May 6, 2003

Dear Glenn,

At the BIBCO/CONSER Operations Committee Meeting, held May 1-2, 2003, the subject of Spanish language records in OCLC was discussed. According to OCLC Enhance policies, when authenticating these non-English language records for BIBCO or CONSER, PCC libraries are to convert any existing 040 subfield $b from “spa” to “eng.” Further, catalogers are to convert any elements—especially the 300 and 5XX fields—into English, thus, in effect eliminating any evidence of the original Spanish language cataloging except for the 6XXs (if any and only if appropriately tagged). This practice has caused concern among colleagues in the CONSER community, especially as we try to build bridges of trust and respect across the borders with our SCCTP workshops. The BIBCO community too has expressed concern and some libraries have decided not to authenticate as BIBCO rather than to remove the Spanish data.

We are proposing that OCLC adopt a policy of allowing for “parallel records” (so as not to say duplicate records), similar to those now created by the National Library of Canada in French. PCC libraries would create a copy of the Spanish language record and add or convert elements that reflect the English language environment without “obliterating” the original record. (Robert Bremer suggested at the meeting that we find a way to tie the records together via OCLC record number and we would certainly be willing to do this.) There is no consideration at this time of adding Spanish-language records to the CONSER database or authenticating Spanish language records for BIBCO. The issue at the moment is on allowing non-PCC Spanish language records to remain in OCLC and not obliterate the Spanish data while performing first time authentication for CONSER or BIBCO.

We do not anticipate requesting this treatment for other languages at this time. We feel that it is important to recognize Mexico as a member of North America and the large number of other Spanish-speaking countries in the western hemisphere who are potential contributors to OCLC. While a number of these libraries have or will soon join NACO, we do not think it is feasible to expect catalogers to discern which records are coming from NACO libraries and suggest that the new policy relate to any Spanish language record, with the exception of those from vendors.

We are making this request as an interim measure, looking forward to OCLC’s future
capability to provide multi-lingual displays of a record. Indeed, we believe that by preserving the Spanish data now, we are enabling this future capability. We believe that this would be a good public relations move both for the PCC and for OCLC. We thank you for your continued support and look forward to swift implementation of this request.

Jean Hirons
CONSER Coordinator

Ana Cristán
BIBCO Coordinator

Glenn Patton
Director, WorldCat Content Management.
OCLC, Inc.
6565 Frantz Rd.
Dublin, OH 43017
MEMORANDUM

May 12, 2003

TO: Cataloging and Metadata Interest Group

FROM: Glenn Patton

SUBJECT: Background for the Non-English Records Discussion

As Interest Group members will remember from previous discussions about plans for WorldCat in the Oracle environment, part of those plans call for the ability to have multiple, parallel “language of cataloging” versions of bibliographic records so that, depending on individual institutions’ needs, catalogers could see a French version (with notes and subject headings in French) or a Spanish version (with notes and subject headings in Spanish). It is still OCLC’s intention to implement this capability as part of the “globalization” of WorldCat but that implementation is at least a year away.

Multiple language versions of records are not new in WorldCat. For almost the entire history of the CONSER Program, the National Library of Canada (because of legal requirements) has added parallel records in French and English to the CONSER database as hosted within WorldCat. When OCLC began to load Books Canada records received from the Library of Congress, the same situation applied. However, when French university libraries began to create new records online, the agreement was that, if an OCLC Enhance library encountered a record with French notes and subject headings, it was acceptable to convert the notes in that record to English and to add LCSH headings while retaining the RAMEAU subject headings used by French libraries. As records from other libraries (and vendors) for which English is not the language of cataloging have been added to WorldCat, this conversion practice has continued.

In the past year or so, it has become increasingly clear that this practice is a barrier to the expansion of the OCLC cooperative into new markets. Subsequent conversions into English cause the perception that the initial record was in some way unacceptable even though the intention of the conversion is to produce a record that is more immediately usable by other OCLC member libraries.

