Indexing


This collection of essays is an essential addition to the library of taxonomists and indexers interested in enlarging their work to include thesauri. Bringing together many of the major theorists and practitioners, the editors have provided a wide view of thesauri and their importance in contemporary search activities. This is truly a resource book: the essays cover many important aspects of thesaurus construction, and list a truly amazing amount of further resources on every aspect. It's a great starting point from which the interested reader can carry on.

Jean Aitchison and Stella Dextre Clarke summarize the history of thesauri, allowing the reader to realize just how much of a foundation was laid 50 years ago. Ms Aitchison's 'Thesaurus construction and use: a practical manual' has been a standard on bookshelves for years, and many have valued her practical advice many times in their own work. She and Ms Clarke cover the major developments and ideas since the first unpublished information retrieval thesaurus and bring the concept quickly up-to-date by discussing ontologies and the needs for new standards.

Alan R. Thomas develops a self-teaching guide for new thesaurus designer wannabes. His 'Teach yourself thesaurus' piece not only breaks down a process for learning the art and craft, but gives a list of resources for reading and analysis that covers each learning phase. James R. Shearer follows this with 'A practical exercise in building a thesaurus'. The use of a real example is invaluable for learning construction and raising the issues that come with categorization in any human endeavor. I have to admit, my arrangement of Shearer's example did not resemble his, and I spent time thinking through the reasons for that result. People take differing approaches to categorization, and the environment you work in can frame the categories you choose. It's a great exercise to work this example without peeking and compare your results.

As if guessing that I would be examining the issues that resulted from Shearer's example and my own interpretation, Marianne Lukke Nielsen's 'Thesaurus construction: key issues and selected readings' follows Shearer. Nielsen discusses issues such as the differences between classical thesauri and searching thesauri, the approaches used by corporate taxonomies, problems in work process, user-centered approaches, term collection issues, organizational concepts, and automatic thesaurus construction. An extensive list of references and readings is summarized.

Leonard Will's 'Thesaurus consultancy' should be mandatory reading before working on a first contract as a taxonomist. His discussion of the role, benefits and expectations of working with a consultant or as a consultant are clear, and will help the potential contractor to redefine the contract and understand potential issues.

Leslie Ann Owens and Pauline Atherton Cochrane then discuss how to analyze the effectiveness of a thesaurus in use. Studies to evaluate the value of thesauri have a long history of their own, and these efforts are summarized for the reader. Formative, observational, constructive, and comparative methods of evaluation are defined, and sample forms and questions from the study are provided.

Continuing the theme of real-world value, Jane Greenberg describes the role of user education in 'User comprehension and searching with information retrieval thesauri'. Her studies indicate a rise in valid retrieval rates after students have a short introduction to thesauri components and their use in searching in highly controlled databases. Thesauri are not the simplest method of search available, but her study finds students are more open to using this method once they understand the concepts and benefits. The thesaurus interface is an important component: it must allow interaction in the process.

But the latest, hottest use of thesauri is occurring on the Wild West of the web. Eric H. Johnson discusses the drawbacks of web searching in browsers and the need for a web services thesaurus application. He describes how XML thesaurus protocols could facilitate the user's search experience, the kind of functionality such an application would need, and the possibilities of cross-thesauri searching through a UDDI registry. He also describes additional benefits of a web services application, such as expanding queries and populating concept spaces with suggested search terms.

Melissa Riesland follows with an overview of the software packages available for constructing thesauri. She begins by defining classification, notation and controlled vocabularies, and walks through the importance of relationships, trees, networks and structures. She gives the reader a good understanding of the continuum of vocabulary structures, ending at XML topic maps. Her software survey and assessment of features and user needs is an invaluable checklist.

Patrice Landry summarizes the efforts to establish a true 'Global Village Library' and details the history of the MACS Project. MACS is an effort to provide multilingual subject heading access to databases by linking or mapping subject heading languages across domains in four languages. Landry outlines a brief history of multilingual projects, lays out the standards involved, and discusses the use of 'linguistic proximity' as the goal for users' searches. MACS Link Management Interface has been tested on several smaller domains, and major work continues to link further knowledge domains. Other projects involving multilingual subject heading languages are outlined as well.

The final piece is an interview with Dr Amy Warner on the status of the NISO standard Z39.19 Guidelines for the Construction, Format, and Management of Monolingual Thesauri. The interview was conducted in June 2003, and describes the goals that the revision team had for the next version of the standard. That revision is available in PDF form at www.niso.org/standards/resources/Z3919.pdf?CFID=5528173&CFToken=27515538

The only disappointments in the book are that it can't cover every issue in thesaurus design. The breadth of the essay coverage is very good, and the volume brings up all the major issues in the field. Some of us would like a bit more on specific topics, such as cross-linking multiple thesauri, but the references direct us to more information. And the index is a bit weak structurally but perfectly adequate; the reviewer is an indexer, and we must always take ourselves with a grain of salt. Overall, the collection is a valuable introduction to thesaurus construction and appreciation, and does a good job of covering many of the issues, standards, benefits and processes of thesaurus and controlled vocabulary use. The sheer
amount of bibliographic resources in the book ensures that whatever gaps the authors may have left will be filled by following the paths laid out in the references.

Jan C. Wright, indexer, taxonomist, and controller of wild vocabularies

Information searching and retrieval


It is over eight years since I last worked full-time as a librarian. Although I do still return to the fray occasionally, temping or doing contract work, my primary occupation now is as a freelance indexer – mostly of books or other print materials, but I have also ventured into website indexing, dabbled in indexing of other electronic texts, and generally tried to stay on the track of emerging ideas and new developments in ways of getting information from electronic resources.

Anyone who has followed a similar trail will know that it is fairly easy to pick up a basic understanding of metadata as a tool for resource description and subject indexing. Going beyond that, though, in breadth or in depth, can be time-consuming. In my case, there were certainly plenty of gaps in the understanding, and some near-bottomless chasms of ignorance about metadata’s other uses, in and outwith the LIS community.

