September 2005

Index Appreciation: A Publisher's Brief Guide

Sylvia K. Miller

Routledge, Taylor & Francis Group LLC, sylviak.miller@gmail.com

Index Appreciation: A Publisher's Brief Guide

by **Sylvia K. Miller** (Formerly Publishing Director, Reference, Routledge, Taylor & Francis Group LLC) <sylviak.miller@gmail.com>

stock), and a sales agent (who visits the suppliers and provides product information). Each of these people get a percentage of the sale, ranging from 5 to 25 percent each. Many companies also pay a PR firm to handle conferences, catalogues, direct mail, and advertising. Given the additional shipping and warehousing costs involved, overseas sales can be prohibitive for smaller publishers unless they find innovative ways to reach customers (and, in the process, cut the length of the food chain).

Need for change

It will be interesting to see how much longer these traditional but inefficient arrangements survive, given the move to online resources and the ease with which overseas customers can order books directly from **Amazon**, **eBay**, and other U.S. retailers, as well as from the international distributors (which now include U.S. companies such as **Baker & Taylor**, **Ingrams**, and others). But that's by no means the only, or the most important, area for change. We are at a watershed, a time when reference has to prove its value to a world in which knowledge is valued, but when consumers often expect it to be free, online, and always up to date.

Online commerce at its best has superb customer service, with quick response, readiness to change based on new information, and direct interaction between producers and their customers. In the reference world, the opposite is true: there's a great reluctance to change (librarians are far more comfortable with technology than most senior editors and publishers), and there is little interaction between producers and customers.

There are some reference publishers who still resist the digital world. I think of them as horse-and-buggy publishers, and they'll grumble all the way to the end of the line. Then there are the Mr Toads, publishers who always have to have the latest and greatest model. (And, like *Mr Toad in The Wind in the Willows*, they're usually rich enough to buy another, and another, and another.) I don't think any of us knows what will be seen as the digital reference Model T — that's how early a stage we're at in the development of online reference. Developments in publishing will echo what's happened in other online businesses: a great deal of consolidation, sometimes successful and often not. As delivery technologies became more sophisticated and affordable, there will be challenges from a variety of small specialist providers who can quickly adapt, focus, and refine their products to meet the needs of a particular audience.

I deeply believe that reference publishers supply "a sound article that is really needed:" knowledge about our world, past, present, and future. The question we face today is how to work with scholars, librarians, and our readers to create an industry that is both profitable and sustainable, a vital strand in today's World Wide Web. 🌳

Surprisingly, the seemingly mundane book index can be the cause of great anxiety, late nights of work on nearly impossible deadlines, and dramatic confrontations. When the index brews trouble, the explosion occurs at the most inconvenient moment, at the end of the production process when the publication date is set in stone and press time is reserved.

For example, if an indexer boasts that she is so pleased with her work that she wants to enter her newly completed index manuscript in an indexing contest, while at the same time the author of the book complains in desperate tones that the indexer has completely misunderstood and misrepresented a lifetime of work, then you have to learn something besides editorial diplomacy. The clearest lessons often take shape from the worst problems. In this case, I learned how important it is that an indexer understand and appreciate the outlook and purpose of a book. If the indexer is incompatible with the book, and the problem is not caught early on, disaster awaits.

Over my years in scholarly publishing, such problems have illuminated important points about indexing that eventually formed a sort of pathway or strategy for commissioning and editing a useful and suitable index. However, I have never taken an indexing course, and I still approach indexers with a measure of awe. With this as a disclaimer, I would like to share with librarians some of the basic decisions that go into creating an index for a scholarly book in the hopes of dispelling a couple of common misconceptions and renewing appreciation for the indexing art.

Type of Index

Historians of the book find the index a fascinating window into the interests and approaches of the supposed reader at the time of publication. The editor first asks, Who will use the index, how often, and for what purpose? A novel almost never needs an index, but a biography does because some researchers may not read the whole book or may return to the book to locate specific events. A monograph needs a decent index for the same reasons, although its index will not have to take into consideration a nonprofessional reader who might not recognize certain terminology, whereas the index to a reference book usually will.

An index of basic terms, or proper names only, will often suffice. However, an analytical index will give the greatest number of potential readers a port of entry to the content of the book.

Separate or multiple indexes are some-

times considered, such as an index of case names in a book on legal history. Fearing that researchers looking for quick answers might not notice multiple indexes in the back of the book and might instead mistakenly assume that the legal case they are looking for is absent, I always argue for a combined index.

