

PCC SCS Non-Latin Script References in Name Authority Records Webinar (2024-06-24): Questions & Answers

Note: This document includes a compilation of most of the questions posed during the June webinar. These answers are pending the finalization of the Task Group's (TG) final report and its approval by SCS/PCC and are therefore subject to change.

Q1. What is a preferred non-Latin script reference/variant and how is it chosen? Is the “preferred” non-Latin reference determined by frequency?

A1. The preferred non-Latin script form is determined by the usage in resources written in those languages/scripts or by the commonly found form in those resources. Resources include the NAR itself, usage in OCLC or other cataloging utilities, national catalogs, reference works, or other relevant sources if necessary. The preferred non-Latin script form may or may not match the romanized form in a 1XX and may include additions to differentiate it from preferred non-Latin script variants in other NARs.

Q2. When we check for predominant usage, it is quite possible that not all OCLC records may have names transcribed in parallel non-Latin form. Do we limit ourselves to non-Latin usage readily available in OCLC, which technically could differ from the form that is actually predominant in case all non-Latin usage is available?

A2. Catalogers are not required to consult beyond their catalog of record (e.g., OCLC). However, catalogers may consult other resources, such as national library catalogs, for particular languages if necessary.

Q3. What will we do if we do not agree with non-Latin qualifiers in NARs?

A3. A cataloger may add another variant with a qualifier they find useful, though this variant may not be designated as preferred. Catalogers will need to respect the previous cataloger's judgment in designating the preferred form similar to how we currently respect the cataloger's judgment with 1XX. Of course, just as 1XX forms are changed when deemed to be wrong or no longer the preferred form, the same would be done for preferred non-Latin script variants. Changes would be justified in 670s as are Roman script changes.

Q4. Will there be a consideration of headings created by national libraries regarding preferred access points? e.g. National Library of Israel (NLI) authority records, etc. The national library form outweighs the more commonly held/found form, right?

A4: Any resource, including national library catalogs, may be considered when deciding the most commonly found form. There is no PCC policy requiring NACO catalogers to use the preferred form found in a national library's authority file.

Cataloger's judgment continues to play a role, too. An example in our webinar (slide 12) showed the NLI form for Avicenna (fuller form of name with patronymic) did *not* outweigh the commonly found form. All these forms found would be cited in the 670s and may be given as variants.

Reminders: A national library *may or may not* follow RDA, may have its own policy statements or interpretations, and the sources they consult or have available may differ from those available to NACO catalogers. The cataloger may choose the form used by national libraries if it appears to be the most commonly used. These forms would need to be justified in the 670s.

Catalogers are always free to look for other evidence, but in most cases they aren't required to look beyond what's already in the record (e.g. 670s) when evaluating an existing record.

Q4b. I would think that for Hebrew cataloging, referring to the original script in Hebrew as established by the National Library of Israel would be a valuable addition to the NAR.

A4b. National library forms can be cited in 670s and given as added variants as appropriate. We don't want much free text (e.g., "\$i Preferred variant in Hebrew (Hebrew script) *per NLI*: \$a...", etc.) in the \$i notes for preferred form references.

Q5. Is the idea of "language of cataloging" breaking down here? With qualifiers etc. translated into other languages, the goal seems to be to construct, e.g. an Arabic heading that would be appropriate in a record with Arabic as language of cataloging.

A5. In authority records, "language of cataloging" partly refers to notes and may also apply to qualifiers. In the interest of promoting internationalization and increasing the use of non-Latin script references, the TG recommends using non-Latin script qualifiers as much as possible and when readily available. We envision a system that allows a user to set a preferred language/script; the designated preferred non-Latin script variant for that language/script would likely be the authorized access point in such a setting.

Q6. What is the reason to recommend using non-English words (or picking translated words) for qualifiers while English is the cataloging language?

A6. The TG discussed this extensively and felt that providing the term used in qualifier in the same language and script as the rest of the text string facilitates user access. Also, the appearance/display of mixed languages/scripts in a text string looks strange.

