Creating Better Subject Access with Multiple Vocabularies: Upgrading the Subject Heading List for the Alzheimer's Association

Marilyn J. Smith and Pauline Atherton Cochrane

Although the Alzheimer's Association's Green-Field Library's catalog has been available to local chapters and interested people for some time through modem access, placing the catalog on the World Wide Web would make it available to casual browsers as well as determined searchers. When a review of the subject list revealed numerous inconsistencies and duplications, a new list was generated, giving preference to Medical Subject Headings (MeSH) where possible. The result was a mix of MeSH and Library of Congress Subject Headings (LCSH), augmented by a few local- and reviewer-supplied terms. The new subject authority list gives the Green-Field Library an authoritative list of terms to use when performing original and copy cataloging. It can also be placed with the library's catalog on the Web to aid users in performing searches.

The Green-Field Library, housed at the Chicago headquarters of the Alzheimer's Association, is a comprehensive repository of materials dealing with Alzheimer's Disease. The collection includes about 110 periodicals and nearly 3,800 books and videos. The library, which opened in November 1991, answers 7,000 to 8,000 reference questions per year. This collection has been available, through modem access, to interested persons and Alzheimer's Association chapters throughout the nation. But Patricia Pinkowski, director of the library, wanted to reach casual browsers as well as determined searchers, and decided to place the catalog on the World Wide Web.

A review of the list of subject headings revealed a confusing sprawl. Knowing that searchers must have quick and fruitful access, Pinkowski suggested a revision of the list of subject headings in the library catalog and contacted Pauline Cochrane to consider the situation. The

ensuing discussion resulted in several projects:

1. To provide a subject authority list for the library to use when the cataloger accepts or creates catalog records using *Library of Congress Subject Headings (LCSH)* or the National Library of Medicine's *Medical Subject Headings (MeSH)*,
2. To provide a list of approved subheadings, and
3. To review all class numbers, with the aim of making suggestions for their use as access routes.

We report here on projects 1 and 2. The review of class numbers has not yet been completed.

THE PROBLEM

In a local library catalog, bibliographic records either have or are assigned subject headings from multiple sources. These headings are from *LCSH, MeSH,* another specialized or locally developed thesaurus. Universities often have two primary cataloging units, one in a health sciences library and one in the general university library, respectively assigning *MeSH* and *LCSH* headings. Weintraub (1992) describes one such system at the University of California at San Diego (UCSD) library, where the different subject headings are indexed separately based on the MAchine Readable Cataloging (MARC) tag. In contrast, the Green-Field Library had only one index that contained subject headings from all sources.

Multiple vocabulary catalogs can cause difficulties for searchers. Limiting the search to one subject heading system at a time can result in a failure to retrieve all relevant materials, and even sequential searching of each system won't necessarily lead to full recall because of inconsistencies and unknowns in how the terms were used for the items cataloged. On the other hand, as Olson and Strawn (1997) point out, a universal search of a mixed-system vocabulary can result in duplication of headings, incomplete or partial retrieval, conflicts between headings of one system and "search under" references in the other, and confusion of terms that have one meaning in one system and quite another in the second.

POSSIBLE SOLUTIONS

Chaplan (1995) has suggested some ways of coping with this problem of multiple indexing languages. One is to map—or link—terms. Chaplan (1995, 41) states that: "In a map, the terms in one vocabulary are listed, with an indication of what the equivalent terms are in the other vocabulary. There may also be an indication of whether the term is identical, is a synonym, or holds some other relationship to the term in the second vocabulary."

Another strategy for coping with multiple languages that Chaplan characterizes is merging, which she describes (1995, 40) as ". . . simply combining terms in their original form, with their accompanying cross references and term relationships, from two or more vocabularies into a single alphabetical list, with indications of which vocabulary contains the term." Chaplan (1995) and Olson and Strawn (1997) both used the mapping technique. Chaplan mapped Laborline Thesaurus terms to *LCSH*, while Olson and Strawn mapped *MeSH* and *LCSH*. Our work conformed more closely to the merging technique.

Our objective was to produce a single authority list that gave preference to one system *(MeSH)*, while borrowing from others when it appeared necessary. We chose *MeSH* as the preferred system because the collection contained many works addressing the medical aspects of Alzheimer's Disease and we expected the user groups to be familiar with these headings. Because of this objective, and time and financial constraints, our revision of the subject list of the Alzheimer's Association was not as broad-based as those of Chaplan or Olson and Strawn.