The Indexer

The indexer for most books is the author, who knows the content of the book best. The production editor usually provides a set of basic indexing instructions, and the author does his or her best, with varied results. Some authors are intrigued by the task and turn out to be talented at thinking of their ideas in this new structure. Others, however, might be wonderful writers but the first to admit that they cannot index, while still others will be affected by weariness with the book project's endless editors' queries and typographical errors generally, impatience to get beyond this stage and enjoy the well-deserved book party, or distraction by a flooded basement and a visit by five grandchildren.

The publisher hires a professional indexer when (a) the author refuses to attempt the task and prefers to see the cost of the index charged against future royalties and (b) the index is such an important feature of the project that none other than a professional indexer will do. We try to choose a professional indexer who has some knowledge and appreciation of the subject of the book. Indexers for scientific and medical books are a separate, highly prized, breed. The indexer must be capable of reading a passage of the book and summing it up in a brief, understandable term or phrase that may not actually appear on that page. (This capability is important to note because a basic electronic search engine cannot do it.)

The indexer must also understand the intended audience. "What will people look up?" must be the indexer's mantra throughout the indexing process. The simple terminology that students often research might occur infrequently in the book, whereas the information that they seek might be there in depth, disguised, as it were, in different, perhaps higher-level, language. The indexer and the editor might discuss connections that could be made for accessibility.

Necessary Tools

The indexer must use indexing software, to allow flexibility in making last-minute changes. For example, the publisher might decide that list style subentries must be changed to run-on style, to save space on the typeset pages, or preliminary page numbers called "false folios" might have to be automatically converted to final page numbers when the index is complete. The preferred spelling of a recurring term might be changed. Or a large portion of the book might
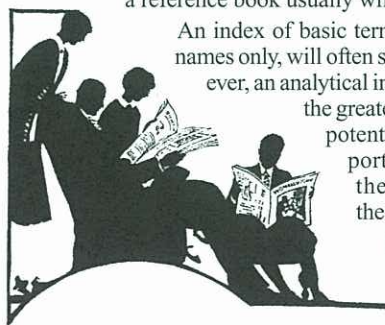
be repaginated late in the production process because of a serious error, such as omission of a crucial illustration or a paragraph of text. Indexers who use only word-processing software, or, heaven forfend, index cards, will not be able to cope with such changes.

If the project is very large, the indexer should have helpers whom she has trained and with whom she has been known to work compatibly in the past. Most indexers cannot take on or improve one another's work, nor can they work together on the same index.

The Cost

The professionally prepared analytical index is expensive. For example, it might cost between ten and twenty thousand dollars to index a three-volume reference set. A substantial sum on its own, this is also an increasingly significant proportion of the total book budget, because while typesetting and some other production costs have been reduced over the last several years due to offshore services and improved automation, indexing is one area in which quality cannot be maintained at lower prices. Copyediting is another, although good people do disagree on these points.

The editor and publisher have to make a business decision about indexing quality and depth in the context of the purpose of the book and the sales income expected from it. A series of useless indexes might damage the reputation of the publishing company, but we must also recognize that an extremely detailed index might not sell more books than a moderately detailed one, while the cost difference might be significant. Furthermore, an overly detailed index might waste the researcher's time with too many page references that do not yield significant information; on the other hand, a truly analytical, in-depth index might provide an invaluable port of entry to the content of a work that has an educational mission, such as an encyclopedia. When we balance these various considerations well, we have set the stage for creation of what we might call the suitable, or appropriate, index.

Length and What to Index

The length of the index is directly related to the cost because (a) many indexers charge by the index entry and (b) an overlong index might add unbudgeted pages and signatures to the book. To the cost of the paper must be added the cost of typesetting and proofreading. Therefore, it is crucial to establish guidelines for what to index and what to omit.

It is a common misconception that an excellent index includes everything in the book, including every passing mention of a proper name. In general, it is impossible to index absolutely everything in a fact-filled book. Just think of it: if we list every fact in the book in alphabetical order, we are simply printing the entire book all over again in list form, minus some connecting phrases. We would have to print almost as many index pages as text pages. There are conflicting opinions on ideal index length as a proportion of the original text, and

when people talk about it in terms of page numbers, these discussions are positively silly, because type specifications can change drastically the number of pages an index occupies; however, no one thinks that 90 percent is a practical proportion.

After a relatively small portion of an encyclopedia has been indexed, I always request a sample index from the indexer, along with a cost estimate. Depending on the budget and the type of reference work, we might discuss ways to reduce the indexing to, say, ten entries per double-column page from twelve.

