Term Selection:
Putting Humpty Dumpty Together, At Last

by Sylvia Coates

here are many aspects to learning how to index. The two main features of the learning process are: first, becoming familiar with the usual indexing conventions, and second, developing term selection skills. Becoming familiar with indexing conventions can be accomplished through any of the numerous good indexing courses available. But the second feature, learning term selection, is more of a challenge.

Term selection instruction can start with teaching rudimentary guidelines. But more finely detailed instruction on term selection is both difficult to develop as well as to learn. This is because term selection is particular to the material being indexed and specific to the needs of the client. An approach to term selection must be tailored to the distinct text being indexed. As such, it is extremely difficult to create a curriculum that adequately conveys the term selection process.

I have, like other experienced indexers, developed my own approach to term selection. I present it here, as a detailed step-by-step process. In no way do I present it as the only feasible approach to term selection, but only as one which has proven useful in my own indexing process.

So, for your consideration, I offer these steps in the term selection process:

1. Term selection begins with *comprehension of the material*.

2. Once the material is understood the information should be broken down into specific components by answering the 'who, what, where, why, when, how come, and under what circumstances' questions. Term selection is based upon the answers to these questions, which, henceforth, will be referred to as the *information questions*.

3. Term selection is irrevocably linked to the *creation of the index structure*.

4. Combining the term selection and index structure processes leads to a *smoother editing process*, an accelerated indexing process, and a cleaner indexing structure.

Comprehension of the Material

Every text has information, or a message, which the author is attempting to convey. This can be found in any kind of text, regardless of subject, density, or intended audience. This is true of the most complex of scholarly material to the most simple of nursery rhymes. The first task of the indexer is to read the text in order to discern the derived knowledge, or information, of that text. Some indexers maintain that it is possible to index material of an unfamiliar discipline. While it may be possible for a very experienced indexer to come up with some kind of term selection there can be no doubt that complete comprehension of the material will result in the superior term selection and index.

It is absolutely required that the indexer comprehends the text material in order to best answer the information questions. Comprehension of the material is necessary to the process of both formulating and answering the pertinent information questions. It is expected that an indexer will have superior reading comprehension skills. But, even with the best of comprehension skills, there may be outside factors affecting the ability to ask and answer the information questions.

There can be quite an array of outside factors contributing to comprehension difficulties. Poorly written or organized text can be a challenge to grasp. Lack of cognizance, due to poorly written text, is why it is more difficult to index a poorly written text than well-written material. It is more problematic to identify the pattern of information in a poorly written

Sylvia Coates has been a freelance indexer since 1989. She lives in the San Francisco Bay area with her husband. Her many hats include mother, grandmother, tennis player, and USDA Basic Indexing instructor.

book and therefore to select index entries. Sometimes the style of writing can also affect comprehension. Anyone who has worked on text written in the deconstruction fashion can testify to the frustration of both understanding and indexing the material. While good term selection is certainly possible for deconstructed text it generally requires a special effort on the part of the indexer. And, of course, an indexer who does not have basic understanding of the discipline being discussed will not be able to ask and answer the information questions required for good term selection.

Asking the Information Questions

The information questions are the "who, what, where, when, why, how come, and under what circumstances" questions of the text. The answers to these questions reveal the information or knowledge derived from the text. These answers can be delineated as main heading term selections.

My experience has been that main headings are most often derived from the "who, what, where" questions. Sometimes, depending on the material, "when" questions can also result in main headings. For example, historical time periods can often be viable main headings. "Why, how come, and under what circumstances" answers often provide the pertinent subheading information. Who and what information, in addition to being used for main headings, can

also be used to provide subheadings which clarify the main headings. For example, a "who" subheading may follow a "what" main heading so as to clarify the "what" main heading information; and, conversely, a "who" main heading may be followed by a "what" subheading.

During the indexing process it is important to differentiate between main headings and subheadings. According to Mulvany's *Indexing Books*, main headings are primary access points for readers. In other words, main headings are where readers will be looking for fundamental units of information. Subheadings present related aspects of the main headings broken down for easier access by readers.[1] It's been my observation that learning the difference between main and subheadings can be difficult for new indexers. It helps to consider the audience or readers of the index. An important aspect of selecting main headings is to anticipate what headings the reader might look up in order to locate specific information. Sub-headings should be selected to help the reader break down that information into units not likely to be otherwise accessed. It must be pointed out here that many main headings are also proper subheadings under other related main headings. Learning to identify main headings and subheadings is an important aspect of good term selection.

As we proceed with the application exercise below, observe that even though all the information questions must be asked, there will not always be answers to all of the questions. Terms are selected only from the information revealed through the questions.

Let us apply this process of identifying the information questions in order to select headings for the familiar children's nursery rhyme *Humpty Dumpty*:

Humpty Dumpty sat on a wall
Humpty Dumpty had a great fall
All the King's horses and all the
King's men
Couldn't put Humpty Dumpty
together again

What is the information, or derived knowledge, in this familiar child's rhyme? Let us assume, just for the benefit of this exercise, that each line is a chapter in a book about Humpty Dumpty. And, to simplify this exercise, I am going to supply hypothetical page numbers to each line.

First, let us examine a poorly constructed index. If the indexer only looks for the "who" or "what" in the text then the index looks something like this:

the fall, 4-6

Humpty Dumpty, 1-3, 4-6, 10-13

King's horses, 7-9, 10-13
King's men, 7-9, 10-13

the wall, 1-3, 4-6

Notice that all the pertinent pieces of information are in the above index. However, there are no relationships between the headings and there is, at best, a poor index structure. Instead of a viable index structure there is only a list of terms. This is an example of an inadequate index.