David Haynes’ book offers plenty of material to help fill those gaps. Its aims are summarized as ‘describing recent progress in metadata standards and applications and focusing on the concepts behind metadata’. The book’s target audience is ‘information professionals who want to develop their knowledge and skills in order to manage metadata effectively, and managers who are faced with strategic decisions about adoption of IT applications that use metadata’.

Haynes first looks at the historical background of the term, and concepts associated with it. He then puts forward his own approach: in place of the usual ‘data about data’, he defines metadata as ‘data that describes the content, format or attributes of a data record or information resource’, applicable to structured or unstructured information, in print or electronic form, and stored either in the resource or in a separate database. Subsequently, he outlines a five-point model of metadata’s purposes: resource description, information retrieval, management of information resources, documenting ownership and authenticity of digital resources, and interoperability.

Two chapters then deal with metadata in general, the first looking at ways of defining, expressing and storing it. Mark-up languages such as XML, with its tags, schemas and Document Type Definitions, embed metadata in the document it refers to; databases of metadata store it separately. Haynes then reviews various contexts in which metadata is used: word processing, cataloguing, records management, e-commerce and content management.

Next, he considers some of the data modelling systems underlying metadata standards. Almost all of those covered were new to me, apart from the Resource Description Framework (to which Haynes supplies the most comprehensible introduction I’ve yet come across). The others discussed are the ABC Ontology, Functional Requirements for Bibliographic Records, the Indecs metadata framework, and the Open Archival Information Standard. This chapter ends with a review of metadata standards themselves: Dublin Core and its extensions/derivatives, MARC, ISAD (International Standard Archival Description), ONIX, and standards for multimedia and educational resources.

Haynes then revisits his model of purposes. Each of the five has a chapter devoted to it, providing both an outline of current developments and an introduction to relevant concepts in each of the specific areas. These are followed by a chapter on managing metadata. The five-point model returns again in ‘Looking forward – the future’, this time as a framework for consideration of trends and possible developments in each of the five areas. Finally, he looks at trends in metadata management, speculates on the durability of metadata, and makes some predictions about what the future may hold for it and those who work with it.

The book is chock-a-block with information, on virtually every aspect of metadata. The one omission I can identify is faceted metadata classification: there is no mention of it or its associated language, XFML. Slightly surprising was the absence of any reference to topic maps in the discussion of metadata’s durability and future development. Topic maps share at least two of metadata’s purposes – resource description and information retrieval – but it is not yet clear whether they will compete with metadata in those arenas, or complement it in some way. A discussion of the possibilities could have made interesting reading.

The text is well organized and well presented, each chapter beginning with an overview of its content and ending with a summary, plus a list of references and further reading. A glossary would have been a useful addition, particularly as the index does not immediately identify where definitions of terms can be found.

The index has other weaknesses, too. Its coverage of the text is patchy: ‘artificial intelligence’ refers only to page 98, for example, ignoring the interesting comment on page 176 about the role of AI in future systems; ‘controlled vocabularies’ refers to a major discussion on pages 152–4, but not to the point made on page 138 about their importance to interoperability.

The book is chock-a-block with information, on virtually every aspect of metadata. The one omission I can identify is faceted metadata classification: there is no mention of it or its associated language, XFML. Slightly surprising was the absence of any reference to topic maps in the discussion of metadata’s durability and future development. Topic maps share at least two of metadata’s purposes – resource description and information retrieval – but it is not yet clear whether they will compete with metadata in those arenas, or complement it in some way. A discussion of the possibilities could have made interesting reading.

If all that sounds like nit-picking, it is not; it is groundwork for a point I think important to make. LIS professionals know that a poor index limits the usability of a book for reference purposes. We also claim, as one of our core skills, expertise in organizing information for retrieval. Yet here we have a professional publication, about a new technique in information management and retrieval, in which the traditional, built-in tool of retrieval has not been made to function as it should.

That the defect blights an otherwise excellent book is disappointing. What bothers me even more, though, is this: if we do not, in our own professional literature, demonstrate the ability to make old methods work, what kind of message does that send about our abilities to cope with the new?

Linda Sutherland, freelance indexer and librarian

This review is reprinted, with permission, from eLucidate, 1(5), Sept/Oct 2004 (www.ukelib.org.uk).

Information retrieval design: principles and options for information description, organization, display, and access in information retrieval databases, digital libraries, catalogs, and indexes. James D. Anderson and José Pérez-Carballo. St.
Don’t let the boring title or the considerable heft of this volume warn you off, or the fact that it is billed as a textbook and database manual. This book provides a fascinating discussion of the principles of information access for both novice and more experienced audiences. Perhaps the last word in the title should come first, for indexing is at the heart of this book: its original title was Indexing for Information Retrieval: The Design of Indexes for Textual Databases.

The book is organized around 20 key issues in information (or ‘message’) retrieval, and is based on James Anderson’s (the lead author) 25 years of teaching, designing and evaluating databases, and doing committee work for the US National Information Standards Organization (NISO). It is gracefully written, insightful, and not without touches of humour. The book provides a comprehensive and thorough analysis of what Anderson defines as its crus, the ‘retrieval of messages’. By choosing to focus on what he calls messages, Anderson avoids linguistic quibbles over terms such as data, information and knowledge. Messages encompass all of these. A message, we are told, is potential information for a user, ‘encoded in a text, which is recorded on a medium, which together with the text constitutes a document’.

Anderson and his co-author, José Pérez-Carballo (who focused on the automatic indexing and interface design sections) seek to describe information retrieval databases, what they must/should include, how they work, and the options available in their design. They methodically and lucidly lead us through all the in and outs of indexing, display and interface. Following are some of the most important topics discussed, each of which commands a chapter: display media, documentary units, indexable matter, analysis and indexing methods, exhaustivity, specificity, displayed vs. non-displayed indexes, syntax, vocabulary management, surrogation, locators, arrangement of displayed indexes, and size of displayed indexes.

