

Проект по оцифровке финно-угорских материалов в рамках Языковой программы Фонда Конне

Jussi-Pekka Hakkarainen

Руководитель проекта

Сотрудник исследовательской библиотеки

Национальной библиотеки Финляндии

10-го апреля 2013 в ЦУНБ им. Н.А. Некрасова



ЦЕЛЬ И ЗАДАЧИ ЯЗЫКОВОЙ ПРОГРАММЫ

Целью **Языковой программы Фонда Коне** является документирование материалов по языкам малочисленных финно-угорских народов, финского языка и языков национальных меньшинств Финляндии, а также упрочение их положения.

Посредством **Языковой программы Фонда Коне** осуществляется научно значимая исследовательская работа. Как научная среда, так и носители языка смогут пользоваться результатами этих исследований и корпусом накопленных языковых данных, и такое взаимодействие приведет к совершенствованию исследовательских методик.

ПИЛОТНЫЙ ПРОЕКТ ПО ОЦИФРОВКЕ 2012-2013 ГГ.

В рамках Языковой программы Фонда Конне на 2012-2016 годы данный проект получил название **Проект по оцифровке родственных языков** (финно-угорских), пилотная фаза 2012-2013 гг.

Национальная библиотека Финляндии (НБФ, Хельсинки) разрабатывает исследовательские материалы и инфраструктуру, необходимые для участия в языковой программе Фонда.

Национальная библиотека несет ответственность за сотрудничество с **Российской национальной библиотекой** (РНБ, Санкт-Петербург) и **Национальным библиотечным ресурсом** (НБР, Москва).

ПИЛОТНЫЙ ПРОЕКТ ПО ОЦИФРОВКЕ 2012-2013 ГГ.

Материалы, которые будут оцифрованы в рамках пилотного проекта, станут доступны в системе управления электронной библиотекой (СУЭБ), поддерживаемой НБФ и основывающейся на открытом исходном коде.

Цифровые материалы станут доступны всем гражданам, а научному сообществу будет предоставлена возможность обращаться к этим материалам в исследовательских целях.

Кроме того, в процессе размещения материалов будут созданы соответствующие инструменты и условия для организации краудсорсинга.



МАТЕРИАЛЫ

Осенью 2011 г. представителями Фонда Конне определены материалы, входящие в коллекции РНБ, которые будут оцифрованы в рамках пилотного проекта.

Эти материалы будут подготовлены в РНБ и предоставлены НБФ в цифровом формате.

Материалы содержат 17 000 страниц изданий на ижорском, вепсском и мордовских языках, в общей сложности 128 монографий. Помимо монографий, будет оцифровано почти 20 000 страниц текста марийских и мордовских газет, опубликованных в основном в 1920-ые и 1930-ые годы.



АВТОРСКИЕ ПРАВА

В пилотном проекте необходимо принять во внимание как российское, так и финляндское законодательство об авторских правах, поскольку материалы будут оцифрованы **в России**, а размещены **в Финляндии**.

Проведенная работа дает все основания полагать, что физических и юридических лиц, которые должны образом могли бы заявить о своих притязаниях на авторские права на произведения, вошедшие в Проект, нет. Следовательно, авторские права на такие произведения с высокой степенью вероятности можно отнести к выморочному имуществу.



АВТОРСКИЕ ПРАВА

Переход произведения в общественное достояние означает, что такое произведение в силу **ст.1282 ГК РФ** может свободно использоваться любым лицом без чьего-либо согласия или разрешения и без выплаты авторского вознаграждения. При этом охраняются авторство, имя автора и неприкосновенность произведения.

Благодаря работе, проведенной НБР, удалось документально обосновать факт принадлежности авторских прав на издания к выморочному имуществу, в отношении которого действует порядок использования произведений, перешедших в общественное достояние.



СИСТЕМЫ И ИНТЕРФЕЙСЫ

Для пилотного проекта будет создано обособленное от остальных цифровых фондов цифровое хранилище **DSpace** (среда СУЭБ), которое позволит удовлетворить потребности как профессионального научного сообщества так и обычных пользователей.

Dspace (коллекция **Фенно-утрика**) расширит доступность оцифрованных материалов и их использование за рубежом и в России.

Интерфейс Фенно-утрика будет реализовано в том числе и на русском языке. Возможно от будет доступен и на других языках.