Q7. If the preferred non-Latin form for a subordinate body is

[Preferred non-Latin form for parent body]. [Non-Latin form for subordinate body],

can we still give a 410 for

[(Latin) AAP for parent body]. [Non-Latin form for subordinate body]?

A7. The latter form is a valid variant, though it would rarely if ever be considered the preferred variant. In other words, you can still provide a romanized form in the field that includes the non-Latin script; non-Latin script qualifiers would not be romanized.

Q8. Will there be guidance on how to deal with words in dates in right-to-left scripts if the cataloger deems it necessary to add to resolve a conflict? Language of the words, order of elements in the subfield are what I am wondering about.

A8. Mixed script variants can be tricky and present display issues in many systems for right-to-left scripts. For this reason the Hebrew script community generally prefers to omit dates. On the other hand, Arabic script catalogers generally include dates in authority records. The Hebrew user community may need to first decide on their best practices then create guidance accordingly.

If a cataloger chooses to include dates, the TG recommends using the same form as in the authorized access point (i.e., including any English words associated with the dates).

Q9. What is considered “correct” when evaluating a reference? My question refers to the structure of a non-Latin equivalent to something like “approximately 1650-approximately 1714.”

A9. For a reference/variant to be considered “correct” (or “valid”), the record must include justification for that reference. There should be no typos, and the construction of qualifiers must follow RDA instructions and PCC/NACO policies. The TG recommends continuing to use English words for dates in most situations. The exception to this is a term such as “jin shi” (进士) in simplified Chinese script and (進士) in traditional Chinese script, which indicates a period of activity related to obtaining a degree.

Q10. Sometimes RDA requires a qualifier, usually to differentiate from other AAPs with the same text. So I was confused by the statement that for non-Latin preferred variants, qualifiers are (always) optional. 1. This contradicts the idea that preferred non-Latin

variants should be formulated according to RDA. 2. I now think that Bob was NOT implying that non-Latin qualifiers are “always” optional. That is, if they are needed for differentiation, you need to include them...

In short, currently, it’s OK if two 4XX fields are identical in different NARs, but should preferred variants in non-Latin scripts be unique?

A10. While a non-Latin 4XX can be the same as another 4XX on another NAR, preferred non-Latin script variants should be unique, similar to how an authorized access point in the 1XX should be unique. Other non-preferred variants can conflict between NARs, although catalogers are also free to differentiate them (such as is often done with identical acronyms for different corporate bodies).

Q11. Should bibliographic records use the preferred 4XX form? And if yes, do we know in what cases a cataloger would need to make changes to the preferred 4XX? Would such changes require BFM?

A11. The preferred variant, if provided in the authority record, should be used in bibliographic records in the 880 field. If, down the road, a cataloger identifies that the designated preferred non-Latin script variant is no longer the a preferred variant, the authority record will be updated and BFM may be involved.

It is hoped that authority vendors will take advantage of variants that are coded as preferred to do bibliographic file maintenance, and that ILS system developers will take advantage of the coding to flip the paired vernacular heading when the preferred heading in an authority record changes. Such hopes may be overly optimistic, but we might see this in future developments.

Q12. Has the use of ISO 639 language codes that are not part of MARC language codes been considered?

(in reference to <https://www.loc.gov/aba/pcc/scs/documents/ISO-639-3-guidelines.pdf>)

A12. In the examples shown on the slides, the Task Group used MARC codes since we knew people would be most familiar with them. No policy has been decided as to what codes will be used. In our testing we ended up using coded subfields, so we need to consider whether MARC codes are sufficient, whether we should use ISO codes instead, or if we should use the ISO code only if a MARC code isn’t available.

Currently, only a limited number of non-Latin script languages are supported in the NACO authority file. We hope the file will expand to include all languages and scripts. After expansion, we would want to include all possible language codes.

Q13. Is it possible for the preferred non-Latin script 4XX to move to 1XX for globalization?

A13. Such coding is not possible in the current NACO environment. NACO authority records can have only one 1XX field, and it must contain the authorized access point in Latin script.

Q14. Question about using non-Latin scripts in original work authority records and the authority record for the translation: Isn't this practice redundant? The question is based on the example of the Japanese variant for the Toni Morrison work. The Japanese form would be the same in the original work record and the expression record for the Japanese translation. Why add the same reference to both?

A14. The Expression AAP should have a language attribute at the end of AAP, so it wouldn't match the Work AAP.

An expression authorized access point that is RDA compliant requires an expression element. Most of us do not currently follow this practice, but we're having more discussions about it. We had robust conversations about variant titles of work, and it's clear to us that a work may be known by many different titles in different languages, and those are legitimate variants that could be recorded in work authority records. Consider for example Tolstoy's War and Peace; it is commonly called "War and Peace" when people are writing English-language dissertations about the work originally called "Voïna i mir." They are not writing about the English expression of "War and Peace;" they are writing about the original work that they call "War and Peace." This is the same practice shown in the Toni Morrison examples.

Q15. Is the TG considering non-Latin variants in SARs? Possibly in the future?

A15. Series authority records are covered under Works and Expressions, and those guidelines would apply as discussed in the webinar.

Q16. Authority work is already so labor-intensive and time consuming. Is having to determine the "preferred non-Latin form," not to mention mastering all of these new guidelines, going to add significant time? Curious to know how/if the TG has grappled with this.

A16. Yes! Authority work can be time consuming, and we understand that implementing these recommendations will require time and training. The TG discussed extensively the need to determine preferred non-Latin forms. The reasons for designating a preferred form are outlined on slides 9 and 10. Designating the preferred language/script form has several benefits as presented in the slides, and the practice has been approved by the PCC Standing Committee on Standards. Furthermore, the practice of not evaluating non-Latin script references was always seen as temporary (and not ideal). These recommendations effectively put non-Latin script

references on equal footing with Latin script references. As such, NACO catalogers should follow the proposed guidelines when creating new authority records or updating existing NARs containing non-Latin scripts.

Q17. I'm concerned about the use of non-Latin script variants for 6XX fields in bibs. Subject Headings Manual (SHM) instruction sheet H 182 says that we cannot use parallel subject fields unless they correspond exactly to the ALA-LC romanization.

A17. The new guidelines apply only to NACO work not to subject analysis and assignment. However, the TG recommends changing the practice advised in H 182, which is to use an “exact correspondence” form, to instead using the preferred non-Latin script form. The SHM rule was probably made to impose some semblance of order to bib records; using the preferred non-Latin script form may be a better way of imposing this order.

Q18. Please continue to encourage folks to include translations/summaries in English in the authority records. They are essential for folks who may be cataloging English titles and are searching for NARs for entities. Having qualifiers in English for variants also helps with this. Often publishers do not include names/titles in the original scripts for translations.

A18. The TG agrees with this and may include such encouragement in its final report, but the form and content of 670 fields in authority records, which we assume this question refers to, does not fall within the TG's charge.

Adding another variant with the qualifier in English is optional when the qualifier in the preferred form is in a non-Latin script language.

Q19. Are the changes to preferred names only a first step and will geographical names follow next? E.g., Munich vs. München

A19. No, preferred forms for geographic AAPs (MARC 151) will almost always be based on the most commonly known (conventional) form of a place name in English, if there is one. Catalogers should follow LC-PCC policy statements and metadata guidance documentation for determining the authorized access point for a place (which varies by language).

Q20. What about field 700 for linking name authority record with title record in both Latin and native script?

A20. This issue does not fall within the TG's charge.

Language specific questions

Note: These language-specific answers depend on the finalization of the TG's report and its approval by SCS/PCC, implementation of MARC coding changes, and creation of best practices by each language community. They are thus subject to change.

Q21. For the Romanov example in slide 16 below, is ‘РОМАНОВЫ’ [= Romanovy] the preferred form of the pre-1917 reform?

100 3# Romanov (Dynasty : #d 1613-1917)

400 3# #w r #i Preferred name in Russian (Cyrillic script): #a РОМАНОВ (Династия : #d 1613-1917)

A21. RDA 10.2.2.3 says “When choosing a preferred name for family, generally choose the name by which the family is commonly known. The name chosen can be the surname (or equivalent) used by members of the family, the name of a royal house or dynasty, or the name of a clan, etc.” The surname used by members of this family is РОМАНОВ. RDA access points for families are established in singular form. This also applies to non-Latin script surnames.

Q22. Can the preferred variant use a shorter but more popular form? For example, 110 China: Nei wu fu; 410: 内府 [= Nei fu] as the preferred Chinese variant.

A22. The preferred form should be the most commonly known form in a given language and script. So yes, if the shorter form is a more popular form in the original script, it will be designated as the preferred variant.

The preferred form used in the 1XX would usually be based on the preferred form in non-Latin, so there would usually be a one-to-one correspondence when the 1XX is romanized from a non-Latin script language. If the forms don't correspond, it could be that the 1XX needs to be reevaluated and changed to reflect the preferred non-Latin script form .

Q23. Doesn't “昭和天皇” [= Shōwa Tennō] consist of a name followed by a title?

A23. “昭和” [= Shōwa] by itself is not a personal name and this is probably why the National Diet Library (NDL) treats “昭和天皇” [= Shōwa Tennō] as a name. If we add \$c before 天皇 [= Tennō], then it would be “昭和, \$c 天皇” [= Shōwa, \$c Tennō] which is not a commonly used form.

Q24. I'm wondering about the Chinese forms of historical people. Would the one preferred form for Sima Qian (c.145-c.86 BC) be traditional Chinese 司馬遷 [= Sima Qian] or simplified Chinese 司马迁 [= Sima Qian]? Was it possible in the Korean example to have two preferred forms: Hancha and Hangul forms, or did I misunderstand?

A24. The TG recommends that both simplified and traditional Chinese forms of the name be designated as preferred variants for each script.

Yes, we can designate both Hangul and Hancha as preferred variants in separate fields. The slides do not show every possible variant.

Q25. Interesting. I would have thought that ם״רמב [= Rambam] would be the preferred Hebrew script form for Maimonides...

A25. Acronymic forms can be given as alternate Hebrew variants if they are not chosen as the preferred variant form. Per the LC-PCC PS on 9.2.2.5.3, Hebrew alphabet, clause 2, the cataloger consulted the following reference sources: *EJ2* online: the entry is MAIMONIDES, MOSES and article begins “(Moses ben Maimon; known in rabbinical literature as ‘Rambam’; from the acronym Rabbi Moses Ben Maimon; 1135–1204)”, while the knowledge card panel has “About this Person... Other Names: Moses ben Maimon; Maimonides, Moses; Abu Imran Musa ibn Maymun ibn Ubayd Allah; Rambam; Moshe ben Maimon”. *EJW*: the entry is Maimonides, Moses and the article begins “Moses ben Maimon, generally referred to as Maimonides or by the Hebrew acronym Rambam...”. Because what is sought is the preferred vernacular form, the vernacular reference sources consulted included: *NLI NAF* online: preferred form is משה בן מימון [= Mosheh ben Maimon]. *EI*: the entry is for משה בן מימון [= Mosheh ben Maimon]. *Hebrew Wiki*: the entry is for ם״רמב [= Rambam]; the article begins with רבי משה בן מימון [= Rabi Mosheh ben Maimon]. Based on all the reference sources consulted, the cataloger judged that the fuller form משה בן מימון [= Mosheh ben Maimon] is preferred and gave a variant to the acronymic ם״רמב [= Rambam].

The Hebrew user community may want to consider what Hebrew reference sources should be used in determining the preferred Hebrew variant.