REVIEW OF THE GREEN-FIELD LIST

The list of subject headings and subheadings furnished by the library consisted of a computer printout of the topical subject fields (tags 65x) in the MARC records that formed the catalog. It consisted of approximately 3,600 different terms, including *MeSH, LCSH,* locally produced headings, and headings of unknown origin. A few

headings were proper names. There had been no apparent preference given to either *MeSH* or *LCSH* headings, and there had been no attempt to resolve any inconsistencies between them. Some of the library's resources need medical subject headings to provide for diagnosis and treatment, while others require non-medical terms to describe the social aspects of the disease. We decided to review the current list and derive from it a unique subject heading list using *MeSH* terms, supplemented by *LCSH* and local terms only when *MeSH* headings proved inadequate to describe the content of a given resource.

Non-English terms in the catalog were also listed using their English equivalents, and were therefore left to stand. Proper names were not searched in the national authority file, as we assumed that their form had already been established. We suggest adding this step when similar projects are undertaken.

A single reviewer determined whether each term in the subject list was derived from *MeSH, LCSH,* or was a locally produced term. The reviewer first checked the subject terms in the 1996 edition of *MeSH.* If the term was located there, the checking process was terminated. If the term was not found in *MeSH,* however, the reviewer searched for the *LCSH* term in ILLINET, the online union catalog of the research institutions of Illinois. If this search did not locate the term, the reviewer continued the search in the 1996 edition of *LCSH.* Exact matches of terms found in *MeSH* were retained; those found in *LCSH* were retained only if a suitable *MeSH* term was not available.

Terms that were near-matches revealed four types of problems:
1. Instances of subject headings with similar meaning that differed in: (a) word order or phrasing; (b) punctuation or capitalization; or (c) spelling (figure 1)
2. Instances where several headings were used when one would have been sufficient (figure 2)
3. Instances where there were long strings of main headings (usually in *LCSH* form) with subheadings ap-

pended in various ways (figure 3), and
4. Instances where terms differed somewhat in meaning (figure 4)

PRODUCTION OF THE NEW SUBJECT LIST

After each term in the existing subject heading list had been reviewed, a new subject list was produced. Word order for all headings was standardized according to *MeSH* style and variants were all cross-referenced. Spelling differences were resolved and appropriately cross-referenced. Because *MeSH* headings were given preference over *LCSH*, the existing *LCSH* or local terms found in the list taken from the catalog records were all referenced so that the validation program of the catalog maintenance system could change the 65x fields automatically. The word order of *MeSH* was the standard, and all variants were cross-referenced. We went beyond the "use for" or "see" reference structure found in either *MeSH* or *LCSH* and considered any heading already constructed to be a valid access point even if it were no longer an authorized heading (figure 5). Previously used *MeSH* terms that were found in the MARC bibliographic records were also cross-referenced so that the validation program would change those records during an update of the file or provide the searcher with clues and links between records (figure 6). Possibly-related terms that were noticed in *MeSH* and *LCSH* were inserted as "see also" references to related topics (figure 7).

MeSH proved quite capable of handling specific medical topics and *LCSH* handled most concepts related to the social aspects of disease, but they did not provide equivalent terms. For example, in problem area number 4, above, *MeSH* has a "Chronic Disease" heading for a specific disease syndrome, while *LCSH* has a "Chronically ill" heading applicable to a set of persons. Though the terms are surely related, they are not interchangeable, so both were kept in the new list, with cross references in both directions: "Chronic Disease – see also Chronically ill" and "Chronically ill – see also Chronic Disease."

```
(a) Word Order or Phrasing
    Adaptation, Psychological (MeSH)
    Adjustment (Psychology) (LCSH)

    Age factors in disease (LCSH)
    Age Factors (MeSH)

    Dementia, Presenile (MeSH)
    Presenile dementia (LCSH)

    American Indians (unknown source)
    Indians of North America (LCSH)

(b) Punctuation or Capitalization
    Memory disorders in old age (LCSH)
    Memory Disorders--in old age (unknown source)

    Down Syndrome (MeSH)
    Down's syndrome (LCSH)

    Language disorders in old age (LCSH)
    Language disorders--in old age (unknown source)

    Terminal care (LCSH)
    Terminal Care (MeSH)

(c) Spelling
    Data Base Management Systems (MeSH)
    Database management (LCSH)

    After care (unknown source)
    Aftercare (MeSH)

    Language disorders in old age (LCSH)
    Language disorders--in old age (unknown source)

    Long-term care
    Long term care
```

Figure 1. Terms Similar in Meaning but Differing in: (a) Word Order or Phrasing; (b) Punctua-

```
Record 1:
    Aged--Housing (unknown source)
    Housing for the Elderly (MeSH)
    Homes for the aged (MeSH)
    Old age homes (LCSH)

Record 2:
    Hospice Care (MeSH)
    Hospices (Terminal care) (LCSH)

Record 3.
    Parent and child (LCSH)
    Parent-Child Relations (MeSH)
    Parent-Child Relationships (unknown source)
```

Figure 2. Instances Where Several Headings Were Used When One Would Have Been Sufficient.