For each topic, Anderson discusses historical background, as well as pertinent research studies and theory, and provides copious examples drawn from real books and databases, along with an abundance of illustrations of model indexes and web pages. He gives a good sense of the debates that go on in the world of information science research, including controversies concerning alphanumeric sorting, human vs. automatic indexing, Boolean vs. weighted retrieval, and vocabulary control. He deplores the lack of standards for alphanumeric arrangement, and of rules for indexing, while acknowledging the absence of a solid research base in these areas.

Chapter 8, one of the longest in the book, goes deeply into indexing theory. Anderson draws on, inter alia, Wittgenstein’s writings, queer theory, gender studies and more in his analysis, relating indexing to larger philosophical, social and cultural issues (as he also does with cataloging and classification). This chapter compares human and computer indexing, finding the results about equal (although again, research findings are inconclusive). Anderson believes that while automatic indexing works, and human indexing is not always economically feasible, human indexing is better than automatic at capturing qualitative and nonexplicit messages.

Anderson closes each chapter by applying the principles he has just laid out to ‘three prominent types of IR [information retrieval] databases’: a book (actually this book, which counts as a database by virtue of its two indexes—the table of contents and the one at the back); a scholarly indexing and abstracting database; and a full-text electronic encyclopedia.

The book is a marvel of organization abetted by clever typographical layout. Each topic and subtopic is numbered; each paragraph is numbered; and printed directly above the paragraphs is a brief annotation (which is identical to the index entry/entries in the back of the book). The paragraph, not the page, is the documentary unit for indexing—locators in the index refer to chapter and paragraph number, such as 12.149 (except for a few that refer to entire sections or chapters). Anderson strongly recommends that paragraph rather than page referencing be more widely adopted. The index to this book, prepared by a human being using NEPHIS (Nested Phrase Indexing System), an ad hoc string system developed by Timothy Craven in 1986, is excellent. The book also includes a full bibliography and a good glossary.

Unlike most dry-as-dust manuals and textbooks, this book often has the flavour of a personal statement by Professor Anderson, who draws on his deep knowledge and obvious love of the information retrieval field, writes freely yet cogently about these issues, and often chooses illustrative examples that reflect his long-time support of gay and lesbian rights. The book comes off rather like a good (long) conversation with a knowledgeable and informative colleague. I recommend it highly.

Clare Imholtz, librarian and freelance indexer

Reference works


When the original Dictionary of National Biography (DNB) first appeared it marked an intellectual achievement on a par with the publication of Diderot and d’Alembert’s great Encyclopédie, and together with its supplements has proved an indispensable work of reference. Now, more than a hundred years later, we celebrate a new version which is at the same time a revision, a cumulation and an update. It has already received a variety of notices, generally admiring, though inevitably there is controversy about its merits, including a continuing correspondence in the Times Literary Supplement. A short review can do little more than nibble at the edges of such a monument of scholarship. I adopt an indexer’s view, which should be narrowly focused on particular aspects: form, structure and arrangement; reference value as an authority; and the work’s recognition of individual indexers.

Before all else, though, this is a handsome and pleasing piece of book production, clearly printed in two columns, in volumes that sit well if not in the hand at least on the desk. For the first time it includes portrait illustrations, not for all the subjects, but profuse enough to add a significant pictorial dimension to the work.

The original DNB with its supplements should need little explanation to indexers of all kinds, and this new version follows essentially the same plan: an alphabetical series of narrative accounts of notable British individuals, no longer living (the cut-off date now being 31 December 2000), but covering all periods of history.

Great reference works often open with an Introduction, too often overlooked, but the preliminary matter to the Oxford DNB (as we must now call it) repays careful reading for its explanations on methodology, scope and authority. Clearly acknowledging the basic soundness of the original, it outlines the innovations and changes. Apart from the illustrations, these include naming the contributor at the end of each article, and (to me rather a curious
feature) ‘wealth at death’. There is also now a small number of articles about families and groups, an innovation I find potentially useful, including as it does entries for ‘the Grey family’ and ‘the Tolpuddle Martyrs’. The Introduction states that in scope the DNB is ‘not merely a roll-call of the great and the good but also a gallimaufry of the eccentric and the bad’. So we find entries for ‘Cooper, Thomas Frederick [Tommy] (1921–1984), comedian’ and for ‘Christie, John Reginald Halliday (1899–1953), murderer’. Nationality is a defining condition for entry, but the Introduction confesses to problems of definition and claims a flexible policy. Thus, quite rightly, ‘Conrad, Joseph [formerly Jozef Teodor Konrad Korzeniowski] (1857–1924), master mariner and author’, but also, perhaps more surprisingly, ‘Loizikey (c.1845–1919), Ndebele queen of Bulawayo’, ‘Abbas Halmi II (1874–1944), last khedive of Egypt’, and ‘Omai (c.1753–c.1780), first Tahitian to visit England’. The old imperial umbrella continues to offer shade. However, the contents remain predominantly a gallery of significant Britons. Following the Introduction, a Guide to Articles describes the arrangement: alphabetical of course, though this receives due explanation and illustration. Then, entries are ‘generally . . . under the final word of the surname’, so ‘Bannerman, Sir Henry Campbell’. This is rather at variance with the usual convention for conjoint names and even with the style of many entries in the original and its supplements. However, there can be no hard-and-fast rule about this, except the sovereign one of consistency within the work itself. Similarly, like the original DNB, peers are entered under family or surname, with a cross-reference from the title. The reference value to indexers is incalculable. As an authority for identification and for accuracy of names and spellings, dates and other factual matters, it is one of the first sources we consult for British figures. The new version has enriched this source by increasing the number of people included from 29,333 to 54,922 (the number of women increasing from 1758 to 5627). This total breaks down into 49,705 individuals and 408 family or group articles. The editors have taken the opportunity to capture a number of people omitted from the original, so we now find ‘Hopkins, Gerard Manley, (1844–1889, poet’ and ‘Beeton, Isabella Mary (1836–1863), writer on household management and journalist’, both absent from the old DNB but achieving celebrity after its publication. Old articles have been subject to revision and improvement. I was keen to see whether a small error I had noticed and marked in the original had been picked up, where the two William Lowthers had been labelled second and third ears of Lonsdale. Indeed it had: they now appear correctly as first and second ears. Confidence is restored. Revised forms of names are used, thus it is Boudicca, with a cross-reference from the old form of Boadicea, as used in the old DNB.

