

MLA-BCC Genre/Form Task Force Business Meeting

MLA, Dallas, Feb. 2012

Submitted by Beth Iseminger, chair

Hierarchy Issues

There are a lot of redundant terms between several of the broader categories, including songs/vocal/sacred music, instrumental/art music, and especially between folk and world music. We need to create scope notes for the broad categories, so that we are working with the same definitions. Do we actually need both the folk and world music categories? If world music means “any non-western music,” it will include a lot of folk music. What would the intellectual, hierarchical distinction be between folk and world music? Is folk music “world music” of North America? “Folk music” also refers to the pop music genre ‘folk music’ and to a few instrumental music terms. We also need to make it clear that “Art music” is a comprehensive term for all musical traditions having an ‘art music’ concept. Combining all the hierarchies into one large set will help us see the extent to which terms overlap between categories, which will give us a clue to which top terms are really necessary.

Are the top terms “Instrumental music” and “Vocal music” medium terms, or are they also appropriate as genre terms? Would these terms be helpful for works that have no other genre term, or do they duplicate other categories? We may be assuming the existence of a layer in the hierarchy that we don’t actually need. Perhaps instead of “Instrumental music” as a broader term for genres like symphonies or sonatas, the next higher term could be “Art music” or the node label “[Forms].”

Will users assume if they see “Instrumental music” in the genre thesaurus that they can find “Piano music” in the same thesaurus? This is a case where user education will be important. We will need to make sure users understand when to search under medium. (For example, a user looking for a CD of vocal music would look under “Vocal music” in the medium hierarchy.)

The new medium thesaurus will need to be user friendly, and systems need to use the hierarchical design in their search mechanisms. Systems currently do not utilize hierarchies, and that’s our basic problem. Eventually, it should be relatively simple to search for specific mediums. What workarounds will be needed to find broader mediums like “all instrumental music?” Systems should be able to retrieve “nothing in the vocal hierarchy and everything in the instrumental hierarchy.” Would it be easier to have a system designed with many algorithms to do this search, or easier to have a genre term (like “Instrumental music”) that is a broader term and subsumes many well-known instrumental genres? We also have to remember that many genres include both instrumental and vocal versions.

Non-Music Facets

Hermine reported on the work that the ALA-SAC Subcommittee on Genre/Form Implementation is doing with facets that are related to genre. These include geography, ethnicity, language, chronology, category of persons, and audience.

The group is working on defining audience versus characteristics of creator/contributor. Current LCSH headings treat these facets in many different ways. One potential solution is to create two new MARC 3XX fields for characteristics of creator/contributor and for audience. Vocabulary should be in the form of plural nouns and could come from LCSH or a new thesaurus. The group is also suggesting potential MARC fields for chronology (648), language (377) and nationality (752 or 370).

The MLA genre task force tends to think this approach results in too much fragmentation. For example, does it really make sense to take all religion and ethnicity terms apart? In our discussion of Chinese opera, we realized it was incorrect to dissect the term in to “China” and “Opera”, since Chinese opera is a genre unto itself.

While it is very good to have these facets, the effectiveness starts to break down when we start dividing so many facets into separate MARC fields. Rather than focusing on MARC, look at the vocabulary itself. It could function very well faceted, but coding it separately in MARC doesn't function as well. We have to make the MARC coding work right now for medium and genre, but in the big picture, we need to look at what the vocabulary needs and make the system do that, not the other way around.

Moving toward post-coordination, the possibilities are great for describing resources with single works very well. In our world of sound recordings that combine multiple different genres and mediums, post-coordination could be a nightmare unless we have work records. For the genre project to be effective at all, it has to push us toward work records. Genre, medium, and other facets should be in work records and personal name records, and in a linked data environment the relevant information will be pulled for search results. The faceted vocabulary makes more sense outside of MARC, and it is one of the things that are finally breaking the MARC format.

Genre Authority Records

The task force agreed that the lists we compiled of new terms (from Garland primarily) should be saved for an ongoing project for another group. We will make the lists available, since they now reside on the task force wiki. The group taking this on will have to address a lot of language issues and de-duplication.

Creation of the genre authority records will likely begin early this year. The task force will work with the SACO-Music Funnel, and we all need to become part of it if we are able within our institutions. Nancy can also work with folks unable to be official SACO members to enable them to contribute records. The SACO project uses the Minaret system, so we would either use it through our institution's Minaret system or work with Nancy. There are only four members of the SACO-Music Funnel currently, and only 2

of them have submitted headings. Patty Falk at Bowling Green has contributed the most headings currently, and she is interested in helping us with the genre authority records.