СИСТЕМЫ И ИНТЕРФЕЙСЫ



FENNO-UGRICA

[Suomeksi](#) [In English](#) [По-русски](#)

[Главная страница](#) > [Разделы и коллекции](#)

[Поиск](#) [Справка](#)

Фенно-угрика

Фенно-угрика – оцифрованная финно-угорская коллекция Национальной библиотеки Финляндии. Коллекция содержит публикации на ижорском, вепсском, марийских (горномарийский и луговомарийский) и мордовских (Эрзянский и Мокшанский) языках, а также газеты на марийских и мордовских языках, опубликованных в основном в 1920-ые и 1930-ые годы. В целом коллекция содержит 128 монографий и почти 17 000 страниц текста газет.

Представленные здесь электронные ресурсы созданы в рамках [Проекта по оцифровке публикаций на финно-угорских языках](#) из фондов [Российской национальной библиотеки](#) в Санкт-Петербурге. Проект является частью [Языковой программы](#) Фонда Конне. Материалы оцифрованы в Российской национальной библиотеке и публикуются с разрешения [Национального библиотечного ресурса](#) (Москва).

В рамках проекта для лингвистов также разработан OCR-редактор, основанный на открытом исходном коде, который позволяет редактировать OCR-тексты на финно-угорских языках. Право редактировать ресурсы данной коллекции предоставляется, прежде всего, финноугроведам. Право пользования и редактирования можно получить у администрации Проекта по оцифровке финно-угорских языков. Дополнительная информация и контакты по электронному адресу: fennougrica@helsinki.fi

Разделы и коллекции

- [Monografica](#) [15]
- [Periodica](#) [1]

Весь архив

- [Заглавия](#)
- [Авторы](#)
- [Даты публикации](#)
- [Темы](#)
- [Свежие поступления](#)
- [Просмотр по языкам](#)
- [Карта сайта](#)

Мой профиль

- [Войти](#)
- [Зарегистрироваться](#)

[KONEEN SÄÄTIÖ](#)

[Фенно-угрика](#) – оцифрованная финно-угорская коллекция НБФ.

СИСТЕМЫ И ИНТЕРФЕЙСЫ



Главная страница > Monografica > Monografiat > Просмотр ресурса

[Поиск](#) [Справка](#)

Поиск по коллекции Поиск в архиве

[Показать краткое описание ресурса](#)

dc.contributor.author	Попова, Наталья Сергеевна	
dc.date.accessioned	2013-02-04T14:31:25Z	
dc.date.available	2013-02-04T14:31:25Z	
dc.date.issued	2013-02-04	
dc.identifier.other	RU_NLR_ONL_14450-1	
dc.identifier.uri	http://smurffi-kk.lib.helsinki.fi/handle/10024/61212	
dc.description	3-це нолдавк витнезь	
dc.format.extent	74, [2] s., kuv.	
dc.language	fi=Ersäjen=Erzyan ru=Эрзянский язык	en
dc.language.iso	myv	
dc.source	Москов : Учпедгиз, 1935	
dc.title	Арифметика : васьень школань 1 классо тонавтнема книга : 1 пельск	en
dc.subject.ysa	oppikirjat	fi
dc.subject.ysa	aritmetiikka	fi
dc.subject.ysa	ersän kieli	fi
dc.subject.ru	эрзянский язык	ru
dc.rights.management	Moskovaalainen kirjastoalan tekijänoikeusjärjestö National Library Resource (Natsionalnyi bibliotetšnyi resurs / Национальный библиотечный ресурс;)	fi

При каталогизации библиографические метаданные будут перерабатываться так, чтобы они соответствовали определениям для описания цифровых материалов формата **Dublin Core**.

Доступность материалов будет обеспечиваться возможностью поиска через интерфейсы Национальной цифровой библиотеки [Finna](#), и, возможно, через интерфейсы европейской цифровой библиотеки [Europeana](#) в той мере, в какой это позволяют авторские права.

СИСТЕМЫ И ИНТЕРФЕЙСЫ

bx000010800.pdf



erilajaiset, niitä eri ajoil eri viisii arvattii ja senen mukkaa keelees syntyi eri formia, kummat enemmän tali vähemmän tarkast näyttää keelellisil sredstvoil näitä erilajaisia otno-senjoja.

Näitä määrättyihe kysymysii vastaavia ja määrättyjä predmettoin välisiä otno-senjoja näyttäviä formia kutsutaa painutossihoiks tali **prosto** padezoiks.

