Issues in Retrospective Conversion for a Small Special Collection:
A Case Study

by *Fern Hieb*

A small special collection presents a unique problem for the task of retrospective conversion of the catalog to machine-readable form. Unless associated with a college or university library, a small collection frequently does not have a professional librarian on the staff who can formulate a plan for retrospective conversion (recon) and answer related questions.

This paper will identify and explore issues from the viewpoint of a small special collection planning for recon. The Moravian Music Foundation is used as a case study, but most of the questions and recommendations would apply to any special collection.

The Moravian Music Foundation was established in 1956 as a repository for music of the Moravian Church. The archives of the foundation contain music composed or used by Moravians in early America: tune books; band books from the Civil War; hand-copied works of Haydn, Bach, and Mozart; and works by Moravian composers (Charles G. Vardell, Johannes Herbst, Johann Friedrich Peter, and others). The archival holdings are housed in two locations: Winston-Salem, North Carolina and Bethlehem, Pennsylvania. In addition to these archival holdings, the foundation maintains the Peter Memorial Library in Winston-Salem, a reference collection for the use of visiting scholars. The foundation's total holdings are as follows:

Peter Memorial Library	6,000 items
Winston-Salem Archives	4,120 items
Bethlehem Archives	4,900 items
Lowens Tune Book Collection	1,175 items
	16,195 items[1]

Why Automate the Catalog?

There are several basic advantages in having the catalog automated. First, it becomes more easily and quickly accessible from any location. Retrospective conversion will integrate all of the bibliographic records of the Moravian Music Foundation into a single online catalog, including holdings in both Bethlehem and Winston-Salem. The major benefit to the Foundation will be improved access by scholars from remote locations. A researcher working in Boston, for example, will have access to the Foundation's complete catalog from home or office, using a personal computer and a modem.

A second reason for having the catalog online is for ease of updating. When new materials are accessioned or when corrections are made to bibliographic records, updating is easily done online. Other benefits of automation include more manageable inventory control, easier collection development, and bibliographic files that are better preserved and more secure.[2] An online catalog also can support such applications as serials control and circulation.

What Are the Options for the Actual Conversion Process?

Without a professional librarian on the staff to manage the retrospective conversion, two major options are available. The first is to hire temporary staff to manage the operation; the second is to outsource the project. The two alternatives are compared below with regard to cost, staffing, time involved, and desired quality of the converted records.[3]

The first conversion option is to do the project in-house with temporary staff hired for the project, deriving records from OCLC or from a database on CD-ROM such as *The Music Catalog* on CD-ROM from the Library of Congress.[4] The greatest advantage to hiring temporary project staff to work in-house is a much higher level of quality in the converted records.[5] If questions arise, project staff can pull the item from the shelf to check bibliographic data. Using this option, permanent staff are available to answer questions, but will be only minimally involved in day-to-day project work. The timetable generally is set by the duration of funding, but is likely to be more flexible with the temporary staff on site. On the negative side, costs will be high, since project staff have to be hired, trained, and supervised. Data access, project terminals, network charges, and office equipment will all add to the cost.

A second option is to outsource the project to an outside vendor such as OCLC (Online Computer Library Center, Inc.).[6] Shelflist cards are sent to the commercial vendor, who does the conversion process at a remote site. An advantage of this option is that the vendor provides staff with considerable experience in retrospective conversion projects, so

little time is lost in training. Also, the timetable and cost of the project are specified in the contract. Quality control however, is the main reason *not* to go with a commercial vendor. As experienced as a commercial vendor might be, record quality is usually compromised because of time constraints, lack of familiarity with details about the collection, inability to look at the actual item when the shelflist card is lacking information, and inability to ask questions of regular staff on a day-to-day basis.

How Can the Conversion Effort Be Quantified?

A major aspect of planning for either conversion option is determining how many of the library's bibliographic records can be copied from records already in the OCLC database. At the Foundation, a random sample of 100 cards was pulled from the Peter Memorial Library and the archives to check the hit rate on OCLC. The 50 cards pulled from the Peter Memorial Library yielded 34 hits, meaning that approximately 68% of the records would be derivable from existing OCLC records. The hit rate might in reality be higher, considering that the Peter Memorial Library consists largely of standard reference materials, but the cataloging on the shelflist cards was often too minimal to confirm a match with an OCLC record. If the information on the cards is inadequate, the item itself must be pulled from the shelf for comparison with OCLC records. For the records that do not have a match on OCLC, the librarian will have to create original records. If the conversion project is sent to a commercial vendor, only the definite hits will be converted. The remaining cards will be returned for local staff to convert after questions are resolved.