Of special interest to Society members will be the coverage of former indexers. I naturally turned to find ‘Wheatley, Henry Benjamin (1838–1917, bibliographer and editor’, which gives due notice to his indexing endeavours and to the medal named in his honour. This is a new entry, contributed by J. D. Lee. Another new one is ‘Carey, Gordon Vero (1886–1969), indexer and headmaster’ (surely a fair order of importance). I read this elegant and informative piece (by Geraldine Beare) with great enjoyment, learning, among other things, that Carey was a rugby blue and took the first kick-off at Twickenham in 1909; also that he died in the small local hospital in the town where I now live, a mere stone’s throw from the home of another eminent indexer, James Thornton (not included, alas). Geraldine Beare also contributes the entries on Gerald Duckworth, William Dugdale and Gilfred Norman Knight. Under ‘Beer, Esmond Samuel de (1895–1990)’ we read of his definitive edition of the John Evelyn Diary that, ‘the whole [is] made accessible by a large and exemplary index’. That heading further demonstrates the arrangement: elsewhere he would be expected to appear (as in Who’s Who) under ‘de Beer’. De la Mare is similarly treated, appearing under ‘Mare, Walter John de la Mare (1873–1956)’, not, I confess, the first place I would look, but at least consistent, and buttressed by a cross-reference.

At 60 volumes, this is not a work that many of us will be popping to the local bookshop to purchase, and if we do we should remember to take a wheelbarrow, but it will be a magnetic draw in whatever libraries we consult: a dependable and magnificent reference work on which we can all rely. But there is also an online version, updated three times a year, offering sophisticated search facilities, such as people by place, period, occupation, religious affiliation, a references search, a portraits search and a contributor search. (The Introduction to the online version can be found at www.oup.com/oxforddnb/info/dictionary/intro/onlineintro/.)

Douglas Matthews, former Librarian, the London Library, and freelance indexer

The devil’s dictionary. Ambrose Bierce. London: Folio Society, 2003. xviii, 365 pp., 26 ornamental initial letters illustrated by Peter Forster: £18.95 (The dictionary may also be found on the American Literary Classics Website at www.americanliterature.com/DD/DD1INDEX.HTML.)

Ambrose Bierce is less well known in this country than he deserves. His short stories, notably ‘An occurrence at Owl Creek Bridge’, are occasionally anthologized, and there are tantalizing references to The devil’s dictionary. It seems almost to have the status of a cult book. It was a great favourite of Frank Muir who relished its quirkiness and word play. Now the Folio Society has issued an edition with an appreciative introduction by Miles Kington and illustrations by Peter Forster. It is a characteristic Folio Society publication, carefully printed and bound in buckram with the trade mark marbled paper sides and gold-tooled spine, in a slipcase and small enough to pop into a handbag.

Bierce was born in Ohio in 1842 and served in the American Civil War, in which he suffered a serious wound. After the war he became a journalist in California, married and moved to England where he lived from 1872 to 1875 before returning to the United States to resume his writing. In 1913 he went to Mexico at the time of Pancho Villa’s revolution and is there presumed to have ended his life, though how and where remains a mystery. The devil’s dictionary began life as a series of articles in a San Francisco weekly paper, and first appeared, in book form as The cynic’s word book. Since then it has been reprinted many times, as well as constituting a volume in Bierce’s 12-volume Collected works (1909–12).

We do not often read dictionaries for entertainment. They are functional works. Yule and Burrell’s Hobson-Jobson can be opened for enjoyment, but it remains first and foremost a reference work, and much of the pleasure comes from the exotic words rather than the definitions. Samuel Johnson’s Dictionary of the English language is the great exception, though Johnson’s declared aim was to attempt to fix the language. Still, who can resist the appeal of some of his definitions?:

oats – ‘a grain, which in England is generally given to horses, but in Scotland supports the people’.
patriotism – ‘the last refuge of a scoundrel’.
cynic – ‘a philosopher of the snarling or currish sort’.

These, of course, are hardly definitions at all, but self-indulgent expressions of prejudice or put-downs. Modern lexicographers
would hardly dare violate the scientific practices of their profession, though Chambers twentieth century dictionary offers a couple of gems:

jaywalker – ‘a careless pedestrian whom motorists are expected to avoid running down’.
middle-aged – ‘between youth and old age, variously reckoned to suit the reckoner’.

The devil’s dictionary surpasses all these. It is very selective in the words it lists, but is a true dictionary in that it gives definitions, occasional etymologies, and frequent illustrations, often invented by the author. They best speak for themselves, as with this little sequence of short entries:

Kill, v.t. To create a vacancy without nominating a successor.
Kilt, n. A costume sometimes worn by Scotchmen in America and Americans in Scotland.
Kiss, n. A word invented by the poets as a rhyme for ‘bliss’.

