Letters to the Editor

February 28, 2006

Certain opinions about library catalogs and their rules in Joan Aliprand's article, "Scripts, Languages, and Authority Control" (*LRTS* 49, no. 4 [Oct. 2005]: 243-49) differ so much from mine that I must try to refute them lest they mislead the authors of the new edition of AACR.

I begin with some quotations from her article. "The source of authority that defines headings has an explicit or an implied language, and may include instructions on how to coordinate text in other languages—some written in other scripts—with the preferred language of the source of authority" (244). "Locale specific access points include names, uniform titles, subjects and additional parts of headings in the language of the catalog" (244). "Sources of authority that determine the content of a particular headings include cataloging rules, subject heading lists, and thesauri" (244–45). "The basic need of a user is the ability to search the catalog using his or her preferred language and that language must be written in the proper script" (245). "To establish headings in a specific language and script, one must choose a source of authority [i.e., cataloging rules] whose operative language—either explicitly or implied—is the language of the catalog user" (245). "Conceptually a separate logical catalog exists for each language, with separate authority files for names and subjects. In reality, these may be amalgamated into a single physical catalog, with repetitive data being eliminated. In thinking about multilingual and multiscript access, the catalog must be considered conceptually, that is, as a logical model, not as it is in actuality. The *Anglo-American Cataloging Rules,* 2nd ed. (*AACR2*) and Library of Congress subject headings (LCSH) are accepted so much as basic tools that catalogers may forget that the catalog they are building is for English speakers" (245).

I omit topical subject headings to focus on the cataloging rules. (Though why our rules don't include rules for subject headings, as Cutter's rules and the Vatican rules did, is beyond me.) I submit that *AACR2* is intended for cataloging Anglo-American libraries whose collections contain items in many languages and scripts. Her logical model is so contrary to reality that anything built upon it would have a very fragile foundation. Part one of *AACR2* defines multilingual and multiscript descriptions; see rule 1.0E1, "In the following areas give information transcribed from the item itself in the language and script (whenever practicable) in which it appears there: Title and statement of responsibility, Edition, Publication, distribution, etc., Series." Part two defines multilingual but monoscript access points for persons, geographic names, corporate bodies, and uniform titles. These access points must be in the roman script, but most are in their original language, not English. Geographic names are given in English if in general use (rule 23.2). Except for persons entered under given name (that is, forename), personal names are given in their original language, or a romanization of it (rules 22.3B3 and 22.3C). Most personal names are language neutral; for example, the heading, "Verne, Jules, [dates]" is fully adequate for editions of his works in languages using roman script. Only the titles of royalty, their consorts, children and grandchildren, and persons in religion are given in English (rule 22.16). The names and titles of nobility are given in the original language (rule 22.6), "Follow the proper name in the title by the personal name (excluding unused forenames) in direct order and the term or rank in the vernacular." More than three pages are devoted to treatment of surnames, with separately written prefixes in various languages (rule 22.5D). A corporate body is entered under

the name by which it is commonly identified; for example, "Ecole centrale lyonnaise," not under an English translation (rule 24.1). Uniform titles for works created after 1500 use the title or form of title in the original language (rule 25.3). In *AACR2,* appendix A (pages 25–37) gives rules for capitalization of various foreign languages, including Bulgarian and Russian in Cyrillic script. For brevity I simplify things a bit, but the fact remains that except for geographic names, most name headings are in the original language, or romanizations of it when they are in another script; in short, *AACR2* explicitly defines a multilingual catalog.

She also writes, "Canada, for example, has two official languages, English and French. Conceptually, authority control in Canada serves two locales: English speaking users and French speaking users" (245). This is true but somewhat misleading. In catalogs based on the English or the French version of AACR2, corporate bodies with German, Spanish, Italian, and so on language names have their headings in their respective languages. The same is true for persons entered under surname.