In contrast to the above approach, if we ask "who, what, where, when, how come, and under what circumstances" questions my first headings would then be:

For the first line/chapter: Humpty Dumpty sat on a wall

We ask, "Who is doing what?" (Humpty Dumpty is sitting); "Where is Humpty Dumpty sitting?" (on a wall); "When and why is Humpty Dumpty sitting on the wall?" (the text doesn't answer the question when or why). So the main headings, and subheadings which clarify those main headings:

These headings recreate the pattern of information in the first line, or chapter, of this rhyme. The who and what questions have been answered and used to create both the main and subheadings above. Moving along to the second line our headings become:

Humpty Dumpty
    wall sitting by, 1-3

the wall
    Humpty Dumpty sitting on, 1-3

For the second line/chapter: Humpty Dumpty had a great fall

We ask, "What has happened?" (Humpty Dumpty had a great fall); "and to whom?" (Humpty Dumpty); "and where?" (from the wall); "and why?" (we don't know). So the main and subheadings:

the fall
    suffered by Humpty Dumpty, 4-6

Humpty Dumpty
    fall suffered by, 4-6

the wall
    Humpty Dumpty's fall from, 4-6

For the third line/chapter: All the King's horses and all the King's men

We ask, "What happens after the fall? (the King's horses and men arrive after Humpty Dumpty's fall)' "who was the action taken by?" (the King's horses and men); "and what was the result of that action?" (we don't know yet); "and why was this action taken?" (we don't know). The main and subheadings:

the fall
    King's horses/men arrive
        following, 7-9

King's horses
    arrival after the fall by Humpty
        Dumpty, 7-9
King's men
    arrival after the fall by Humpty
        Dumpty, 7-9

For the fourth line/chapter: Couldn't put Humpty Dumpty together again

We ask, "What was the result of the fall? (Humpty Dumpty suffered permanent damage from the fall); "who suffered from the fall?" (Humpty Dumpty did); "and who else was involved?" (King's horses and men couldn't fix the damage); "and why?" (we don't know why the damage couldn't be fixed). The main headings and subheadings:

the fall
    permanent nature of damage
        from, 10-13

Humpty Dumpty
    permanent damage suffered by,
        10-13

King's horses
    failure to repair Humpty Dumpty
        by, 10-13
King's men
    failure to repair Humpty Dumpty
        by, 10-13

Putting all of the entries together the index becomes:

- the fall
    - Kings' horses/men arrive following, 7-9
    - permanent nature of damage from, 10-13
    - suffered by Humpty Dumpty, 4-6

- Humpty Dumpty
    - fall by, 4-6
    - permanent damage suffered by, 10-13
    - wall sitting by, 1-3

- King's horses
    - arrival after the fall by Humpty Dumpty, 7-9
    - failure to repair Humpty Dumpty by, 10-13
- King's men
    - arrival after the fall by Humpty Dumpty, 7-9
    - failure to repair Humpty Dumpty by, 10-13

By asking and answering the information questions we end up with qualitative main headings and subheadings for this "index." We were unable to ascertain answers for some of our questions, so therefore there are no entries to that missing information.

This is an opportune time to discuss the link between information and the type of material being indexed: Application-oriented text will generally contain more what, where, when, and under what circumstances information and less why or how come information. Application-oriented text is often found in computer software manuals, how-to books, or corporate-type publications. In contrast, scholarly material generally contains equal amounts of these questions and certainly a higher percentage of why, how come, and under what circumstances information than other types of text.

Term Selection is Irrevocably Linked to Index Structure

We have included in the term selection for the main and subheadings the information revealed in the text. And, tempting though it may be, we have *not* added to that information or interpreted it in any manner.

These entries not only recreate the pattern of information in the nursery rhyme but they are also interrelated. The index also includes multiple access points for retrieval of the same information. Altogether, these term selections create a strong index structure accurately reflecting the information found in the nursery rhyme. The essence of the *Humpty Dumpty* nursery rhyme is instantly revealed through even a casual examination of the index.

Good Term Selection and Editing Process

When good term selection is implemented the editing process is greatly reduced in terms of both effort and time. Since the index structure is created simultaneously with the selection of both main and subheadings there are seldom 'loose ends' to the index. Entries are generally already in their proper place with the appropriate relation to other entries. There is no need to reorganize entries into a cohesive collection or to create parallel structure as entries are already placed in a viable structure with appropriate parallel structure in place. The editing process thereby becomes more of a copyediting nature. Time is saved by not having to restructure the index while editing. The end result is an efficient indexing process with a clean index structure and user-friendly term selection.

Last thoughts

A good index is more than a long list of headings. A good index requires a strong index structure made of headings accurately reflecting the information (or author message) found in the text. A reader should be able to discern, from the superior index, the type and depth of information available in the text. Good term selection, resulting in a strong index structure, is required.

Good term selection requires strong comprehension skills, asking the right questions, and headings based on the answers to those questions. The type of material being indexed must be considered in designing those questions and answers. And finally, selecting appropriate main headings and subheadings will result in a superior index, a happy client, and a job well done! *i*

[1] *see* Mulvany, Nancy, *Indexing Books.* (Chicago: University of Chicago, 1994), p. 70-75.