Not surprisingly, the archives yielded a very low hit rate since the collection is largely manuscripts. An off-site commercial vendor would be unable to convert much of an archival collection.

What Will Retrospective Conversion Cost?

Recon is a costly venture, whether done in-house with temporary staff or outsourced to a commercial vendor. In general, the lower the hit rate on OCLC, the higher the cost, since the records that are definite "no-hits" will require original cataloging. But it is difficult to determine definite hits when a shelflist card has minimal or questionable information.

In the title area, for example, it is sometimes unclear whether the title was transcribed from the item or was synthesized by the cataloger.[7] Were titles translated into English by the cataloger or transcribed as they appeared on the item? Were any title words moved around, added, or deleted? If the title came from the cover, was that indicated? These questions make it difficult to determine from the shelflist card whether or not it matches the OCLC record.

In the publication area, the date of publication often appears in brackets on the shelflist card, meaning that the date was not on the item, but was supplied by the cataloger. If the record that appears to be a match on OCLC does not have brackets around the publication date, then the item will have to be checked to see whether this item really is a match with the OCLC record.

As to physical description, many old shelflist cards lack pagination. Even when pagination is given, it often differs by a few pages from what appears to be a match on OCLC.

Subject headings and added entries are critical access points in bibliographic records, but these often are lacking on old shelflist cards. The cards that *do* provide subject headings and name-added entries have to be checked against authority files for accuracy.

Another major issue is the staff time required per record conversion. The author has had personal involvement in estimating the proposed recon project for manuscripts on microfilm at the Isham Memorial Library at Harvard University. In that proposal, a total of 38.5 minutes was allocated for each hit on OCLC (22.25 minutes for students who search the database and do data entry, and library assistants who edit the records and do authority work, and 16.25 minutes for a professional cataloger who revises the records). For records not found on OCLC, a time of 50.5 minutes was projected for converting each record, since the record would have to be created, not just revised.[8] Harvard's projected cost in 1994/95 for that project was $21/record.

If the Moravian Music Foundation chose to hire a temporary staff to do the project in-house, students from nearby colleges could be hired for such tasks as searching and data entry. With a full-time librarian as supervisor and general manager, the project could be completed in two or three years. Printed catalogs of parts of the collection would be excellent resources (i.e., Frances Cumnock's *Catalog of the Salem Congregation Music*[9] and Marilyn Gombosi's *Catalog of the Johannes Herbst Collection.*)[10] Another time-saver might be the downloading of Foundation holdings that are already a part of the database *Repertoire International des Sources Musicales* (RISM).[11] Online cross references to these catalogs would aid the researcher.

If the Foundation outsourced the recon project to OCLC, the costs would be based on such factors as the estimated number of hits, the type of material (scores, books), the language of the material, the type and number of special local requirements, the percentage of shelflist cards containing an LC card number or an OCLC number, and other factors related to editing.

If a small collection chooses to hire an outside vendor to convert the catalog, the shelflist cards are mailed to the vendor. If hits cannot be verified, those cards are marked as problems (exceptions) to be resolved at the local level. Vendors' charges are calculated on searches, not hits, so if the vendor has to conduct extra searches because of minimal shelflist information, the cost will increase.

An analysis of the options for the Moravian Music Foundation indicated that hiring an in-house staff was the preferred approach. The decision was based on two major factors: 1) an overriding concern for good quality records, and 2) the large percentage of archival holdings that will not have copy on OCLC.

How Is In-house Retrospective Conversion Actually Done?

Generally, shelflist cards are used as the source of cataloging information, rather than catalog cards. Shelflist cards are more likely to provide subject headings and added entries. Also, patrons are less likely to be inconvenienced since they use the catalog instead of the shelflist.

Student workers and/or library assistants pull shelflist cards, search for hits on OCLC, and then edit the derived records into machine-readable form. Editing a derived record involves adding local information and updating access points to bring the record into conformity with national bibliographic standards. Producing original records has to be done by a professional librarian who has a broad understanding of cataloging rules.