It is customary to omit passing references. In other words, the indexer focuses on listing the most useful — that is, substantive--discussions of, a topic. Therefore, if Michael Apple, Susan Blueberry, Jacob Grape, and Horatio Kiwi all attended a meeting in Geneva in 1953, the indexer might legitimately choose not to index this list of names. Unfortunately, if the published book is assigned to a reviewer whose favorite person in history is Susan Blueberry, it is all too likely that the published review will call the index to task for inexcusable omissions. The indexer's goal, however, was to avoid wasting researchers' time with too many index entries that will be insignificant to most readers.

The reviewer might wish to ascertain if legitimate choices were made consistently; if so, the indexing cannot be considered carelessly, although of course there can be many legitimate differences of opinion on the best indexing choices. Some tend to be more customary, but no decision is easily taken; after all, the people drawn to the indexing profession are rarely cavalier or inconsistent types!

Structure and Design

As a reference publisher, I am fascinated by the challenge of categorizing and organizing knowledge, the ever-unsurmountable task of applying structure to chaos. I appreciate the indexer as a consummate information architect. The index breaks the text into small pieces and reshuffles them. The new order of content created in this daunting process is not only alphabetical but also hierarchical.

Inexperienced indexers will occasionally outline, rather than index, the book. This type of list is similar to a topical table of contents and will not be as useful as a real index.

It is important to include subentries that identify in words the purport of the discussion on the referenced page. Otherwise, strings of page numbers result (for example, Religion and civil liberties, 35-36ff, 72, 85, 103, 256, 258, 390, 424, 435-436, 445, . . .), which are worse than useless because they waste the researcher's time in tedious and mostly fruitless page-shuffling.

Indexers often ask how many levels of subheadings the editor prefers. I always choose one only (that is, main headings will have subentries, but no sub-subentries or sub-sub-subentries). I cannot accept the necessity of more complexity than that; somehow over the years, order and simplicity have become joint concepts in my mind.

An item of lively contention is the method of alphabetization used. Believe it or not, experienced editors can spend hours debating the

advantages and disadvantages of the word-by-word versus the letter-by-letter method! In years past I expounded the virtues of the letter-by-letter method with verve, but eventually I recognized that different methods can be preferable for different fields or topics and that ultimately, consistency is more important than the method chosen.

It is important that type specifications take into consideration all the levels of subheading plus turnover lines for each. If turnover lines do not indent further than the lowest level of subheading, the index will look very confusing and be difficult to use. The likely frequency of turnovers should be taken into consideration when the width of the column is chosen, to avoid a great deal of hyphenation and narrow strings of turnovers that are confusing to the eye.

The type size and number of columns are often chosen to fit the index neatly within the last signature (standard gathering of 16 or 32 pages) of the book. The index is wonderfully shrinkable and expandable via its type design, and production editors gleam with satisfaction when the book has no blank pages at the end.

The Schedule

The index is cursed with a perpetually, unavoidably tight schedule because it must be created last, usually after page proofs are corrected and finalized. This stress is somewhat alleviated by electronic tagging of terms before pages are typeset — resulting in an inferior index, in my view — or, in encyclopedia publishing, indexing letter sections as they are completed, using false folios such as Q-1, Q-2, etc. (the letter Q tends to be complete first, as few subjects generate many Q articles!) and converting them to the final folios as a last step.

However, indexing pages proofs that are not finalized causes much additional correspondence between editor and indexer, as typographical and even editorial errors are corrected and these corrections must be reflected in the index. Correspondence is often generated in the other direction, too, as the indexer inconveniently but inevitably picks out inconsistencies of spelling or fact heretofore unnoticed by copy editors and proofreaders. (For example, on page 22, Blueberry's birth year is 1928, but on page 505 it's 1929; which is correct? Or, the index contains two main entries, each with subentries: "Blueberry, Susan" and "Blueberry, Susan B." Are they the same person? If so, the two entries should be merged.) Editing an index consists mostly of resolving such queries. Sometimes editing also involves adding more blind entries and cross-references to the index. Without these finishing touches, a basically solid index can give an unfortunately careless impression.

Index quality can suffer from rushing, but not as much as you might think. Experienced indexers can reorganize and re-connect the content of a book with amazing speed; quality more often suffers from a too-tight budget or an inexperienced indexer.

Placement

The index should always be placed at the very end of the book. We try to avoid the temp-

tation of placing late-breaking news such as copyright acknowledgments following the index, because what good is the index if the researcher turns to the back of the book and quickly assumes there isn't one?

Into the Future — The Index Goes Electronic

How old-fashioned, you are perhaps musing at this point, to go on at such length about a soon-to-be-obsolete enterprise! Many scholarly books, including reference books, have been loaded into electronic databases without their indexes on the assumption that the automated full-text search will replace the index function — which indeed it does, although often rather poorly. All too often, excessive numbers of search results that are not sensibly prioritized lead researchers through labyrinths of not particularly pertinent text passages, obscuring rather than illuminating the location of the substantive information that they really need.