Sandwiched between them is a long disquisition on King’s Evil, which is practically a history of the scrofula and its supposed treatment by royal touch, and gives an unexpectedly serious tone to the work. However, mostly the definitions are sardonic, curmudgeonly, world-weary, cynical, eccentric, unconventional, rational in a topsy-turvy way, thought-provoking and nearly always very comical. A few more examples will illustrate the character of the work:

Conservative, n. A statesman who is enamored of existing evils, as distinguished from the Liberal, who wishes to replace them with others.
Erudition, n. Dust shaken out of a book into an empty skull.
Hope, n. Desire and expectation rolled into one.
Pray, v. To ask that the laws of the universe be annulled in behalf of a single petitioner confessedly unworthy.
Price, n. Value, plus a reasonable sum for the wear and tear of conscience in demanding it.

I have heard that the hardest test for any lexicographer is the treatment of the common particles (words such as ‘by’, ‘in’ and other similar prepositions, conjunctions and the like). Bierce makes no attempt at these, but does go to town on some individual letters. His account of ‘J’ achieves a kind of surrealism, but perfectly illustrates his literary fluency, his learning and his penchant for the morbid:

J is a consonant in English, but some nations use it as a vowel – than which nothing could be more absurd. Its original form, which has been but slightly modified, was that of the tail of a subdued dog, and it was not a letter but a character, standing for a Latin verb, jacere, ‘to throw’, because when a stone is thrown at a dog the dog’s tail assumes that shape. This is the origin of the letter, as expounded by the renowned Dr Jocol Bumer, of the University of Belgrade, who established his conclusions on the subject in a work of three quarto volumes and committed suicide on being reminded that the j in the Roman alphabet had originally no curl.

So now we know. But finally, a peep at what Bierce himself says about a Johnsonian word:

Patriotism, n. Combustible rubbish ready to the torch of anyone ambitious to illuminate his name. In Dr Johnson’s famous dictionary patriotism is defined as the last refuge of a scoundrel. With all due respect to an enlightened but inferior lexicographer I beg to submit that it is the first.

This has long been one of my favourite books. Put in a thumb anywhere and it will come out with a juicy, refreshing plum, generally on the tart side, but always full of flavour and satisfying. If you give one for a present be sure to buy a second for yourself, because you will never want to give it up.

Douglas Matthews, former Librarian, the London Library, and freelance indexer

English language and reading


The author of The Cambridge guide to English usage, Pam Peters, is the Director of the Dictionary Research Centre and of the Graduate Program in Editing and Publishing at Macquarie University (Sydney, NSW), and wrote six chapters of the Style manual (John Wiley Australia, 2002). Consultants for this book included David Crystal, whose book The stories of English was reviewed in the last issue of The Indexer.

The Guide contains a mixture of topics (for example, indexing) as well as discussions of specific words or parts of words (e.g. ‘Croatia’, ‘cross-‘, ‘crumby or crummy’, and ‘-etic/-xic’). The discussion is based on statistics from databases (corpora) of written and spoken English.

In addition to the alphabetic section, the book contains nine appendixes including a perpetual calendar, proofreading marks, and formats for letters, memos and e-mail.

Because the book is based on statistics from English usage recorded in the corpora, it is descriptive (telling us what happens) rather than prescriptive (telling us what we should do). One of the strengths of the book is its explanations that enable us to better understand our own usage of English. For example, I have trouble differentiating ‘lose’ and ‘loose’, and Peters explains that lose is a ‘spelling headache’ because there are no spelling analogies for it except ‘whose’.

International Englishes

British usage has come from the British National Corpus (BNC), and American English from the Cambridge International Corpus (CACE). Smaller databases exist for Australian, New Zealand and Canadian English. South African English is more complicated, as it is made up of Afrikaans English, Black South African English, South African Indian English and the English of the British community in South Africa. In addition to data from the corpora, a range of dictionaries, grammars and style guides were researched, and survey data is included.

The only area in which Peters veers towards the prescriptive is in boxed recommendations in which she suggests the best usage for international audiences. For example, she recommends ‘pediatrics’ rather than ‘paediatrics’ because its use is more widespread around the world (Americans, Canadians and Australian doctors prefer ‘pediatrics’, while Australian writers and the British prefer ‘paediatrics’). ‘Catalogue’ is preferred to ‘catalog’ as it is well established in both American and British English and is more linguistically regular as a base for the verb forms ‘catalogued’ and ‘cataloguing’. (‘Catalog’ is widely used in American libraries, but in other fields ‘catalogue’ is used. The two forms
have equal status in Webster's Third, and are equally current in written texts in CCAE.)

Indexing

There is half a page on indexing, starting with the need for an index for almost any nonfiction book whose material is not already presented in alphabetical order. The Guide itself is arranged alphabetically and does not have an index. This almost works, but there are not quite enough cross-references within the text. For example, there is no reference from 'conjunctions and conjuncts' to 'zero conjunction', and there is no direct reference from 'indexing' to the entry 'Mac or Mc', where there is a substantial section on 'Indexing names with Mac and Mc' (there is an indirect chain via 'alphabetical order'). There is also no reference from 'titles to names' (which has a section 'Titles and names'), although again there is a circuitous route via the entry 'forms of address'.

Pam Peters briefly mentions the location and type size of indexes, the use of specialized indexes (such as of names) and the selection of index terms. She describes index entries being set 'broken off' or 'run in' (American) or 'run on' (British). I have never heard the phrase 'broken off', and it is not in the indexes to Mulvany's Indexing books (which uses 'indented style') or Wellisch's Indexing from A to Z (first edition) which has a reference 'indented style of subheadings see line-by-line style'). The AS/NZS 999:1999 standard (based on ISO 999:1996) has the heading 'Set-out (indented) style versus run-on (paragraph) style for layout'.

There is also content relevant to indexes under 'numbers and number style', where the content on spans of numbers seems to be derived from style guides rather than the database, and under 'alphabetical order' (or should that be alphabetic? – See ‘ic/ical’).

The five-page bibliography does not include the three current standard indexing textbooks, but does include Indexing the art of by G Norman Knight (1979), which is now out of print. The URL of one of the societies of indexers’ websites would have been a useful addition.