As a result, OCLC has already changed the guidelines in use by Enhance libraries to discourage the conversion of non-English records except in specific cases such as vendor-supplied records. (See the Enhance Training Outline on the OCLC web site at http://www.oclc.org/worldcat/enhance/trainingoutline.shtm). Still, the perception is that members of the cooperative undervalue records using other “languages of cataloging”. In addition to comments received from Latin American member libraries, OCLC has also been contacted by members of SALALM (the Seminar on the Acquisition of Latin American Library Materials) and by participants in the Program for Cooperative Cataloging. Both groups favor finding some kind of solution to this issue that would preserve the integrity of the original language record while providing for other versions that meet other libraries’ needs.

OCLC staff are now considering a further interim solution that would allow the creation of multiple, parallel “language of cataloging” versions in the current WorldCat environment.
There are a variety of factors that need to be considered before deciding to allow the input of parallel “language of cataloging” records, including these:

- Many users, especially end-users of the FirstSearch interface, complain about “duplicate” records in WorldCat even when those records are legitimately created according to standard cataloging practice. Adding “language of cataloging” versions will likely exacerbate this perception since major changes to interfaces to make the versions clearer may not be immediately implementable. This would have an obvious impact on WorldCat’s image as a usable, quality database.
- Multiple, parallel records created in the current “flat file” environment will, in the future, have to be converted into the new database structure. There is no obvious, “legal” way in MARC 21 to link records together so that future conversion will be easier. We are considering how existing linking fields (fields 76X-78X normally used to link bibliographic entities) might be “misused” to link records but consideration must be given to what kind of message this sends about OCLC’s adherence to standards.
- Creation of a larger number of new records has an obvious budget impact for OCLC.
- A record created by the Library of Congress (or a PCC institution) batchloaded to WorldCat would likely cause the set of parallel records to be merged into one until such time as Batchload processing is reimplemented in the Oracle environment (since we have passed the point where extensive changed to existing software can be cost-justified).

OCLC staff are currently working to gather data about the number of records that have been changed in the past year or so and the kinds of changes that have been made (addition of call numbers versus complete recataloging of the record, for example). We hope to have that data available to assist in our discussion.

Please give some thought to these issues and discuss them with your colleagues so that you can provide advice on this topic at the Tuesday morning meeting.
Spanish WorldCat Records Created by Libraries in Mexico Converted to English (2003 May 15)

The following statistics represent new records added to WorldCat since January 2000 (beginning with OCLC # 43,000,000) in which the language of cataloging (Field 040 $b) has been changed from Spanish to English by another OCLC member library. The five libraries represent the top 5 creators of new records in Mexico in fiscal year 2002/2003 year-to-date.

<table>
<thead>
<tr>
<th>Name of Institution</th>
<th>Symbol</th>
<th>Total Input</th>
<th>Records Now Lacking 040 $b spa</th>
<th>Percentage of Total Input</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universidad Autonoma de Ciudad Juarez</td>
<td>DQ$</td>
<td>16,080</td>
<td>82</td>
<td>0.5%</td>
</tr>
<tr>
<td>El Colegio de Mexico</td>
<td>EL$</td>
<td>93,524*</td>
<td>561</td>
<td>0.6%</td>
</tr>
<tr>
<td>Universidad Autonoma de San Luis Potosi</td>
<td>UXS</td>
<td>2,434</td>
<td>22</td>
<td>0.9%</td>
</tr>
<tr>
<td>Universidad de Monterrey</td>
<td>UN@</td>
<td>3,829</td>
<td>53</td>
<td>1.4%</td>
</tr>
<tr>
<td>Benemerita Universidad Autonoma de Puebla</td>
<td>UBP</td>
<td>4,461</td>
<td>34</td>
<td>0.8%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>120,328</td>
<td>752</td>
<td>0.6%</td>
</tr>
</tbody>
</table>

* The total records for EL$ includes a large retrospective batchload project.