

Appendix C: ScriptShifter

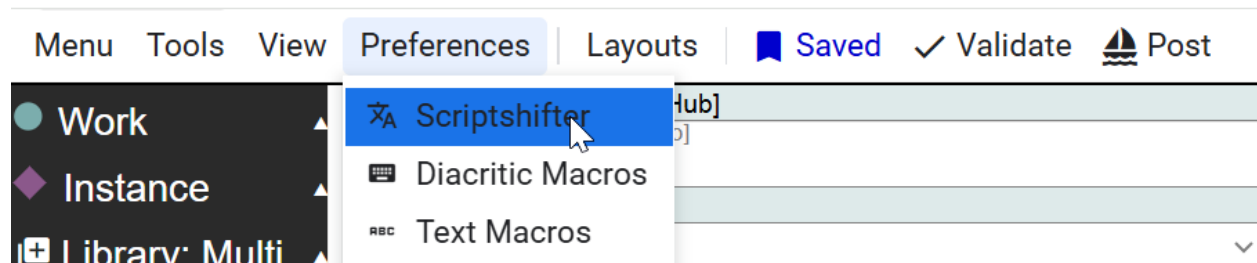
About ScriptShifter

ScriptShifter is an open source tool that is used in Marva to romanize non-Latin script according to the ALA/LC Romanization Tables, and to generate non-Latin scripts from data romanized according to the ALA/LC Romanization Tables.

ScriptShifter uses AI and user input to improve its responsiveness. Anyone working with ScriptShifter in Marva is welcome to provide feedback to the developers on ScriptShifter/

Setting ScriptShifter Scripts and Script Options

Access ScriptShifter under Preferences in the Navigation Bar in Marva.



In the pop-up box, you will see multiple options for selecting Scripts and Actions.

X

Visit bibframe.org/scriptshifter to test these languages.

☐ Capitalize first letter of transliteration




Name	Script to Roman	Roman to Script
Abkhaz (Cyrillic)	<input type="checkbox"/>	<input type="checkbox"/>
Altai (Cyrillic)	<input type="checkbox"/>	<input type="checkbox"/>
Arabic	<input type="checkbox"/>	<input type="checkbox"/>
Arabic-to-Roman transliterator using the ArabicTransliterator external library.		
Armenian	<input type="checkbox"/>	<input type="checkbox"/>
Asian Cyrillic		
Multi-purpose transliteration for non-Slavic Cyrillic scripts: Abaza, Abkhaz, Adygei, Aisor, Altai, Avar, Azeri, Balkar, Bashkir, Buryat, Chechen, Chukchi, Chuvash, Dargwa, Dungan, Eskimo, Even, Evenki, Gagauz, Ingush, Inuit, Kabardian, Kalmyk, Karachay, Karachay-Balkar, Karakalpak, Karelian, Khakass, Khanty, Komi, Komi-Permyak, Koryak, Kumyk, Lak, Lapp, Lezghian, Lithuanian, Mansi, Mari, Moldovan, Molodstov, Mordvin, Nanai, Nenets, Nivkh, Nogai, Ossetic, Permyak, Romanian, Romany, Selkup, Shor, Tabasaran, Tat, Tuva, Udekhe, Udmurt, Yakut.		
assamese	<input type="checkbox"/>	<input type="checkbox"/>
Azerbaijani (Cyrillic)	<input type="checkbox"/>	<input type="checkbox"/>

Capitalize first letter of transliteration is “off” by default. If the script and language you are working with uses capitalization, you can click on this box so ScriptShifter will capitalize the first letter of the text that is generated.

In the Name, Script to Roman, and Roman to Script columns, you can select the languages and scripts that you will be working with in Marva from the A-Z list.

For most languages and scripts, ScriptShifter offers both transliteration from non-Latin script to romanized script, and the generation of non-Latin script from romanized script. Check the appropriate boxes. There is no limit on the number of boxes you can check.


For some languages and scripts, ScriptShifter does not offer both transliteration from non-Latin script to romanized script, and the generation of non-Latin script from romanized script. For these scripts, one of the boxes will be greyed-out, meaning that the option is not available in ScriptShifter. For example, ScriptShifter does not have an option to generate Korean script from romanized Korean script, and ScriptShifter does not have an option to romanize Kurdish script.

Korean (last + first names only)	<input type="checkbox"/>		<input type="checkbox"/>
Korean S2R for strings ONLY containing personal names formatted as last + first name. Separate multiple names with a comma or a center-dot (U+00B7).			
Korean	<input type="checkbox"/>		<input type="checkbox"/>
Korean S2R for strings NOT containing any personal names.			
Koryak (Cyrillic)	<input type="checkbox"/>		<input type="checkbox"/>
Kurdish		<input type="checkbox"/>	<input type="checkbox"/>
Kyrgyz (Cyrillic)	<input type="checkbox"/>		<input type="checkbox"/>

Using ScriptShifter in Marva

Script to Roman

Set the value in ScriptShifter. Click the Esc button to exit the ScriptShifter panel.

Rajasthani (Devanagari)	<input type="checkbox"/>	<input type="checkbox"/>
Romani (Cyrillic)	<input type="checkbox"/>	<input type="checkbox"/>
Russian		<input checked="" type="checkbox"/>
Sanskrit (Devanagari)	<input type="checkbox"/>	<input type="checkbox"/>
Serbian	<input type="checkbox"/>	<input type="checkbox"/>

In Marva, the options that you selected in the ScriptShifter panel will be visible in the Lightning Bolt feature for each literal component.

Title information	
Work title	
Non-sort character count	0
Preferred title for Work	
Part number	1 Add Another Component
Part name	3 Delete Component
	4 Send to Instance
Contributors	
Person	5 Additional Literal
Search LCNAP	6 Set Language
Relationship designator(s)	7 → ⑈ Russian S2R
	8 ⑈ Russian R2S
Notes about the Work	
Note	0 Debug
Note type	Add To Library
Contents	
Contents note	- View Documentation

<marc:sul

In a Literal field, key in or paste Russian script. If you are using the MARC Preview Panel, you will see that the Russian title in Cyrillic script shows up in the MARC 245 field.

Title information	
Work title	
Non-sort character count	0
Preferred title for Work	
Русские объединяются с крокодилами и нападают на Европу	⚡
Part number	
Part name	

```

008 250318|||||||xx 000 0
245 00 $a Русские объединяются с крокодилами и нападают на Европу
758 $4 http://id.loc.gov/ontologies/bibframe/instanceOf $1 http://id...
758 $1 http://id...
884 $a DLC bibframe2marc v2.10-dev (libxslt) $g 20250318 $q DLC
$u https://github.com/lcnetdev/bibframe2marc/releases/tag/v2.10-dev
  
```

Click on the Lightning Bolt and select the **Russian S2R** option.

Title information	
Work title	008 250318
Non-sort character count	245 00 \$a P
0	крокодилам
Preferred title for Work	758 \$4
Русские объединяются с крокодилами и нападают на Европу	http://id.loc.g
Part number	1 Add Another Component
Part name	3 Delete Component
	4 Send to Instance
Contributors	5 Additional Literal
Person	6 Set Language
Search LCNAF	7 ⌘ Russian S2R
Relationship designator(s)	8 ⌘ Russian R2S
Notes about the Work	0 Debug
Note	Add To Library
Note type	- View Documentation
Contents	
Contents note	

The romanized title will be added to a second Preferred title for Work in the Marva Title information component, and the values in the MARC Preview will reflect current cataloging policy, with the romanized title in the 245 field, and the Russian Cyrillic script title in the corresponding 880 field. The non-Latin title and the romanized title will be connected with a bracket in Marva.

Title information	
Work title	007 t
Non-sort character count	008 250407 xx 000 0
0	066 \$c (N
Preferred title for Work	245 00 \$6 880-02 \$a Russkie ob'edini' a i u tsi' a s krokodil
Русские объединяются с крокодилами и нападают на Европу	880 00 \$6 245-02/(N \$a Русские объединяются с крокодил
Европу	758 \$4 http://id.loc.gov/ontologies/bibframe/instanceOf \$1 t
Preferred title for Work	758 \$1 http://id.loc.gov/resources/instances/e0975249
Russkie ob'ediniiautsia s krokodilami i napadaiut na Evropu	884 \$a DLC bibframe2marc v2.10-dev (libxslt) \$g 2025040
Part number	

Roman to Script

Set the value in ScriptShifter. Click the Esc button to exit the ScriptShifter panel.

Rajasthani (Devanagari)	<input type="checkbox"/>	<input type="checkbox"/>
Romani (Cyrillic)	<input type="checkbox"/>	<input type="checkbox"/>
Russian	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Sanskrit (Devanagari)	<input type="checkbox"/>	<input type="checkbox"/>
Serbian	<input type="checkbox"/>	<input type="checkbox"/>

In Marva, the options that you selected in the ScriptShifter panel will be visible in the Lightning Bolt feature for each literal component.

Title information	
Work title	
Non-sort character count	0
Preferred title for Work	
Part number	1 Add Another Component
Part name	3 Delete Component
	4 Send to Instance
Contributors	
Person	5 Additional Literal
Search LCNAF	6 Set Language
Relationship designator(s)	7 ⌘ Russian S2R
	8 ⌘ Russian R2S
Notes about the Work	
Note	0 Debug
Note type	Add To Library
Contents	
Contents note	- View Documentation

<marc:su

In a Literal field, key in or paste romanized Russian script. If you are using the MARC Preview Panel, you will see that the romanized Russian title in Cyrillic script shows up in the MARC 245 field.

Title information	
Work title	
Non-sort character count	0
Preferred title for Work	Russkie Russkie ob"ediniatsia s krokodilami i napadaiut na Evropu
Part number	

008 250318|||||||xx 000 0
 245 00 \$a Russkie Russkie ob"edini`a`i`u`tsi`a` s krokodilami i napadai`u`t na Evropu
 758 \$4 http://id.loc.gov/ontologies/bibframe/instanceOf \$1 http://id...
 758 \$1 http://id...
 884 \$a DLC bibframe2marc v2.10-dev (libxslt) \$g 20250318 \$q DLC \$u https://github.com/lcnetdev/bibframe2marc/releases/tag/v2.10-dev

Click on the Lightning Bolt and select the **Russian R2S** option.

Title information	
Work title	
Non-sort character count	0
Preferred title for Work	Russkie Russkie ob"ediniatsia s krokodilami i napadaiut na Evropu
Part number	1 Add Another Component
Part name	3 Delete Component
	4 Send to Instance
Contributors	
Person	5 Additional Literal
Search LCNAF	6 Set Language
Relationship designator(s)	7 ⌘ Russian S2R
	8 ⌘ Russian R2S
Notes about the Work	
Note	0 Debug
Note type	Add To Library
Contents	- View Documentation

008 250318|||||||xx 000 0
 245 00 \$a Russkie Russkie ob"edini`a`i`u`tsi`a` s krokodilami i napadai`u`t na Evropu
 758 \$4 http://id.loc.gov/ontologies/bibframe/instanceOf \$1 http://id...
 758 \$1 http://id...
 884 \$a DLC bibframe2marc v2.10-dev (libxslt) \$g 20250318 \$q DLC \$u https://github.com/lcnetdev/bibframe2marc/releases/tag/v2.10-dev
 n="1.0" encoding="UTF-8"?>
 d xmlns:marc="http://www.loc.gov/MARC21/slim">
 er xml:space="preserve"> cam a22 i 4500</marc:leader>
 rolfield xml:space="preserve" tag="001">id1</marc:controlfield>
 rolfield xml:space="preserve" tag="003">DLC</marc:controlfield>
 rolfield xml:space="preserve" tag="005">20250318153746.0</marc:controlfield>
 rolfield xml:space="preserve" tag="007">t</marc:controlfield>
 rolfield xml:space="preserve" tag="008">250318|||||||xx 000 0</marc:controlfield>
 field tag="245" ind1="0" ind2="0">245 00 \$a Russkie Russkie ob"edini`a`i`u`tsi`a` s krokodilami i napadai`u`t na Evropu</marc:controlfield>
 field code="a">Russkie Russkie ob"edini`a`i`u`tsi`a` s krokodilami i napadai`u`t na Evropu</marc:controlfield>
 field tag="758" ind1="0" ind2="0">758 \$4 http://id.loc.gov/ontologies/bibframe/instanceOf \$1 http://id...</marc:controlfield>
 field tag="758" ind1="0" ind2="0">758 \$1 http://id...</marc:controlfield>
 field tag="884" ind1="0" ind2="0">884 \$a DLC bibframe2marc v2.10-dev (libxslt) \$g 20250318 \$q DLC \$u https://github.com/lcnetdev/bibframe2marc/releases/tag/v2.10-dev</marc:controlfield>
 </marc:datafield tag="758" ind1="0" ind2="0">

The Russian script title will be added to a second Preferred title for Work in the Marva Title information component, and the values in the MARC Preview will reflect current cataloging policy, with the romanized title in the 245 field, and the Russian Cyrillic script title in the corresponding 880 field. The non-Latin title and the romanized title will be connected with a bracket in Marva.

The screenshot shows the Marva interface with two components. The first component, 'Title information', has a 'Work title' field with the text 'Russkie ob"ediniäiutsia s krokodilami i napadaiut na Evropu' and a 'ru-Latn' tag. The second component, 'Preferred title for Work', has the text 'Русские объединяются с крокодилами и нападают на Европу' and a 'ru-Cyrl' tag. The MARC preview on the right shows the following fields:

007	t
008	250407 xx 000 0
066	\$c (N
245	00 \$6 880-02 \$a Russkie ob"ediniäiutsia s krokodilami i napadaiut na Evropu
880	00 \$6 245-02/(N \$a Русские объединяются с крокодилами и нападают на Европу
758	\$4 http://id.loc.gov/ontologies/bibframe/instanceOf \$1
758	\$1 http://id.loc.gov/resources/instances/e0975249
884	\$a DLC bibframe2marc v2.10-dev (libxslt) \$g 2025040

Language and Script Tags

Component Level

When ScriptShifter is used in Marva, language and script tags are added by default to both romanized text and text in Non-Latin script.

The screenshot shows the Marva interface with two components. The first component, 'Title information', has a 'Work title' field with the text 'Russkie ob"ediniäiutsia s krokodilami i napadaiut na Evropu' and a 'ru-Latn' tag. The second component, 'Preferred title for Work', has the text 'Русские объединяются с крокодилами и нападают на Европу' and a 'ru-Cyrl' tag. The MARC preview on the right shows the following fields:

007	t
008	250407 xx 000 0
066	\$c (N
245	00 \$6 880-02 \$a Russkie ob"ediniäiutsia s krokodilami i napadaiut na Evropu
880	00 \$6 245-02/(N \$a Русские объединяются с крокодилами и нападают на Европу
758	\$4 http://id.loc.gov/ontologies/bibframe/instanceOf \$1
758	\$1 http://id.loc.gov/resources/instances/e0975249
884	\$a DLC bibframe2marc v2.10-dev (libxslt) \$g 2025040

These values can be changed if needed by clicking on the Lightning Bolt and selecting Set Language.

The screenshot shows the Marva interface with the 'Title information' component. The 'Work title' field has the text 'Russkie ob"ediniäiutsia s krokodilami i napadaiut na Evropu' and a 'ru-Latn' tag. A dropdown menu is open, showing the following options:

- 1 Add Another Component
- 2 Delete Component
- 3 Send to Instance
- 4 Additional Literal
- 5 Set Language
- 6 Russian S2R
- 7 Russian R2S
- 8 Thai S2R
- 9

Russkie Russkie ob"edini~a~i~u~tsi~a~ s krokodilami i napadai~u~t na Evropu@ru-Latn

Russian

Latin

Clear Language & Script

Русские Русские объединяются с крокодилами и нападают на Европу@ru-Cyrl

Russian

Cyrillic

Clear Language & Script

Done

Under each literal value, there will be a Language tag and a Script tag. Use the dropdown lists under Language and Script to change the values.

Russkie Russkie ob"edini~a~i~u~tsi~a~ s krokodilami i napadai~u~t na Evropu@ru-Latn

Russian

Latin

Clear Language & Script

Русские Русские объединяются с крокодилами и нападают на Европу@ru-Cyrl

Russian

Cyrillic

Clear Language & Script

Script

Language

Global Level

You can select Language and Script tags for all literal values in a description by clicking on Tools à Non-Latin Literals in the Navigation Bar.

Below are non-Latin literals in the record, you can quickly assign a language and script to them with this list. To add a new option expand "Add Language Option"

► Add Language Option

Non-Latin literals in the record:

Set all literals below to: None

http://id.loc.gov/ontologies/bibframe/mainTitle

Russkie Russkie ob"edini~a~i~u~tsi~a~ s krokodilami i napadai~u~t na Evropu@ru-Latn

Set individually: None

http://id.loc.gov/ontologies/bibframe/mainTitle

Русские Русские объединяются с крокодилами и нападают на Европу@ru-Cyrl

Set individually: None

Menu Tools View Preference

Wor

Inst

Libr

Libr

Defa

Shelf Listing Browser

AutoDewey

Non-Latin Literals

Non-Latin Agents

Copy Mode [off]

Paste Content

Add All Defaults

E

C

F

S

R

Ti

V

N

C

P

R

P

This will open a box that will show all non-Latin Literal values in the description and you can manage all of them in one screen.

Non-Latin Agents

Menu Tools View Prefere

Wor

Inst

Libr

Libr

Defa

Shelf Listing Browser

AutoDewey

Non-Latin Literals

Non-Latin Agents

Copy Mode [off]

Paste Content

Add All Defaults

2025-05-09

[Back to Table of Contents](#)

Appendix C
ScriptShifter
8

Non-Latin Agents (Tools → Non-Latin Agents) allows you to manage access points in Non-Latin scripts. This feature works with non-Latin variant access points in authority records.

When a name authority has multiple Non-Latin authorized labels. For example: [n2010057779](http://id.loc.gov/authorities/names/n2010057779) :



(id.loc.gov is “selecting” a Non-Latin variant label to be the “authorized” label for that script)

When added to Marva the editor will try to select which Non-Latin authorized label to use to build an 880 field. It does this by looking at the language tags set for the literals in the record. It selects the most common one and then chooses that language + script combination to build the 880. This works fine if there is only one Non-Latin authorized label to select. But if there are multiple it could potentially select the incorrect one to use. This modal window allows you to overwrite that logic and force Marva to use a specific script to build the 880. After I added this name authority, and a Non-Latin title, and went to Tools -> Non-Latin Agents I see this:

X

The system will build Access Points using values found in the script: **Kore**

You can override which script to use to build specific access points below:

bf.contribution
Hyŏn, Mu-am, 1969-
Use script to build access points: Hani ▾

For this name it has selected the Hani script to build the 880. I can force it however to use the Hang script instead:

X

The system will build Access Points using values found in the script: **Kore**

You can override which script to use to build specific access points below:

bf.contribution
Hyŏn, Mu-am, 1969-
Use script to build access points: Hani ▾

Hani

Hang

This will force the conversion to use the Hang script label when building the 880 in the MARC output. To summarize you will likely never need to modify these selections as Marva should pick the best fit. But if it doesn't this is how you can overwrite it.

[Back to Table of Contents](#)