I heartily agree with her that "romanization is not good enough!" It is a disservice to those seeking nonroman script materials, but the concept of separate catalogs and separate rules for every language is not the path to a better catalog.

Since the 1980s, many catalogers of materials in some languages using nonroman scripts (Japanese, Arabic, Chinese, Korean, Persian, Hebrew, Yiddish [JACKPHY]) have assigned nonroman script access points so readers can find the authors and titles they want without first guessing how catalogers have transliterated their names into roman script. They have done so without any guidance from the cataloging rules. Needed are cataloging rules to allow nonroman access points and to provide authority control for them. At present, these JACKPHY languages nonroman access points lack effective automated authority control. But they are infinitely more helpful to those seeking nonroman resources than the romanized access points alone on which those seeking resources in languages using other nonroman scripts must rely.

This brings us to authority records as the means to authority control of access points. Again I begin with several quotations. "The limitation of one established form per record is imposed by the fact that in the *MARC 21 Format of Authority Data,* certain key data elements are singly occurring. Field 008 *Fixed-length data elements,* which is not repeatable, includes two positions that are each only one character long" (page 246). Were the cataloging rules expanded to allow nonroman script access points there would be no problem identifying the rules used in the MARC authority format field 008, fixed length data elements—the single character position would suffice. For authority control of these headings, I propose that one authority record be defined for each *entity* (person, corporate body, and so

on), with an authorized form in as many scripts and languages as that entity has on title pages of items in a library's collection. As noted before, most modern personal names are language neutral—but none are script neutral, so one authorized per script would usually suffice. (From an earlier article of hers I am aware that some authors who write in Hebrew and Yiddish [both in Hebrew script] use different names for each.) An authorized heading would have references from variants in the same script or languages using that script. Early in her article she writes, "Unlike bibliographic records, where nonroman data can just be ignored or discarded when it cannot be handled, all parties working with synchronized authority files have to be able to see authority records in their entirety" (243). Though it would require some changes in software—possibly using the field link and sequence number, subfield $8, with script or language codes to cluster authorized headings and their references in the same language or script—the method I advocate would better enable a cataloger to see all of the headings and references for an entity than separate linked records for different scripts. Making 1XX fields and the linkage subfield, $6 repeatable, might also be part of the solution.

"When, in the future, all systems contributing to NAF have multiscript support . . ." (247). Whether the requirements for synchronized authority data justify the delay in improving local catalogs until all who participate in the synchronized authority file can see all scripts is not for me to say. It would seem to involve a cost-benefit analysis of cooperative cataloging.

She continues, "The benefit of precluding a single composite authority record with multiple syndetic structures is that such a record would be complicated, and difficult to process and update" (246). Because (as quoted above) it is important "to see authority records in their entirety, including data in nonroman scripts," dividing it into multiple records will make seeing the whole picture much more complicated for catalogers. If a choice must be made between more simplicity for computers or catalogers, simplicity for catalogers seems preferable. And multiple authority records for an entity may be more complex for the computer, too: "The other option is to have multiple authority files (for example, one for each language) and to link them all together to provide multilingual and multiscript access" (247). Simultaneous display of authority records from separate, linked files would be even more complicated. Attempting to maintain synchronization of updates of such files across multiple networks would be fraught with errors. A single, sometimes multiscript authority record with everything pertaining to an entity is preferable to either separate linked authority records in one file or separate linked records in different authority files. This is because a single record requires less synchronization and is thus less prone to corruption. Note that were this suggestion adopted, author-

ity records would be more complex due to the use of the already defined $8 control subfield only when an entity has names in more than one script or, more rarely, language.

I believe her statement "Including links to other established headings in the record is also possible. . . ." (247) must mean: Including links in the record to established headings *in other records* is also possible. The reference in the next paragraph to "the last line in figure 1" should refer to figure 2.