On the other hand, if automated algorithms are customized to the content through the creation of thesauri to connect similar ideas expressed in different words; if professional evaluations of content authority and usefulness are applied; if the book index is incorporated in the search metadata or, as a second choice, in the offerings of links from one content tidbit or "digital object" to another; then electronic searches can work amazingly well. However, without the analytical choices made by indexers, full-text searching is often no more than a promising mishmash.

In fact, top-notch databases and library online catalogs use similar methods to provide access to large amounts of information. It is exciting to think that as these techniques improve, full-text searching will provide links to relevant material with the ease of a good book index. As automated full-text searching becomes more intelligent — incorporating more and more levels of indexing-like human analysis to connect questions with their most useful answers — new types of meta-searching will increasingly bring valuable content to the light of day. Information will be not only accessible, but locatable as never before, and indexers will never be out of a job! 🌳

*Sylvia K. Miller has twenty years' experience in scholarly publishing, at **Macmillan**, **Scribners**, and **Routledge**.*

ATG Interviews Diane Kovacs

President, Kovacs Consulting Internet & Web Training <diane@kovacs.com>

by **Tom Gilson** (Head, Reference Services, Addlestone Library, College of Charleston, 66 George Street, Charleston, SC 29401; Phone: 843-953-8014; Fax: 843-953-8019) <gilsont@cofc.edu>

ATG: Diane, you began at Bucknell University in 1989 as a reference and government documents librarian and four years later you started your own consulting business. What prompted you to take such a leap of faith? What opportunities did you see that maybe some of the rest of us missed?

DK: Well I compare myself to the penguins that get pushed off the iceberg to test the waters for Sea lions and Orca. But, the real story is that I had so many people asking me to develop training that I felt it necessary to form a business to cope with it and keep it separate from my day job — by 1991 I was at **Kent State University Libraries**. Also, around that time my husband was doing a lot of computer networking consulting and we formed the business together. So initially it was both of us — him doing networking consulting and me doing training.

ATG: It's great that you and your husband were able to match talents enabling you to branch out. By the way, what were your job duties at Kent State? How did that experience play into your decision?

DK: I was just Reference Librarian — my hiring area was Humanities Reference. What I actually did was to do a great deal of the Dialog searching, and then Internet training for both, the library faculty, and the faculty in my own areas of collection responsibility. About 1991 spring — I think — the **Kent State University Library** School Dean **Rosemary DuMont** asked me to design and teach an Internet course for librarian continuing education. Also, the University put me in their list of referral consultants and I began doing Internet training for community organizations/businesses. **ALA**, **MLA**, **SLA** folks all asked me to do programs

on various Internet related topics at conferences. I did a lot more of those back then than I do now — I don't have the travel budget for more than one conference per year. I tend to specialize with **MLA** now.

ATG: You were a pioneer in advocating that librarians collect and organize Websites, and in a sense, create electronic reference collections like we create print reference collections. What common criteria exist for collecting, evaluating, and selecting in both the print and digital world? What criteria are different?

DK: All criteria used in selecting print resources apply in the digital world. If you keep your mind flexible — even the size of the resources matters — e.g., the size of the books matters in terms of how they fit on a shelf — the size of the digital resource matters in terms of how much disk space and bandwidth you have to support it.

The most important aspect that is unique to

electronic resources is archiving. So much of what we use is in e-form only and may disappear if proper attention is not given to making sure it is archived and available. Access is the second most important — libraries provide access to books by buildings being open during certain hours. E-resource access might be 24/7 *if* the technology infrastructure is set up properly to do that and that requires attention be paid to a patron's ability to access resources. Not everyone has a computer at home/office. Not everyone has an understanding of all the hoops you have to jump through to get to a virtual library — especially password /proxy protected e-resources. I'm getting a sharp reminder of this, this summer, as I am teaching an undergraduate information literacy course for nontraditional students. I'm still trying to get them to use email.

ATG: Who are these nontraditional students? In our experience it is highly unusual for students not to be somewhat comfortable with technology. Do all of your non-traditional students have concerns about using technology? What are those concerns?

DK: These are adult learners — working people who are returning to school to pursue further education e.g., business associates/bachelors degrees, administrative assisting, medical assisting, and criminal justice associates degrees. I have a couple of IT bachelors/associates students as well — but ALL of my students are first year/first semester returning adults. Their main concern is time, I think. Out of 30 students — two sections — only one of my students is a regular computer gamer. And none of them were regular users of Internet Messaging — those latter two activities are what I think