Plurals

A good example of the background given in the discussion of language choices is the plural form of ‘thesaurus’. The section headed ‘thesaurus’ notes that database evidence shows that British writers prefer ‘thesauri’ while American writers prefer ‘thesauruses’. There is a reference from ‘thesaurus’ to the more general ‘us’ section, which points out that there are four categories of words ending in ‘us’:

- Latin plural ending in ‘i’ (e.g. fungus/fungi)
- Latin plural the same as the singular (e.g. apparatus/apparatus) – in this case the English plural uses ‘es’ to avoid ambiguity
- Latin plural ending in ‘ra’ (e.g. opus/operas)  
- non-Latin nouns which need not take a Latin plural, although some do (e.g. thesaurus/thesauri/thesauruses).

Should you buy it?

Does an indexer need this book? If you work on multi-authored works without centralized editorial control, and need to make stylistic decisions about word forms (particularly hyphenation), you will find this book useful. The distinctions between usage in different countries could also be relevant.

There are also many topics of general interest including dating systems, inclusive language/political correctness, report writing, and emoticons (including Asian emoticons, which work in the vertical plane, such as (^_^) as a smiley and (Y_Y) for crying).

I have found this book useful and stimulating, and I think most indexers will enjoy dipping into it for its background information on the way the English language works, and will find it enhances their use of language within indexes.

Glenda Browne, indexer, writer and teacher


Hastily on the heels of Lynne Truss’s success story of last year, Eats, shoots and leaves, comes journalist and broadcaster John Humphrys’ Lost for words: the mangling and manipulating of the English language, an ardent denunciation of clichés that arguably manages to get two into its own title. This is, of course, not entirely fair: the need for a snappy title and some neat alliteration must have been irresistible, particularly for a journalist; and Humphrys’ book ranges beyond cliché to cover a spectrum of modern-day grammatical solecisms. But it is clearly a rushed and opportunistic production that at best – to risk another hackneyed phrase – is a missed opportunity.

The development of English (or rather Englishes), the characteristic capacity of the language to absorb and assimilate foreign interlopers and necessary neologisms, and the constant evolution of its rules and usages to serve and reflect changing needs and different, increasingly global, functions, are all subjects of fascinating importance, worthy of careful study. See, not least, Christine Shuttleworth’s review of David Crystal’s The stories of English in The Indexer (24(2): 114–115). Occasionally Humphrys offers the glimmerings of such analysis, as in a section loosely covering the ethical obfuscations that result from euphemistic expression, where he argues (not altogether conclusively) that the replacement of ‘abortion’ with ‘termination’ focuses attention on the mere ending of a process, and allows one to distance oneself from any thoughts of the foetus. But such instances of thought provocation are all too rare. Perhaps we shouldn’t be surprised: early on Humphrys asserts – in an astonishing phrase for a book studying the use of language – that ‘Language is always on the wing. It cannot be examined or analysed to see whether it works.’ This will come as news to David Crystal, who gets the tiniest of mentions, complete with a characteristic Humphrys anti-intellectual jibe. This side-stepping of analysis allows Humphrys to concentrate on a coy but excluding concept of ‘our common language’ – by which he seems to mean English as it is written by people who think like John Humphrys – that is as simplified and specious as that other favourite of his, ‘common sense’.

Too often Humphrys’ argument fails to distinguish between long-term trends in the development (or decline, as he would have it) of English and the most ephemeral or transient of British and American teenage usages. Anyone reading his book in five years’ time will struggle to know what he’s talking about when he rails against the phonetically rendered contraction ‘innit’ or the over-extended exclamation ‘whatever’. But the blackest of Humphrys’ bêtê noires are management jargon and double-speak and their ineluctable spread into all areas of government and the civil service. So much of the book is little more than a catalogue of quotations from official and corporate reports and pronouncements, ranging from the infelicitous to the impenetrable. More fun, though – and as much meaning – could be had from flicking through the fortnightly rubric in Private Eye ‘Pseudos Corporate’.

Unfortunately it is hard to escape the idea that Humphrys is part of the malaise that he is charting. His relentless journalisme is all so insensitive; anyone who has ever switched off the Today programme in exasperation will be wearily familiar with the constant carpings, why-oh-whyness of his tone. It needn’t be this way: the writers and
academics who formed the Society for Pure English, active in the first half of the twentieth century, demonstrated in their tracts on grammar, pronunciation and etymology that one could be a purist without being dogmatic. Humphrys' volume bears very unhappy comparison with the work of a true prose stylist on the subject, such as the Society's co-founder Logan Pearsall Smith.

For all its lightness of analysis, *Lost for words* is not even that amusing, despite Rod Liddle's dust-jacket assertion that 'Humphrys is passionate about language – and very funny too', which may be repayment from Humphrys' 'friend and former editor' for the claim in the book that Liddle is 'one of the most promising writers of his generation'. The only passage that raised a wry smile on my reading reached Humphrys third-hand, via that friend to indexers Bernard Levin: the translation by the euphemism-puncturing writer Marghanita Laski of "simple, inexpensive gowns for the fuller figure" into "nasty, cheap dresses for fat old women".

Oh, and one final thing: the book doesn't have an index.

Christopher Phipps, librarian


Writing a book review is like making love to a beautiful woman. I won't try to justify the comparison, but Swiss Toni would no doubt be able to do so. He is the smooth-talking car salesman from the BBC's *Fast Show*, who applies this concept to everything from selling a car to making a cup of coffee. If he wished for some variety in his conversation, he could do no better than to study this superb collection of catchphrases, clichés, euphemisms, nannystms, format phrases (see below), idioms, quotations, sayings, slogans and stock phrases compiled by Nigel Rees, the well-known broadcaster and writer of more than 50 books on the popular aspects of the English language.