It is not possible to define all the intricacies of multi-script authority control in even a long letter. An insoluble part of the problem is that a cataloger can determine the language of the item in hand to be cataloged, but not the language of every future catalog user. Basing the catalog's definition on the notion that our cataloging rules are for English users only and concept of language-specific rules for language-specific catalogs will not bring about access equity for those seeking nonroman library resources.—*James E. Agenbroad, retired systems analyst, Library of Congress*

April 12, 2006
The purpose of my paper, "Scripts, Languages, and Authority Control," and my original presentation in 2003, was to shift the focus of discussion on the use of nonroman scripts in library data from the scripts themselves to languages.

In October 2005, the ALCTS Executive Committee established the Task Force on Non-English Access to address "access to library resources in all languages and scripts" (www.ala.org/ala/alcts/alctspubs/alctsnewsletter/vol16no5/announcements/TFnonenglish/tfnoneng.htm). That is, the ALCTS Executive Committee recognizes that a key issue for ALA is providing library service to people who use a language other than English, whether the language is written in a nonroman script or not.

In his letter to the editor, Mr. Agenbroad covers two topics: "English" headings, and his view on how an authority record should be structured.

Using foreign language names or terms in an English language environment is normal practice—just read any newspaper written in English! Similarly, headings established according to *AACR2* for a catalog to be used by people who read English may include, or consist entirely of, names or terms from a foreign language (sometimes converted to Latin script by romanization). So in an English language environment whether the words are written or spoken, borrowings from foreign languages may occur.

How are foreign names or terms occurring in headings formulated according to *AACR2* incorporated into a catalog? In an English-language environment, the ALA Filing Rules is the usual standard for ordering bibliographic records. The rules do provide for separate ordering of text in "nonroman alphabets," but, as Mr. Agenbroad noted (in the third paragraph of his letter), access points defined by *AACR2* "must be in the roman script" (that is, Latin script).

Foreign language text written in Latin script is not given special treatment under the ALA Filing Rules. What would be significant differences in the language of origin are either ignored or converted to English alphabet equivalents, so that foreign language text can be forced into the A–Z order of English. But, of course, this is the order expected by people who read English, so this cavalier treatment of foreign language text is understandable in what I called an "English-speaking locale."

The second part of Mr. Agenbroad's letter (beginning with paragraph 6) takes issue with the model for MARC 21 authority records presented in my paper, and describes an alternative with the following features:

> one authority record for each entity (person, corporate body, etc.) with an authorized form in as many scripts and languages as the entity has on titles pages of items in a library's collection.
> An authorized heading would have references from variants in the same script or languages using that script.

The MARC 21 Format for Authority Data specifies that the data elements identifying the descriptive cataloging rules (008/10) and subject headings/thesaurus (008/11) used to formulate the 1XX heading are singly occurring. To accommodate the multiple authorized forms that his model contains, Mr. Agenbroad proposes: "Were the cataloging rules expanded to allow nonroman script access points there would be no problem identifying the rules used in the MARC authority format field 008, fixed length data elements the single character position would suffice."

The record model proposed by Mr. Agenbroad will not work unless the cataloging rules are expanded to allow nonroman access points. Whether this will happen is questionable, given the Joint Steering Committee for Revision of AACR: RDA's (JSC) current focus on resource description and access (RDA) as well as the history of a previous attempt. In 2000, ALA submitted a proposal to JSC: to modify AACR2 to provide an option for assigning nonroman access points. The ALA proposal had its origin in a proposal submitted to the Committee on Cataloging: Description and Access (CC:DA) by Mr. Agenbroad in 1999. The "Outcomes of the Meeting of the Joint Steering Committee Held in Washington, D.C., USA, 2–4 April 2001" (http://www.collectionscanada.ca/jsc/0104out.html) report that the ALA proposal "was greeted with largely negative responses." In 2001, ALA withdrew its proposal.

There are many other problems with the record model. For example, it is unclear whether the proposed alternative authorized forms are language- or script-based. Mr. Agenbroad writes: "An authorized heading would have references from variants in the same script or languages using that