Project staff need access to basic cataloging support materials. These include the *Anglo-American Cataloguing Rules* (*AACR2R*),[12] OCLC's *Bibliographic Formats and Standards*,[13] LC's *Subject Cataloging Manual* (*Shelflisting*),[14] LC's *Descriptive*

Cataloging of Rare Books,[15] and Hensen's *Archives, Personal Papers, and Manuscripts*.[16]

In the case of the Moravian Music Foundation, the Peter Memorial Library probably will be converted first because more hits will be found on OCLC from that collection than from the archival collections. This will launch the recon effort with initial success and speed.

Unique Aspect of Automating the Foundation's Music Archive

A single text often was set to different tunes by various Moravian composers. The musical incipit (first few notes of a tune) is often the only way to differentiate among various settings of the same text. The Foundation currently relies heavily on the musical incipits found on the back of many catalog cards. The recon effort needs to include a system for coding these musical incipits. The appendix of this paper describes the method recommended for the Moravian Music Foundation, based on Barry Brook's *A Plaine and Easie Code System for Musicke*.[17]

What Are the National Standards for Online Records?

The USMARC[18] standard for online records varies somewhat according to the format of the item being cataloged. Formats for scores, books, serials, etc., are all integrated into the *Bibliographic Formats and Standards* by OCLC. The standard for library cataloging is the *Anglo-American Cataloguing Rules* (AACR2R) and, for archival description, *Archives, Personal Papers, and Manuscripts* (APPM).

The standard for authority records is the LC authority file on OCLC. Since authority control provides standardization of access points throughout the database, every access point should be checked against the authority file. These access points include proper names, titles of works, subject headings, and added entries. Published musical works often are found in different manifestations: as a score, a score and parts, parts alone, a vocal score, an arrangement, etc. Authority control will be critical for collocating these different manifestations of the same work under a uniform title. A music librarian will be needed to understand the nature of the music itself and to recognize names, titles, and musical forms for authority work.

Although *AACR2R* is slanted heavily toward data found in published books, the chapters describing formulation of names for persons, geographic places, corporate bodies, uniform titles, and cross references will be of value to archival catalogers.[19]

What Is the Recommended Library Automation System?

Once the organization has chosen the recon strategy, then it must select an automated system. It is essential to choose a library automation vendor that has experience and stability. The vendor should be committed to providing support and to enhancing the software.

When choosing an automated library system, the following factors should be considered: functionality, price, ease of use, workflow requirements, standards, training, performance and reliability, and expandability.

Library automation vendors tend to sell to specific niches of the market. Innovative Interfaces Inc., for example, is the premier automated library system focusing on large U.S. academic libraries. Ameritech and DRA generally sell to smaller academic institutions.[20]

What Are the Options for Remote Access?

The automated catalog can be made available to patrons in several ways. One possibility is to have a personal computer with a modem at the site of the collection. Bulletin board system software makes it possible for patrons to dial into the personal computer. Unfortunately only one patron, or a very limited number of patrons, can access the data simultaneously.

Another possibility is to distribute the catalog in CD-ROM format. This works best if the collection accessions only a limited number of items each year. Tapes of the database can be printed onto CD-ROM disks for distribution to interested patrons.

Perhaps the best approach for remote access, but also the most expensive, is the Internet. A large number of patrons can access the catalog simultaneously, and online updating is visible as soon as it is completed. The library purchases an account through an Internet access provider.

Conclusion

The decision to undertake a retrospective conversion project involves a number of complex decisions. One of the most critical first decisions concerns desired quality of the finished records. A project staff on site will be able to consult the collection or the permanent staff when questions arise. A librarian as project manager will ensure that standards are followed. The resulting high quality of completed records is the most important aspect of the project.

The decisions about remote access and choice of automation software also are critical to the success of the retrospective conversion project. Retrospective conversion merits a substantial investment of time and money because it will ultimately be a major determinant of the collection's usefulness and accessibility.

Status of the Moravian Music Foundation Project

At the time of this writing, the Moravian Music Foundation is projecting that retrospective conversion of its catalog will begin in 1998. Currently, a new building is under construction that jointly will house the Southern province offices and the Foundation's library and archives. The new building will be equipped with computer hardware to support an automated catalog and remote access.