Mr Rees knows his onions. But even a book of 6,000 phrases has its omissions. I found accident waiting to happen, but not walking disaster area; end of an era, but not blessed release or it's what he/ she would have wanted; bells and smells, but not happy-clappy; dream on, but not in your dreams or you wish. A number of other familiar phrases, such as tell me about it and (he/she) scrubs up well (looks more attractive when dressed up to go out) are not there. How weird is that?

The sources of catchphrases are various: literature has provided some, such as Evelyn Waugh’s *Up to a point*, Lord Copper, and James Forsyte’s lament in Galsworthy’s *Forstye Saga*, nobody tells me anything. Many come from radio or television programmes past and present, as well as from commercials. *Calm down, dear – it’s only a commercial* is probably too new to have been included, though I would have expected to find Rolf Harris’s *can you see what it is yet? But is that your final answer? and phone a friend from ‘Who Wants to Be a Millionaire?’ are there, as is you are the weakest link – goodbye, and there is a generous selection of older phrases from, for example, *ITMA* and *The Goon Show*. Anyone of a certain age who used to watch the ITV show *Sunday Night at the London Palladium* will remember the ‘Beat the clock’ section with the instruction to rearrange the following words into a well-known phrase or saying.

The world of journalism has produced such phrases as shock horror and mercy dash (one is in, the other not), and the satirical magazine *Private Eye* has been a rich source, popularizing, if not always originating, references such as tired and emotional to describe a politician’s embarrassing behaviour under the influence of alcohol, or the joke headline phew, what a scorch! The euphemistic term for sexual relations, [Ugandan discussions](#), is in fact to be found here as (to) discuss Ugandan affairs – there is some cross-referencing, but often tracking down the particular version of the expression you want is a matter of guesswork.

Politics provides terms such as spin doctor, expletive deleted and economical with the truth. The remark you might think that; I couldn’t possibly comment originated from Michael Dobbs’s TV dramas *House of Cards* and *To Play the King* but, Rees tells us, was subsequently self-consciously poached by real-life politicians, adding to its popularity.

Some phrases seem to come from we know not where. Are you a fashion victim? Who first said or wrote been there, done that, bought the T-shirt or declared that another person was a few sandwiches short of a picnic? This latter expression is given here as a few vouchers short of a pop-up toaster, and variants of this and many other phrases are recorded in discursive and entertaining entries. *It’s past its sell-by date*, first recorded in 1973, has long been past its own sell-by date and has been superseded by *it’s so last week* (it’s so last century was popular for a brief few weeks in January 2000).

The word ‘cliché’ rears its ugly head. One type is what Rees calls the format phrase, where the blanks are filled in according to need, such as –gate to denote a political or royal scandal (originating in the Watergate affair, but becoming increasingly meaningless), happiness is – (whatever you choose), or – from Hell (something really ghastly, such as the office party from Hell). Then there are all-purpose nouns such as situation (everything used to be an ongoing situation), experience (a marketing cliché – could it possibly be based on the band name, the Jimi Hendrix Experience?), syndrome and solution. Private Eye keeps a handy eye (another phrase not in this book) on such trends and currently has two columns, ‘Solutions’ (suitcases are advertised as ‘Samsonite travel solutions’) and ‘The Neophiliacs’ – ‘nasty is the new black’, ‘comedy is the new rock ’n’ roll’, ‘60 is the new 40’, and so on.

Samuel Johnson is quoted on the book’s dust jacket as follows: ‘A great part of their [the common people’s] language is proverbial. If anything rocks at all, they say it rocks like a cradle; and in this way they go on.’ And how true those words are, even today.

Christine Shuttleworth, freelance indexer and translator


This is a volume devoted to the question of form in poetry, aiming ‘to define the range of formal possibilities available to poets today’. Eighty-five modern writers each contribute a poem together with an essay on its composition and form; there is also an 18-page glossary describing and illustrating 29 different literary forms, from acrostic to word golf. Among the 85 poems, the editor states, ‘you will find villanelles, pantoums, prose poems, sonnets, songs, narratives, commentaries, rhymed poems, free verse, a poem in the form of a musical fugue, a poem in the form of baseball lineups, a poem in the form of an index to a non-existent book, a poem based on a principle of alliteration, and a sonnet containing fewer than 14 words’.

Yes – ‘a poem in the form of an index’. An *Indexer* article of 1994 proposed, ‘The indexer uses language in a creative way, not unlike a poet . . . The indexer, like the poet, focuses intensely on his subject in order to present a concept in an index . . . using language in a condensed form is a skill common to the indexer and the poet.’ 2 Paul Violi, adjunct professor at New York University, in his contribution to this volume, a poem entitled ‘Index’, illustrates these features, and besides, as he explains in his accompanying essay:
... an index, with its fragmentary lines, suggested a way to catch both the quick, haphazard changes a character [who was not quite the master of his fate] would endure and his increasingly scrambled perception of them. As I assembled the poem it began to resemble a chronology... and gave the static index, which was developing imagistically, a linear movement as well... I’d set-off and continue to play-off an ‘argument’ between the neutral if not deadpan tone and the wild particulars of the life it describes.

‘Index’ itself, 59 lines long, in authentic index format, resembles the short story by J. B. Ballard, titled ‘The index’, which likewise wryly presents the perplexing narrative of a life laid out as an index.3 So now indexes (albeit to imaginary books) have appeared in two distinct literary forms.

These are some enticing entries from Violi’s poem, illustrating his claims above:

Hudney, Sutej...
Enteres academy, honors 84
Arrest and bewilderment 85
Marriage 95
Weakness of character, inconstancy 101
First signs of illness, advocacy of celibacy 106, 107
Advocates abolishment of celibacy 110
Consequences of fame, violent rows, professional disputes 118, 119
Disavows all his work 120
Bigamy, scandals, illness, admittance of being ‘easily crazed, like snow’ 128
Arrested for selling sacks of wind to gullible peasants 146
Impression and bewilderment 147
Last words 184, 185, 186, 187, 188, 189, 190

The presentation of items in this volume is not well organized. The contents list gives only the names of poets contributing; the poems’ titles are nowhere listed. Shamefully, the book lacks an actual index. The glossary includes many fascinating examples of literary forms by various writers – Auden, Eliot, Herbert, Pasternak – none traceable, alas.