Headings were properly capitalized. For *MeSH* headings all significant words in the headings are capitalized; for *LCSH*, only the first word in the heading. Local headings, comprising mostly names of persons or institutions, were retained. Foreign language headings were retained.

Inexplicably, neither *MeSH* nor *LCSH* authorizes the term "care" of the person, yet professionals and lay people alike

```
Aged, Physically handicapped--Care--United States
Aged--Dwellings--Public Opinion
Alzheimer's Disease--Patients--Care--Handbooks, manuals, etc.
Cerebrovascular disease--Patients--Care--Handbooks, manuals, etc.
Charitable uses, trusts, and foundations--United States--Political
  activity
Consumer protection--United States--Information
  services--Directories
Dementia--Patients--Home care
Portable data bases--Catalogs--Periodicals
Rural aged--United States--Economic conditions--Statistics
Sick children--Respite care--United States--States
```

Figure 3. Instances with Long Strings of Main Headings and Subdivisions in Variant Order.

```
Chronic Disease (MeSH)
Chronically ill (LCSH)
```

Figure 4. Instances Where Terms Differed Somewhat in Meaning.

```
Dementia, Presenile (MeSH)
Presenile dementia (LCSH)--see Dementia, Presenile

Accidental Falls (MeSH)
Falls (Accidental) (LCSH)--see Accidental Falls

Elder Abuse (MeSH)
Abused aged (LCSH)--see Elder Abuse
Aged--Abuse of (LCSH)--see Elder Abuse

Homes for the Aged (MeSH)
Housing for the Elderly (MeSH)--see Homes for the Aged
Aged--Housing (unknown source)--see Homes for the Aged
Old age homes (LCSH)--see Homes for the Aged
```

Figure 5. Examples of Older Headings with Variant Word Order, or Spelling Variations Used as References.

make heavy use of the term, sometimes as part of a combination of terms such as "respite care," " caregivers," and "long-term care." Such terms cover the set of behaviors involving physical and emotional support of patients. We found 75 different headings for the concept "long term care." This was caused by the variation in spelling between *LCSH* and *MeSH* (see figure 1), and by the variations in subdivisions and subheadings used by the two lists. Such an important concept was thus scattered in the list; we normalized these headings by following the procedures described here. We also added the access term "care" to the list as a local term and provided cross references to related terms in the list.

A NEW LIST IS FORMED

In place of the old subject list of 3,600 headings, the new subject list presented 1,220 main headings; of these, 775 (64%) were *MeSH*, 285 (23%) were *LCSH*, and 160 (13%) were local terms or proper names. In addition, 125 "see" and "see also" references were added. Lists of approved *MeSH* topical subheadings (with allowable *MeSH* Tree Structure categories), form subheadings, and geographic subheadings were provided in an appendix to the list.

RESULTS

Because neither author was or is on the staff of the Green-Field Library, this was

```
Tranquilizing Agents, Major--see Antipsychotic Agents

Tranquilizing Agents, Minor--see Anti-Anxiety Agents
```

Figure 6. Previously Used *MeSH* Terms as References.

```
Medicare (MeSH)
--see also Medical Indigency (MeSH)
--see also Social Security (MeSH)


Minority groups (LCSH)
--see also Blacks (MeSH)
--see also Hispanic Americans (MeSH)
--see also Pacific Islander Americans (MeSH)
```

Figure 7. "See Also" References for Possibly Related Terms.

a limited project. We hope that our work will form the basis for an ongoing maintenance of the subject list. When we had to stop work on the project, we felt we had provided the Alzheimer's Association with a better tool to:

1. revise the existing subject headings in their catalog,
2. select appropriate headings as new resources are being cataloged,
3. consult a syndetic structure whenever their catalog had an authority/verification system and a search system that provided automatic links between old access terms, references, and established headings, and
4. showcase the holdings of the library on the Web with some consistency in subject access.

As part of a continuous subject access improvement project, we would suggest relating class numbers and subject headings. The captions for class numbers provide the broad term access so often missing in subject heading lists. In small special libraries the size of the Green-Field Library, this work could form the basis for special signage near the open shelving and on the Web page for the library, as well as cross references in the catalog file itself.

WORKS CITED

Chaplan, Margaret A. 1995. Mapping Laborline Thesaurus terms to *Library of Congress Subject Headings*: Implications for vocabulary switching. *Library quarterly* 65, no. 1: 39–61.

Olson, Tony and Gary Strawn. 1997. Mapping the *LCSH* and *MeSH* systems. *Information technology & libraries* 16 (March): 5–19.

Weintraub, Tamara S. The dual-thesaurus catalog: *MeSH* and *LCSH*. *Technicalities* 12, no. 12: 10–12.