von anderen höhercrn Programmiersprachen, wie ALGOL, COBOL und FORTRAN vor allem dadurch, daß sie das Pattern-Matching, die Definition von symbolischen Mustern und den Test auf deren Vorkommen in den Daten, besonders unterstützt. Auch die benutzerdefinierten Datentypen von SNOBOL sind interessant.

Den Schwerpunkt des Buches bildet eine gründliche und gut verständliche Einführung in SNOBOL. Viele Beispiele, Orignialprogramme und Aufgaben mit Musterlösung machen das Buch als HIlfe auch zum Selbststudium geeignet. Erklärt werden Grundkonzepte von SNOBOL, das Pattern-Matching, die Kontrollstruktur von SNOBOL-Programmen, eingebaute Funktionen, Vergleichsoperatoren, benutzerdefinierte Funktionen und Datentypen, die Fehlersuche in SNOBOL-Programmen.

Zwei Beispiele für SNOBOL-Anwendungen schließen den Band ab: eine sprachstatistisch-stilistische Analyse der Dichtung von Sylvia Plath und eine Berechnung der lexikalischen Dichte (type-token ratio) von Texten aus dem London-Lund-Korpus des gesprochenen Englisch.

Die Beispiele kennzeichnen den Schwerpunkt des wissenschaftlichen Interesses von Ch.Butler und den Lesern, die er anspricht: Im Blickpunkt stehen die computerunterstützte quantitative Stilistik, Autorenschaftsuntersuchungen, Wörterbucherstellung, Textedition und computerunterstützter Sprachunterricht. Computer dienen vor allem zur Produktion von Indices und Konkordanzen, zur Auszählung von Frequenzen, zur Errechnung statistischer Kennwerte, zur Lemmatisierung. Innerhalb dieses Interessenbereiches ist die Darstellung von Ch.Butler technisch und konzeptionell modern. Der Ausblick auf andere computerlinguistische Forschungsrichtungen (automatische Übersetzung, Textproduktion, Simulation menschlichen Sprachverhaltens) bleibt begrenzt. Eine Computerlinguistik, die sich um die Gestaltung von Mensch-Maschine-Dialogen, um die Repräsentation von Texten, um die Verarbeitung nicht transkribierter, sondern akustisch aufgenommener Sprache und immer wieder um das zentrale Problem des syntaktischen und semantischen Parsings bemüht, kommt bei Ch.Butler so gut wie nicht vor.

Brigitte Endres-Niggemeyer

Prof.Dr.B.Endres-Niggemeyer, Fachhochschule Hannover,
FB BID, Hanomagstr. 8, D-3000 Hannover 91.



https://doi.org/10.5771/0943-7444-1987-2-107 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY