



Non-Latin Variants in Name Authority Records: Different Perspectives, New Practices

PCC Informational Discussion/Online Forum
with NACO Catalogers Working with Non-Latin Scripts

Hosted by: PCC SCS Task Group on Evaluation Guidelines for Non-Latin Script
References in Name Authority Records

Presented by: Iman Dagher, Shi Deng, Bob Maxwell, Adam L. Schiff

June 20, 2024

TG statement: The Task Group members have been working together for almost a year. It took a long time for us to reach consensus, and our work is still in progress. We ask that you please hear our proposals with an open mind. Nothing is final, and we will solicit your input. The overall goal is to help users by increasing the presence of non-Latin script references in library records and, ultimately, improving discovery of resources.

Task Group membership

- **Iman Dagher** (Co-chair) (University of California, Los Angeles), Representative for Standing Committee on Standards, Arabic NACO Funnel, and ALA/Core Committee on Cataloging: Asian and African Materials
- **Shi Deng** (Co-chair) (University of California, San Diego), Representative for CJK NACO Funnel
- **Erica Chang** (University of Hawai'i at Mānoa), Representative for CJK NACO Funnel Reference Project
- **Bob Maxwell** (Brigham Young University), Representative for Greek script, NACO trainers, and NARDAC
- **Adam L. Schiff** (University of Washington), Representative for NACO trainers and NARDAC
- **Jasmin Shinohara** (University of Pennsylvania), Representative for Hebraica NACO Funnel
- **Naomi Shiraishi** (University of California, Berkeley), Representative for CEAL Committee on Technical Processing
- **Larisa Walsh** (University of Chicago), Representative for ACRL/ESS Slavic Cataloging and Metadata Committee
- **Jessalyn Zoom** (Library of Congress, Acquisitions and Bibliographic Access Directorate, Asian and Middle Eastern Division), Representative for Library of Congress and ALA-LC Romanization Review Board
- Representing languages and scripts in MARC-8 repertoire in NAF (Arabic, Chinese, Cyrillic, Greek, Hebrew, Japanese, Korean, Persian, Yiddish)
- Representing different Non-Latin Communities:
 - Arabic NACO Funnel
 - CJK NACO Funnel
 - Hebraica NACO Funnel
 - Cyrillic script
 - Greek script
- Representing different committees/groups:
 - PCC Standing Committee on Standards
 - PCC NACO Trainers
 - ALA Core Committee on Cataloging: Asian and African Materials (CC:AAM)
 - North America RDA Committee (NARDAC)
 - LC and ALA-LC Romanization Review Board
 - CEAL Committee on Technical Processing
 - CJK NACO Funnel Reference Project
 - ACRL/ESS Slavic Cataloging and Metadata Committee

2

This presentation represents the work of this task group, we are presenting on their behalf.

This slide shows our task group members. As you can see, this group represents the language expertise of the non-Latin communities limited to the MARC-8 repertoire. And we would like to start by acknowledging the extensive work of each member of the group and for their valuable and indispensable contributions.

We would also like to acknowledge the PCC Policy Committee and the PCC Standing Committee on Standards for supporting the Task Group's work.

We will also remind everyone that our work and examples are limited to the non-Latin languages and scripts in MARC-8 repertoire.

Task Group charges

Task Group Charges

- Consider and respond to the questions and issues raised in the PCC White Paper and SCA and SCS reports
- Survey and communicate with stakeholders in the non-Latin script NACO funnels and cataloging communities, to identify existing guidelines and best practices; to determine gaps, needs, and desired outcomes; and to understand the extent to which consensus and alignment of practice across script-based cataloging communities is possible or desirable
- Develop and propose guidelines needed to support evaluation, formulation, and revision of non-Latin script references in Name Authority Records, including supporting documentation (such as FAQ, examples, or training materials, consulting with the Standing Committee on Training (SCT) as needed)
- Make recommendations regarding timeline and strategies for evaluation and updating of existing records where appropriate (for example: establish a “Day One” for implementation; suggest coordinated projects; identify categories of records amenable to machine processing)

Five questions from the charge

1. What does “evaluation” mean, and does it mean something different for non-Latin script references than it does for Latin-script references?
2. Should there be separate guidelines for non-Latin script references and Latin script references, or do they need to be aligned?
3. Are the varying practices in formulation of qualifiers and dates in non-Latin access points in bibliographic records, including those formulated according to the PCC Guidelines for Creating Bibliographic Records in Multiple Character Sets, acceptable or unacceptable as references in evaluated NARs?
4. Is there any need to indicate a particular reference in an original script as “preferred,” and if so, how?
5. Are non-established forms of elements in variants for hierarchical corporate names and name/title NARs acceptable or unacceptable as non-Latin references in evaluated NARs

The TG has been working on discussing and addressing all five questions, but today will focus on questions 3 and 4, as decisions on these two questions will affect the proposed practice in dealing with non-Latin script references in name authority records and will affect other related cataloging practices

Current practices with non-Latin script references: Quick overview

Rec stat	c	Entered	19810515	Replaced	20170310073919.0				
Type	z	Upd status	a	Enc lvl	n	Source	400	0	Mohamed Hassanein Heikal
Roman	■	Ref status	b	Mod rec		Name use	400	0	Muhammad Hasanayn Haykal
Govt agn	■	Auth status	a	Subj	a	Subj use	400	1	הייכל, מוחמד חסנין
Series	n	Auth/ref	a	Geo subd	n	Ser use	400	1	هيكل، محمد حسنين
Ser num	n	Name	a	Subdiv tp	n	Rules	400	1	هيكل، محمد حسنين
							400	1	هيكل، محمد حسنين
010		n 81045286					400	1	هيكل، محمد حسني
040		DLC #b eng #e rda #c DLC #d OCoLC #d UPB #d					400	1	هيكل، مُحَمَّدُ حَسَنِين
046		#f 1923-09-23 #g 2016-02-17 #2 edtf					400	1	محمد حسنين هيكل
100	1	Haykal, Muḥammad Ḥasanayn					667		Machine-derived non-Latin script reference project.
							667		Non-Latin script references not evaluated.
							670		His Autumn of fury, 1983: #b t.p. (Mohamed Heikal)

Before we get into what will be our future practices, let's look at our current practices:

- Currently the Name Authority File includes non-Latin script variants that have been added by catalogers and/or the machine-populated project in 2008 based on bibliographic headings that match entities in the authority file. These latter NARs are flagged with a 667 note "Machine-derived non-Latin script reference project" as shown in the example on this slide, the first 667 note.
- NACO catalogers creating authority records for non-Latin names can *optionally* add non-Latin script variants to their records. When adding non-Latin script variant to NARs, we should code reference status in fixed field 008/29 "b" and add a 667 note: "Non-Latin script reference(s) not evaluated." as shown in this slide. They were also added to the NARs with machine populated non-Latin script variants.
- These non-Latin script variants are essential for enhanced discovery since romanization is not always efficient in discovery. However, the machine-populated non-Latin script variants are not always correct and need to be evaluated
- Currently, LC/PCC have limited instructions regarding non-Latin script references in RDA and DCM Z1, and no guidelines for evaluating non-Latin script references have been established
- Each non-Latin community may have developed its own best practices, and even within a community, best practices have not been consistently applied. Practices between communities are also not consistent

- Also, not all non-Latin script groups have developed a community practice

NACO CJK Funnel reference project

- CJK NACO Funnel launched the project in May 2019 in three phases (currently in phase 2 and 3). Project completion is expected in 2025.
- Project volunteers (CJK NACO catalogers) are reviewing differentiated CJK personal name authority records against mainly OCLC (LC, VIAF, etc. as necessary)
- Any incorrect CJK non-Latin script variant(s) are removed
- After completion of a record, a 667 project note, “Non-Latin script references reviewed in NACO CJK Funnel References Project” is added.
- Other two existing 667 notes, “Machine-derived non-Latin script reference project” and “Non-Latin script references not evaluated”, and the Reference Status (008/29) code “b” still remain in the evaluated records unless the review of a record results in the removal of all non-Latin script variants, in which case the 008/29 value will be changed to “a” or “n” and the corresponding 667 field “Non-Latin script reference not evaluated” will be removed
- With evaluation, non-RDA record is re-coded to RDA

6

I would like to talk very briefly about the NACO CJK Funnel Reference Project.

- It was launched in May 2019, focusing on evaluating machine populated CJK non-Latin script variants for a pre-identified set of 80,956 differentiated CJK NARs that have more than one non-Latin variants and with the pair of 667 notes “Machine-derived non-Latin script reference project” and “Non-Latin script reference(s) not evaluated.” The project is structured into three phases with the current focus being on phases 2 and 3. The project completion is expected in 2025.
- Project volunteers are about 40 CJK NACO catalogers. They review these NACO records based on the usage(s) found in bibliographic records cited in the existing 670(s) either in OCLC or LC database or other sources, and remove any incorrect CJK non-Latin script variants.

Future evaluation process: New coding

- Fully evaluated records will be coded 'a', not 'b' in 008/29 ["Ref status" in OCLC]. No 667 note "Non-Latin script reference(s) not evaluated" will be added
- **Existing records:**
 - For records with one or more non-Latin references, apply the new coding when *all* references have been evaluated: 008/29 'a', delete 667;
 - If *not all* variants have been evaluated, the 667 note should be revised to note which ones have not yet been evaluated; 008/29 'b', revise 667;

Examples

 - 667 ## \$a Chinese and Korean script references evaluated. Hebrew and Japanese script references not evaluated.
 - 667 ## \$a Arabic script references evaluated. Other script references not evaluated.
- **New records:** Apply new coding when adding new records with non-Latin variants (i.e., all new non-Latin variants should be evaluated).

7

In the future, the coding will change upon evaluation. Fully evaluated records will be coded "a" not "b" in the fixed field, and no 667 note, "Non-Latin script reference(s) not evaluated" will be added.

For existing records: if the record has one or more non-Latin script references, we apply the new coding when *all* references have been evaluated.

If not all references have been evaluated, the 667 should be revised to note which ones have not yet been evaluated as shown in the examples here.

The format of the 2nd example may be used if the cataloger is not familiar with the language/script of the remaining variants.

We apply the new coding for all new records, which is to say that henceforth, all non-Latin script variants in new records should be evaluated.

Indicate a particular reference in non-Latin script as “preferred”: Guidelines

What:

- One preferred non-Latin form for each language/script
- Other evaluated non-Latin variants for the same language/script can also be added (but not designated as preferred)

Who:

- Catalogers with language expertise evaluate the non-Latin reference

How:

- Follow recommendations of the task group for each language/script and for each type of entity
- Add special MARC coding for designating the preferred non-Latin script variant
- Change 008/29 (“Ref status” in OCLC) from ‘b’ to ‘a’
- Delete 667 note “Non-Latin script reference(s) not evaluated” when evaluation is complete
- Use preferred form as needed in other NARs and in the parallel access points in bibliographic records (880 fields)

8

Back to the main question from our charge: Is there a need to indicate a particular reference in an original script as “preferred,” and if so, how?

The Task Group discussed this question in our charge thoroughly and extensively, and we all agree that designating One preferred non-variant reference for each language and script would be beneficial.

Records could also include other evaluated non-Latin script variants for the same language/script that are not designated as preferred.

Catalogers with language expertise are encouraged to evaluate the non-Latin script reference and designate the preferred variant.

That would be done by:

- Following the guidelines recommended by the task group for each language/script and for each type of entity
- Adding special MARC coding for designating the preferred non-Latin script variant (covered in a few slides)
- And as mentioned earlier, we change the 008/29 (“Ref status” in OCLC) from “b” to “a as applicable
- Delete the 667 note “Non-Latin script reference(s) not evaluated” when evaluation is complete

- The preferred form will be used as needed in other NARs and in the parallel access points in bibliographic records (880 fields)

Preferred non-Latin script variants: Why?

Preferred Variants: Indicate a particular variant in a non-Latin script as “preferred” for a given language and script in order to

1. Promote consistent practice among NACO catalogers
2. Enable the use of the preferred non-Latin script form(s) in variant access points in related authority records (e.g., in qualifiers; name-title; parent and subordinate corporate bodies)

110 2# Ittiḥād al-Jāmi'āt al-'Arabīyah
410 2# [Preferred Variant] اتحاد الجامعات العربية
410 2# اتحاد جامعات الدول العربية

110 2# Ittiḥād al-Jāmi'āt al-'Arabīyah. #b
Jam'iyat Kulliyāt al-'Ilām al-'Arabīyah
410 2# اتحاد الجامعات العربية. #b جمعية كليات الاعلام العربية

100 0# Avicenna, #d 980-1037

400 0# ابو علي بن سينا

400 0# ابو علي الحسين بن سينا البخاري

400 0# ابو على سينا

400 0# [Preferred Variant] ابن سينا، #d 980-1037

400 0# الشيخ الرئيس ابن سينا

400 0# بو علي سينا

100 0# Avicenna, #d 980-1037. #t Hidāyah

400 0# هداية #d 980-1037. #t ابن سينا،

This new practice of designating a preferred non-Latin script form would :

1. Promote consistent practice among NACO catalogers
2. Enable the use of the preferred non-Latin script form(s) in variant access points in related authority records (e.g., in qualifiers; name-title; parent-subordinate corporate bodies) as illustrated in the examples on the slide.

Preferred non-Latin script variants: Why?

3. Ensure consistent forms are used in bibliographic records parallel (880) fields, thereby enhancing collocation of non-Latin script access points for searching and retrieval

100 1# #a Chekimoglou, E. A. #q (Euangelos A.)

400 1# #a Χεκίμογλου, Ευάγγελος Α.

400 1# #a Χεκίμογλου, Ε. Α. #q (Ευάγγελος Α.) [Preferred Variant]



100 1# #a Χεκίμογλου, Ε. Α. #q (Ευάγγελος Α.), #e author

100 1# #a Chekimoglou, E. A. #q (Euangelos A.), #e author.

245 10 #a Θεσσαλονικη : #b τεκμηρια φωτογραφικου αρχειου, 1900-1980 / #c κειμενα

Ευαγγελος Χεκιμογλου.

245 10 #a Thessalonikē : #b tekēmēria phōtographikou archeiou, 1900-1980 / #c keimena
Euangelos Chekimoglou.

4. Facilitate potential system configuration for conducting searches using the preferred non-Latin script form corresponding to a user's preference

And would:

3. Ensure consistency in form used in bibliographic records parallel (880) fields. This will enhance collocation when using non-Latin script access points as illustrated in the example.

4. Facilitate potential future system configuration for conducting searches using the preferred non-Latin script form corresponding to a user's preference

Options for coding preferred non-Latin script variant access points in MARC

- **Option 1.** Use \$w r and \$i in 4XX fields

Requires NACO nodes validation changes

400 0# \$w r \$i Preferred variant in Russian (Cyrillic script): #a Абу Али ибн Сина, #d 980-1037

400 1# \$w r \$i Preferred variant in Chinese (Traditional Chinese script) and Korean (Hancha script): #a 林語堂, #d 1895-1976

- **Option 2.** Use \$7 data provenance subfields in 4XX fields to encode the language, script, and preferred status of non-Latin script variants.

Requires implementation of \$7 in NACO

The data provenance category codes that we could use in \$7 are:

dpeloe = Data provenance element language of expression

dpes = Data provenance element script

dpecou = Data provenance element context of use

400 0# #d 980-1037 \$7 (dpeloe)ara \$7 (dpeloe)per \$7 (dpes)Arab \$7 (dpecou)Preferred non-Latin form

400 1# 林語堂, #d 1895-1976 \$7 (dpeloe)chi \$7 (dpes)Hant \$7 (dpeloe)kor \$7 (dpes)Hani \$7 (dpecou)Preferred non-Latin form

- **Option 3.** Establish a new code to use in \$w to indicate a preferred form as well as a way to code language and script (using \$7 or some new subfield)

Requires MARC discussion paper and proposal

Option 1. Allow use of \$w r & \$i in 4XX fields in authority records so that we can test our guidelines. These subfields are already available in the MARC Authority Format but NACO policy does not currently allow them. This would not require any MARC code changes but it would require changes to the DCM Z1 4XX See From Tracings – General Information and LC Guidelines Supplement Tracings and References 4XX as well as to the NACO nodes to allow for validation of records containing these two subfields (see <https://www.loc.gov/marc/authority/adtracing.html>). In this scenario, we would include a 667 note during the test period to alert PCC participants not to edit 4XX fields with \$w and \$i. The 667 field could be deleted later if option 1 is implemented fully, or it could be used to identify records that should be changed to the coding in option 2. A con for this option is that it is not very machine actionable.

Option 2. Allow the use of existing data provenance category codes in subfield \$7 in 4XX fields to encode the language, script, and preferred status of non-Latin script variants. This would require the implementation of \$7 in NACO records, and for catalogers to learn to use data provenance codes and know correct language and script codes. The slide shows one possible implementation of this option. Another possible implementation will be discussed at the upcoming MARC Advisory Committee meeting on June 25. The MARC Discussion Paper No. 2024-DP11 proposes the addition of category codes for use with MARC data provenance subfields to accommodate transliteration scheme codes and BCP (Best Current Practice) 47 tags in the MARC 21 Authority and Bibliographic Formats.

Option 3. Propose a new code to use in \$w to indicate a preferred form. Also propose a means of coding the language and script of a variant in non-Latin script in a 4XX field,

perhaps using a new repeatable subfield or subfield \$7. This will require a MARC discussion paper and proposal, and will take longer to implement, but coding language and script will be much more machine actionable and is more desirable in the long term.

Indicate a particular reference in an original language/script as “preferred”: Use cases, examples

100 0# Avicenna, #d 980-1037

400 0# أبو علي بن سينا

400 0# أبو علي الحسين بن سينا البخاري

400 0# ابن سينا

400 0# ابن سينا, #d 980-1037

400 0# ابن سينا، أبو علي الحسين

400 1# ابن سينا، الحسين بن عبد الله، #d 980?-1037

400 0# ابن سينا, #d 980-1037 [= Ibn Sīnā; cataloger determines this is preferred Arabic script form: the common form in Arabic resources; cataloger has chosen to include dates]

400 0# الشيخ الرئيس ابن سينا

400 0# אבן סינא [= Ibn Sina; cataloger determines this is preferred Hebrew script form; cataloger has chosen not to include dates]

400 1# אבן סינא, [National Library of Israel preferred Hebrew script form; includes patronymic, i.e., fuller form of name]

400 0# #w r #i Preferred variant in Arabic (Arabic script): #a ابن سينا, #d 980-1037

400 0# #w r #i Preferred variant in Hebrew (Hebrew script): #a אבן סינא

12

This slide shows Arabic and Hebrew script references found in the current authority record for Avicenna. Many of these references were added through an automation process and need to be evaluated. In the process of evaluating the references, an Arabic-language cataloger determines which form in Arabic is most commonly found in Arabic resources. This form can be labeled as the preferred variant access point in Arabic in Arabic script. A Hebrew-language cataloger could do the same thing to determine the preferred name in Hebrew in Hebrew script. Note that neither the Arabic preferred name nor the Hebrew preferred name romanize to the authorized access point. This is not unusual for well-known persons whose names appear in English-language resources and English reference sources. The commonly found form of name in English is selected as the preferred name when determining the authorized access point for an English-language catalog.

Note that including dates in scripts that read from right to left has been considered optional by some NACO language communities. In the example here, the Arabic cataloger chose to include the dates while the Hebrew cataloger chose not to.

Indicate a particular reference in an original language/script as “preferred”: Use cases, examples

- 100 0# Confucius [most common form in English; not a transliteration of the Chinese name]
- 400 0# Kongfuzi
- 400 0# 孔夫子
- 400 0# Kongzi
- 400 0# Kōshi
- 400 0# Kongja
- 400 0# 孔子 [cataloger(s) determine this is the preferred form in Chinese, Japanese Kanji, and Korean Hancha scripts: most commonly found form]
- 400 0# 공자 [cataloger determines this is the preferred form in Korean Hangul script: most commonly found form]
- 400 1# Kong, Qiu
- 400 1# 孔丘
- 400 1# Kong, Zhongni
- 400 1# 孔仲尼

400 0 #w r #i Preferred variant in Chinese (Simplified and Traditional scripts) and Korean (Hancha script) and Japanese (Kanji script): #a 孔子
400 0 #w r #i Preferred variant in Korean (Hangul script): #a 공자

13

Here’s another example of a name that does not correspond to the ALA-LC romanization of the non-Latin script. The authority record for Confucius has a number of non-Latin script references in Chinese, Japanese, and Korean languages. A cataloger or multiple catalogers determine that the first 400 shown in red in the slide is the preferred form in Chinese, Japanese in Kanji script, and Korean in Hancha script. The second 400 shown in red is determined to be the preferred name in Korean in Hangul script. These references can be labeled as preferred variants.

One reminder specific to Korean: even if only a Hancha script form is found, the Task Group recommends that the Hangul script form must always be provided. It is a simple matter for Korean catalogers to convert a Hancha script form into a Hangul form. The Hangul form shown in this slide is found in many resources but a different case is shown in slide 22.

Preferred non-Latin script variant: General recommendations

- The preferred non-Latin script variant for a given language and script can function like an authorized access point for that language/script combination and will be used elsewhere in bibliographic or authority records where controlled forms are desired or appropriate.
- In most cases, follow RDA instructions for constructing authorized access points when constructing the preferred non-Latin script variant (there may be some exceptions, particularly for right-to-left languages)

Here are some general recommendations for preferred non-Latin script variants.

Preferred non-Latin script variant: General recommendations

- **Choice of the preferred non-Latin script variant is based on usage (commonly known form in resources or reference sources) in the language and script. This may or may not correspond to the form found in 1XX (authorized access point) of the NAR**

100 0# Hirohito, \$c Emperor of Japan, #d 1901-1989

400 0# #w r #i Preferred variant in Japanese (Kanji script): #a 昭和天皇, \$d 1901-1989

400 0# Shōwa Tennō, #d 1901-1989 [romanized form of preferred variant]

Authorized access point based on most common form in English, plus RDA-prescribed addition of title of royalty

Preferred variant is the most commonly used name in Japan for the emperor; the title is part of the name, and does not include a given name at all

100 1# Maimonides, Moses, #d 1135-1204

400 0# #w r #i Preferred variant in Hebrew and Yiddish (Hebrew script): #a משה בן מימון, #d 1135-1204

400 0# Mosheh ben Maimon, #d 1135-1204

400 1# #w r #i Preferred variant in Arabic (Arabic script): #a ابن ميمون، موسى، #d 1135-1204

400 1# Ibn Maymūn, Mūsá, #d 1135-1204

Authorized access point based on most common form in English (a hellenized version of the name)

Preferred variants based on most common forms in Hebrew, Yiddish, and Arabic

Preferred non-Latin script variant: General recommendations

- **Qualifiers can be recorded as found in a non-Latin form or translated into vernacular if readily available (e.g., fuller form of name, type of corporate body, type of family, other designation, associated place, form of work), or English can be used**

110 2# Jiu lian hao (Research vessel)

410 2# #w r #i Preferred variant in Chinese (Traditional Chinese script): #a 九連號 (海洋研究船)

130 #0 Sifriyat eshkolot (Haifa, Israel)

430 #0 #w r #i Preferred variant in Hebrew (Hebrew script): #a (חיפה)ספריית אשכולות

151 ## Paekche (Kingdom)

451 ## #w r #i Preferred variant in Korean (Hangul script): #a 백제 (왕국)

451 ## #w r #i Preferred variant in Korean (Hancha script): #a 百濟 (王國)

151 ## Chung-gu (Taejŏn-si, Korea)

451 ## #w r #i Preferred variant in Korean (Hangul script): #a 중구 (대전시, 한국)

110 2# Namsanhyŏn Kyohoe (P'yŏngyang, Korea)

410 2# #w r #i Preferred variant in Korean (Hangul script): #a 남산현 교회 (평양, 북한)

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

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

400 3# #w r #i Preferred name in modern Greek (Greek script): #a Ρομανώφ (Δυναστεία : #d 1613-1917)

110 2# Silsons (Firm)

410 2# #w r #i Preferred variant in Macedonian (Cyrillic script): #a Силсонс (Firm)

130 #0 Kaguyahime no monogatari (Motion picture)

430 #0 #w r #i Preferred variant in Japanese (Hiragana and Kanji script): #a かぐや姫の物語 (Motion picture)

For persons, do not use a term as a qualifier that reveals the person's gender.

Preferred non-Latin script variant: General recommendations

- **Additions (e.g., dates or other qualifiers) of non-core elements are optional and are recorded on the basis of the cataloger's judgment and community practices**

100 1# Abayov, Shim'on ben David, †d active 17th century-18th century

400 1# †w r †i Preferred variant in Hebrew (Hebrew script): †a שמעון בן דוד, אבאיוב

Cataloger chose not to add dates, as they are not needed for differentiation and present directionality issues

110 2# Firqat Jadal (Musical group)

410 2# †w r †i Preferred variant in Arabic (Arabic script): †a فرقة جدل

Name in Arabic already conveys the idea of a corporate body, so qualifier not added to variant

- **Use English form if the non-Latin script form is not readily available**

110 2# Apollo 11 (Spacecraft)

410 2# †w r †i Preferred variant in Belarusian (Cyrillic script): †a Апалон-11 (Spacecraft)

130 #0 Rhesus. †I Greek †s (Nauck)

430 #0 †w r †i Preferred variant in ancient Greek (Greek script): †a Ῥῆσος. †I Ελληνικά †s (Nauck)

500 1# †w r †i Editor: †a Nauck, August, †d 1822-1892

No Greek script form found for Nauck

Second example: (In one of its several uses), the Arabic word Firqat means musical or dance group.

Third example: if the cataloger knows the Belarusian word for “spacecraft”, they could use that as the qualifier in the preferred Belarusian variant instead of the English word. However, English qualifiers are still perfectly acceptable.

Preferred non-Latin script variant: General recommendations

- **Hierarchical corporate subordinate variant may have the parent body in its preferred non-Latin form(s)**

110 #0 Hellēniko Anoikto Panepistēmio

410 #0 #w r #i Preferred variant in modern Greek (Greek script): #a Ελληνικό Ανοικτό Πανεπιστήμιο

110 #0 Hellēniko Anoikto Panepistēmio. \$b Scholē Anthrōpistikōn Spoudōn

410 #0 #w r #i Preferred variant in modern Greek (Greek script): #a Ελληνικό Ανοικτό Πανεπιστήμιο. #b Σχολή Ανθρωπιστικών Σπουδών

- **Other hierarchical variants (e.g., series/subseries) may have all parts in their preferred non-Latin forms**

130 #0 Akademiīā`fundamental'nykh issledovanīi

430 #0 #w r #i Preferred variant in Russian (Cyrillic script): #a Академия фундаментальных исследований