Appendix

Several systems exist for the coding of musical incipits into a uniform typewriter code. Since the RISM project already uses Barry Brook's *A Plaine and Easie Code System for Musicke*, it is recommended that the Foundation use the same system. There are several reasons why this system would suit the Foundation's needs:

1) it is simple and accurate as to pitch and rhythm;
2) it is closely related mnemonically to musical notation;
3) it requires only a single line of typewriter characters;
4) it is usable by anyone with some musical training;
5) it is easily recognizable as music from the symbols alone;
6) it is applicable to all western music;
7) it is universally understandable and internationally acceptable.

Pitches are indicated by capital letters, rests by a dash, time values of notes and rests by numbers, and other qualifying terms by symbols and lower case letters. Precise location of pitch is accomplished with a minimum of octave symbols (commas and apostrophes).

References

[1] Claypool, Richard D. "Archival Collections of the Moravian Music Foundation and Some Notes on the Philharmonic Society of Bethlehem," *Fontes artis musicae* 23, 4 (1978): 180.

[2] Jutta Reed-Scott, *Issues in Retrospective Conversion: Report of a Study Conducted for the Council on Library Resources* (Washington, D.C.: Bibliographic Service Development Program Council on Library Resources, Inc., May 1984): 4.

[3] Reed-Scott, 15.

[4] *Music Catalog on CD-ROM* (Washington, D.C.: Library of Congress, Cataloging Distribution Service, 1994-)

[5] Reed-Scott, 21.

[6] OCLC is a major bibliographic utility. The regional membership cooperative for OCLC is SOLINET (Southeastern Library Network, Inc.), 1438 West Peachtree Street, N.W., Suite 200, Atlanta, GA 30309-2955. Tel. 1-800-999-8558. Fax no. (404) 892-7879.

[7] Sue Weiland, "Music Scores: Retroconversion or Recataloging?" *Technical Services Quarterly* 10, 1 (1992): 67.

[8] Grant proposal for 1994-1995 to convert records for 4700 manuscripts of musical sources and 2200 rare early printed and manuscript treatises on music at the Isham Memorial Library, Harvard University.

[9] Frances Cumnock, ed. *Catalog of the Salem Congregation Music* (Chapel Hill: University of North Carolina Press, 1980).

[10] Marilyn Gombosi, comp. *Catalog of the Johannes Herbst Collection* (Chapel Hill: University of North Carolina Press, 1970).

[11] The RISM-U.S. Music Manuscripts Database is copyrighted by the Joint Committee on RISM of the American Musicological Society and the Music Library Association. It is part of *Repertoire International des Sources Musicales* (Kassel: Barenreiter, 1971-). To contact the U.S. RISM Office, send e-mail to RISMHELP@RISM.HARVARD.EDU.

[12] *Anglo-American Cataloguing Rules* (Ottawa: Canadian Library Association; Chicago: American Library Association, 1988).

[13] OCLC, *Bibliographic Formats and Standards* (Dublin, Ohio: OCLC, 1993-).

[14] Library of Congress, Cataloging Policy and Support Office. *Subject Cataloging Manual. Shelflisting* (Washington, D.C.: The Office, 1995).

[15] *Descriptive Cataloging of Rare Books* (Washington, D.C.: Cataloging Distribution Service, Library of Congress, 1991).

[16] Steven L. Hensen, comp. *Archives, Personal Papers, and Manuscripts: A Cataloging Manual for Archival Repositories, Historical Societies, and Manuscript Libraries* (Chicago: Society of American Archivists, 1989).

[17] Barry S. Brook, and Murray Gould. "Notating Music with Ordinary Typewriter Characters," *Fontes artis musicae* 11 (1964): 143.

[18] USMARC format is the standard for representing and communicating bibliographic and authority records in machine-readable form.

[19] Jackie M. Dooley, "An Introduction to Authority Control for Archivists," in *Archives and Authority Control: Proceedings of a Seminar Sponsored by the Smithsonian Institution October 27, 1987* (1989; reprint, Archives and Museum Informatics Technical Report no. 6): 5-18.

[20] Jeff Barry, Jose-Marie Griffiths, Gerald Lundeen. "The Changing Face of Automation," *Library Journal* 120, 6 (April 1, 1995): 44-54.