Notes
1 For an example not of a poem in the form of a index but of an index which turns out a poem, see Indexes Reviewed, Hordern House Rare Books: The encyclopedia of exploration, p. 150.

Hazel K. Bell, freelance indexer

Other subjects


The unsung sixties recounts the development of 36 social organizations founded in Britain in the 1960s. Each of them has a chapter to itself, in a standard format: about ten pages on the development of the organization, followed by five or six on the biography of the founders, told by the founders themselves in first-person narrative. The name of each founder appears at the head of the relevant chapter, as if that of the author.

There is a 21-page index of names only, printed in a single column with no indentations – although plenty of the organizations’ names run over to a second line. Only single page references are given: there are no page runs; and references are given only when the name appears on the page cited, and in exactly the form cited. Thus we have separate entries for organizations’ names with or without their acronyms in brackets following them, and for places beginning St with or without a full stop. The whole chapter devoted to each organization is shown as referred to only on the page where the name occurs, usually just the first page with the chapter title; the founders’ names, too, appear in the index cited only on the first page of the chapter, but not on the pages where they recount their biographies, as there they are designated only ‘I’. So these entries, for example, given thus in the index:

Advisory Centre for Education (ACE) 394, 484
Blackler, Rosamunde 64
Brandon, David 35
Centrepoint 35, 67, 475
Crisis 42, 52, 53, 476
GALS (Girls Alone in London Service) 64
Radford, Jim x, 24
Shearman, Bill 52
Ware, Eileen 19

should in fact be expanded to:

Advisory Centre for Education (ACE) 394–409, 484
Blackler, Rosamunde 64, 71–80
Brandon, David 35, 44–57
Centrepoint 35–51, 67, 475
Crisis 42, 52–63, 476
GALS (Girls Alone in London Service) 64–80
Radford, Jim x, 9–18, 24
Shearman, Bill 52, 57–63
Ware, Eileen 19, 29–34

The selection of entries below shows duplicate entries for the same subject with differing page references – sometimes far separated alphabetically – and misuse of see and see also (as well as no indication of the whole page runs):

Advisory Service for Squatters (ASS) 475
Campaign Against Racial Discrimination x
Campaign Against Racial Discrimination (CARD) 169
Campaign for Homosexual Equality xi
Campaign for Homosexual Equality (CHE) 230, 479
Community Transport 144
Community Transport (CT) 477
Friend [sic] 234; Society of 372
Hoodless, Elisabeth 435
Hoodless, Liz 416 [yes, they do appear to be the same woman]
Indian Workers Association 171. See Southall Indian Workers Association
Joseph, Keith 360, 363
Joseph, Sir Keith 310, 311 [yes . . .]
Mountbatten Report 112
Mountbatten Report 202
National Association for the Welfare of Children in Hospital 321; Leicester Branch 325
NAWCH 325, 364. See National Association for the Welfare of Children in Hospital
OCRI 180. See Oxford Committee for Racial Integration (OCRI)
O’Malley 160
O’Malley, John and Jan 166
Indexing classics

The Society of Indexers has published limited edition facsimile reprints of two indexing classics that encapsulate with wit and humour the very spirit in which indexers still work.

Wheatley, H. B. What is an index? A few notes on indexes and indexers.
First published 1879 by Longmans, Green & Co. for the Index Society. Facsimile edn. Society of Indexers, 2002, 132 pp. £15.00 (£17.00 overseas)* for members of indexing societies.

Wheatley, H. B. How to make an index.
First published 1902 by Elliott Stock. Facsimile edn. Society of Indexers, 2002, 236 pp. £18.00 (£20.00 overseas)* for members of indexing societies.

* Both books for £30.00 (£34.00 overseas)

To order these and other SI publications, visit www.indexers.org.uk for a downloadable order form or contact: Sales Administrator, Society of Indexers, Blades Enterprise Centre, John Street, Sheffield S2 4SU; UK Tel: +44 (0)114 292 2350 Fax: +44 (0)114 292 2351 email: admin@indexers.org.uk

As for alphabetical order, the following disastrous sequence is clearly attributable to a failure to instruct the computer on the sorting of St., St and Saint:

St Joseph’s 81
St Mungo’s 35
St. Anne’s Church, Soho 35
St. Barnabas Church, 53
St. Botolph’s Church 198
St. Joseph’s 84

However, there are some signs of human intervention in this index. One can see how a computer would first have arranged the following three entries, but presumably ‘Family’ was added to ‘Service Unit’ after the computer had done its work, but without anyone thinking to put the revised entry in its proper place.

Serota, Baroness 363
Family Service Unit 120
Sexual Offences Act 1967 233

But this group:

Labour Party 31, 117, . . . [a full line of page numbers]
New Labour 143
Laing, RD 187

is more baffling: perhaps the second entry was intended to be a subheading, and some disaster befell its semi-colon?

Harold Evans 399

appears sic, in the H section, between Harding, Dr and Harris, Rufus. As he is referred to as Harry Evans at the foot of page 399, to become Harold Evans in the top line of page 400, there has clearly been some sort of editorial intervention here too.

But then, if a human being was making amendments to this index, why stop at these?
The unsung sixties’ full page of acknowledgements offers appreciation of ‘invaluable reading and copy-editing’ and of the ‘enthusiasm’ of the publisher, but makes no mention of an indexer.

This index is, surely, a terrible example of the dangers of relying on a computer to do the indexing automatically, barely checking the result, and being left with a job that appears almost untouched by human mind.

Hazel K. Bell, freelance indexer

Publications received and publications noted