130 #0 Akademiīā`fundamental'nykh issledovanīi. #p Istorīiā`

430 #0 #w r #i Preferred variant in Russian (Cyrillic script): #a Академия фундаментальных исследований. #p История

Preferred non-Latin script variant: General recommendations

➤ **Name/titles for work/expression can have the name portion in a preferred non-Latin variant form**

100 1# Maḥfūz, Najīb, #d 1911-2006

400 1# #w r #i Preferred variant in Arabic (Arabic script): #a محفوظ، نجيب، #d 1911-2006

100 1# Maḥfūz, Najīb, #d 1911-2006. #t Qushtumur

400 1# #w r #i Preferred variant in Arabic (Arabic script): #a محفوظ، نجيب، #d 1911-2006. #t قشتمر

~~~

100 1# Morrison, Toni

400 1# #w r #i Preferred variant in Japanese (Katagana script): #a モリスン, トニ

100 1# Morrison, Toni. #t Tar baby

400 1# #w r #i Preferred variant in Japanese (Katagana script): #a モリスン, トニ. #t タール・ベイビー

100 1# Morrison, Toni. #t Tar baby. #l Japanese

400 1# #w r #i Preferred variant in Japanese (Katagana and Kanji script): #a モリスン, トニ. #t タール・ベイビー. #l 日本語

Note: this decision (including the language name in ||) has not been finalized yet.

## Preferred non-Latin script variant: General recommendations

➤ **Dates: continue to use Gregorian calendar following RDA 9.19.1.1, standard Arabic numerals, and English terms**

100 1# Aḥmadzay, Gītā, #d 1994 or 1995-

400 1# #w r #i Preferred variant in Pushto (Arabic script): #a احمدزی، گیتا، #d 1994 or 1995-

**not** 400 1# #w r #i Preferred variant in Pushto (Arabic script): #a احمدزی، گیتا، #d ۱۹۹۴ او ۱۹۹۵ -

100 1# Khmel'nyts'kyi, Bohdan, #d approximately 1594-1657

400 1# #w r #i Preferred variant in Ukrainian (Cyrillic script): #a Хмельницький, Богдан, #d approximately 1594-1657

**not** 400 1# #w r #i Preferred variant in Ukrainian (Cyrillic script): #a Хмельницький, Богдан, #d приблизно 1594-1657

100 1# Tzetzes, John, #d active 12th century

400 1# #w r #i Preferred variant in modern Greek (Greek script): #a Τζέτζης, Ἰωάννης, #d active 12th century

400 1# #w r #i Preferred variant in Korean (Hangul script): #a 트제트제스, 이오안네스, #d active 12th century

**but**

100 1# Zheng, Shaowei, #d jin shi 1088

400 1# #w r #i Preferred variant in Chinese (Simplified Chinese script): #a 郑少微, #d 进士 1088

400 1# #w r #i Preferred variant in Chinese (Traditional Chinese script): #a 鄭少微, #d 進士 1088

*Since "jin shi" is a Chinese term, the Chinese script form may be used in Chinese script variants*

The "Hindu-Arabic" numeral form should *not* be used and deleted if found. Any other non-Arabic numerals should not be used either.

## Other (non-preferred) non-Latin script variants

- May be added/maintained if they are considered to be helpful
- May include qualifiers or other additions in non-Latin script or in English similar to what is available in the 1XX or other 4XX fields
- When evaluating existing records: non-Latin script variants generally should follow RDA instructions; those with errors in format, language, etc., should be deleted or enhanced based on the existing 670(s) or other sources

## Some details for all non-Latin script variants

### Mixed scripts

Forms may contain a mixture of scripts (e.g., Latin and non-Latin or two or more different non-Latin scripts). Follow found usage.

Form found (personal name): C.S. 루이스

Variant access point: 400 1# 루이스, C. S., †d 1898-1963

Form found (series title): Πολις historia

Variant access point: 430 #0 Πολις historia

Form found (corporate name): あべのハルカス美術館

Variant access point: 410 2# あべのハルカス美術館

Form found (corporate name): 코리아 민족 통일 연구소

Variant access point: 410 2# 코리아 민족 통일 연구소

Also make a variant access point all in Hangul: 410 2# 코리아 민족 통일 연구소

Color key for scripts:

Greek  
Hancha  
Hangul  
Hiragana  
Kanji  
Katagana  
Latin

22

1. Hangul and Latin script personal name = Korean form of C.S. Lewis. Last name first followed by a comma for Westerner's name
2. Greek and Latin script series (work) title
3. Japanese in Hiragana, Katagana, and Kanji scripts corporate body name
4. Korean in Hangul and Hancha scripts corporate body name. As Adam mentioned already in slide 13, even if there is no resource carrying only Hangul like this case, provide a variant in all-Hangul to help search; both can be set as preferred non-Latin references

## Some details for all non-Latin script variants

### Romanization

When romanizing non-Latin script variant access points that have parenthetical (cataloger-added) qualifiers, romanize the preferred name, but *not the qualifiers*

151 ## Macau (China : Special Administrative Region)