The next steps we need to take regarding genre authority records are: 1) Combine the hierarchies into one; 2) Agree on the top terms and define them; 3) Examine each term, to see what belongs in the genre authority record, and take out what doesn't belong. Nancy created an example record for concertos. She discovered that the notes in the authority record changes drastically. We may want to add some data, like chronology, geography, and nationality, to records as we create them. This is something to add to our workflow. Some of these facets can be assigned in separate MARC fields, and we should make sure the fields are valid for genre authority records. The 670s from topical authority records can be kept in genre records as long as they are relevant. They tend to define the genres and could function as scope notes. Old records may not have 670s, however. Will we need to add them? What music decides may help other projects as well. In the end, scope notes are the most important things to add.

The hierarchies provide some of the content for the new authority records, in terms of broader/narrower/related terms, cross references, and scope notes. Will the reciprocal scope notes that now exist remain when the LCSH and LCGFT databases are finally separate, or are these notes useful only for distinguishing use while both terms are in one database? (I.e. Use this form for music; use this form for work about the topic.) When proposing a new genre heading currently, proposing the reciprocal subject heading is not required.

There is some concern about losing the 053 class number data when we move to genre authority records. The current class numbers often represent a combination of medium plus genre. 053 numbers could be kept in any remaining LCSH records, or if reference records are created, they could still retain the 053 (though would that be valid?)

Reference records may be helpful, but the same term cannot be in a reference record at the same time it is in an authorized form in the same character string. We will have to keep in mind whether the references needed will conflict with existing headings. This affects the timing of the project and conversion. Reference records would be temporary and would be used to facilitate machine deconstruction. There should be an algorithm to help with deconstructing medium from genre.

There are two options we could consider moving forward:

Option 1: Deconstruct the pre-coordinated genre and medium strings for terms that are simple to do now. Move the genre and medium terms into separate 150s. We're still formulating the question. The work should be done first in LCSH coding, because it is hard for systems to split the headings and also put them in two separate, new fields at the same time.

This is similar to the process that was done for headings for waltzes and polkas, which used to be pre-coordinated with medium. The authority record for "Waltzes" contains

references to all the medium of performance headings that used to be combined with it. We have better technology now than we did then. We'll be trying to programmatically insert fields in bibliographic records, when the algorithm retrieves certain records that qualify to be deconstructed. It will be easier to do this if the genres and mediums are already split. This means we would need to make two passes through the database. We could start this at any time, with existing headings. Medium might need to be cleaned up, but it would be easier to clean up medium terms when they appear in headings in only one form.

Option 2: Create new records from the vocabulary in the hierarchies, rather than starting with existing authority records. The existing authority records could be mapped to the new records after that.

The bottom line is this: We need technical advice. Our next step needs to be talking with people who have experience working with large database projects like this. We need to talk with them about how to make this work. We probably also need to talk with authority vendors and OCLC. Jay Weitz recommended some people at OCLC who we might approach about this.

Process

There are two different issues we have to address: the need for new authority records for genre terms, and the disposition of all the headings in existing bib records. The conversion of existing headings is a longer and different project.

There are so many things to consider in this process, including unqualified headings that may be used topically, and geography, chronology, and other facets that need to be retained. The mapping and conversion of headings is complex enough that it should be the very last thing we do, after we've created the basic set of genre authority records.

Theoretically, all topical headings would remain valid, but many of those headings may be too specific to apply to literature about the topic. It may be useful to consolidate these headings. We will need to review the topical headings to see if the parsing we do for musical works is actually useful when the heading is only used topically. Reviewing topical headings may be a very important step. If all topical headings were to remain without review, it would create the opportunity for institutions to continue using the topical headings for musical works and to forego use of the genre and medium vocabularies, resulting in split practice. Also, if both sets of headings existed, it would be harder to manage conversion of headings.

Genre authority records and the topical records that remain need to be connected to each other. A lot of the genre headings have the LCCN/control number of the former subject heading in their records. We should ask Janis Young at PSD how these numbers were treated in other projects.

One of our next steps should be categorizing the terms in the agreed-upon list into groups such as: terms that can be easily converted to genre coding (like waltzes); terms that cannot be easily converted, like pre-coordinated strings that need to be made into reference records or cancelled; and terms that need to be retained for topical use. After we have gone through the list systematically and determined the disposition of the headings, the actual work of disposition will be the last step.

This part of the project makes sense as a phased approach: create all the genre authority records; deal with the old subject records (retain, change to a reference record, or cancel); convert headings in bib records. The new records can be created before they are used. It is the distribution of new records that authorizes the headings for use, not their creation.

Future of the Genre Task Force

The group agreed that our work should continue through reviewing the hierarchies and figuring out the processes of authority record creation and bib heading conversion. When we get to training and documentation, we can review who will do this and how at that point. Not everyone would have to sign up, though we have a body of people who are very knowledgeable about the project and remain invested in it. We will likely pass the steps of reviewing topical headings and establishing the new terms (Garland terms) on to a new group, which could include current task force members if they wish to continue.