451 ## #w r #i Preferred variant in Chinese (Simplified Chinese script): #a  
澳门 (中国 : 特别行政区)

451 ## Aomen (China : Special Administrative Region)

**not** 451 ## Aomen (Zhongguo : Tebie Xingzhengqu)



## Some details for all non-Latin variants

### Number of conference

Continue to use English ordinal numbers in non-Latin script variants for conferences

111 2# Mizhnarodnaia navukova-praktychnaia kanferentsyia "Stylistyka, mova, maŭlenne i tэкst" #n (4th : #d 2017 : #c Minsk, Belarus)

411 2# #w r #i Preferred variant in Belarusian (Cyrillic script): #a Міжнародная навукова-практычная канферэнцыя "Стылістыка, мова, маўленне і тэкст" #n (4th : #d 2017 : #c Мінск, Беларусь)

110 2# Nihon Fujinka Gakkai. #b Sōkai #n (38th : #d 1940 : #c Tokyo, Japan)

410 2# #w r #i Preferred variant in Japanese (Kanji script): #a 日本婦人科學會.  
#b 總會 #n (38th : #d 1940 : #c 東京, 日本)

We are proposing to continue using English ordinal numbers in all access points for conferences, including non-Latin script variant access points.

## Challenges, or, “Why is this taking so long?”

- Inconsistency of practice across non-Latin communities and cultures
- Variety and extent of categories and situations that the TG needs to address: all the entities and all the language/script practices
- Complexity of certain issues related to languages/scripts (bi-directionality issue, for example)
- Need to compile comprehensive guidelines including examples

The magnitude of the work simply requires a lot of time...

26

The TG is proposing these guidelines that are the result of extensive discussion and testing, and yet this comes with many challenges, some of which are:

- Inconsistency of practice across non-Latin communities and cultures.
- The variety and extent of categories and situations that the TG needs to address across all the entities and all the language/script practices
- Complexity of certain issues related to languages/scripts (bi-directionality issue, for example)
- The need to compile comprehensive guidelines including examples

The magnitude of the work simply requires a lot of time... Time consuming: weekly two hour meetings in addition to off-line work to examine the various situations and discuss the possible implications of each question and recommendation

## Survey: What is expected from the NACO community?

We need practitioners' feedback on:

- General recommendations
- Some details on non-Latin script variants with practices specific to a specific non-Latin community
- Any other questions, concerns, use cases not covered, etc.

Survey will be sent shortly after ALA. Input will be considered and included in TG report.

27

Community feedback will be crucial to our proceeding with the recommendations. Therefore we will be sending a survey soon (hopefully after ALA).

We ask you to please watch for the survey and provide your thoughts on these recommendations!

## Next steps

- Evaluate survey results
- Resolve outstanding questions (e.g., mandatory requirements vs. optional, language-script specific guidelines)
- Send recommendations to PCC Standing Committee on Standards (SCS)
- Pursue complete NACO node validation changes and further MARC development
- Test guidelines, noting the dependence on NACO node implementation
- Draft documentation with examples to demonstrate various recommendations
- Receive SCS and PCC Policy Committee approval
- Create training for implementation of the new guidelines
- Determine implementation date (2025?)

28

Our work is not yet complete.

We need to evaluate the survey results and try to incorporate them into the proposed guidelines

We still have some questions that need to be addressed. For example, one of the main questions is whether designating a preferred non-Latin script variant should be optional or required.

Once our charge is fulfilled, we will send our recommendations to the PCC Standing Committee on Standards; then it will be shared with PoCo for approval.

Also, the recommendations require NACO node validation and/or MARC development.

The TG will also need to provide documentation with examples to demonstrate various recommendations, as well as provide training for using new guidelines to the various communities since some languages have some special considerations.

We hope that this will be implemented by next year!

And that concludes our presentation for today.



# Thank you!

Presentation will be linked to the TG wiki page  
<https://wiki.lyrasis.org/display/PFCCP/SCS+Task+Group+on+Evaluation+Guidelines+for+Non-Latin+Script+References+in+NARs>