Joka painutossiihal (padezaal) ono oma määrätty loppu, ja omat määrättyt merkitökset.

Eri keeliis painutossihoiin luku ono erilain. Toisiis keeliis ono niitä paljo, toisiis ono vähemmän, a ono i mokomia, kummiis neet veel evät oo i syntyneet. Joka keeleel ollaa omat sredstvat arvata ympäröittävää maailmaa, joka keeleel ono ommia omintuksia, vaik keelet kehityksessää i pittiijää yhteisiä zakonoja.

Vennäen keelees, niku tiijettä, painutossihoja ei oo kovin paljo. Vennäen keelt vart ono 6 padezaa (painutossihaa). Neet padezat ollaa, kuin tiijettä, именительный, родительный, дательный, винительный, творительный, предложный.

Neet sihat ollaa i izaran keelees. Mut izaran keelees ono veel paljo muita painutossihoja, mitä vennäen keelees ei oo ja vennäen keelen painutossihoilkiä ollaa izaran keelees veel i toiset merkitökset.

Ylempään oli saottu jo, jot otno-senjoja ono erilajaisia ja jot senen mukkaa i niitä näyttäviä painutossihoja ono paljo. Kaik painutossihat sen peräst möö voimma jakkaa peenii gruppii, ja joka mokoma gruppa näyttää otno-senjoja, kummat keskenää ollaa sihoilu, ollaa likimäisiä. Kaikkus hyväst painutossihoiin merkitös näkky silloin, ku joka painutossi-halle annetaa kysymys, kuhu painutossiha vastajaa. Painutossihoiin nimet möö otamma internatsionalnoist terminologiast, kumpaa pittiissä lingvistises literaturas (literaturas, kumpa tutkii keeltä). Literaturaa neet terminat tali sanat joutuit vanhast latinan keelest, kumpa satoja voosia oli kansainväliseen, internatsionalnoin tiitokeeleen. Katsomma, mitä painutossihoja ono izaran keelees ja mil viisii neet grupittuut. Joka painutossihalle paatnma kysymyksen, mille se vastajaa.

Nominativa. Ken? Mikä? Ket? Mit?
Nominativa näyttää predmetan, rissan nimmiä eikä mit-tää otno-senjoia, mis predmetta ono toisiil predmettoihe. Sen-

erilajaiset, niitä eri ajoil eri viisii arvattii ja senen mukkaa keelees syntyi eri formia, kummat enemmän tali vähemmän tarkast näyttää keelellisil sredstvoil näitä erilajaisia **otno-senjoja**.

Näitä määrättyihe k y s y m y k s 11 vastaavia ja määrättyjä predmettoin välisiä **otno-senjoja** näyttäviä formia kutsutaa painutossihoiks tali **prosto** padezoiks.

Joka painutossiihal (padezaal) ono oma määrätty loppu, ja omat määrättyt merkitökset.

Eri keeliis painutossihoiin luku ono erilain. **Toisiis** keeliis ono niitä paljo, toisiis ono vähemmän, a ono i mokomia, kummiis neet veel evät oo i syntyneet. Joka keeleel ollaa omat sredstvat arvata ympäröittävää maailmaa, joka keeleel ono ommia omintuksia, vaik keelet kehityksessää i pittiijää **yhteisiä zakonoja**.

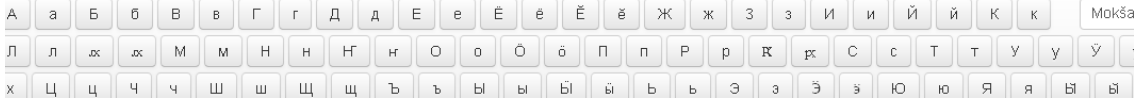
Vennäen keelees, niku tiijettä, painutossihoja ei oo ko-vin paljo. Vennäen keelt vart ono 6 **padezaa** (painutossihaa). Neet padezat ollaa, kuin tiijettä, именительный, родительный, дательный, винительный, творительный, предложный.

Neet sihat ollaa i izaran keelees. Mut izaran keelees ono veel paljo muita painutossihoja, mitä vennäen keelees ei oo ja vennäen keelen painutossihoilkiä ollaa izaran keelees veel i toiset merkitökset.

Ylempään oli saottu jo, jot otno-senjoja ono erilajaisia ja jot senen mukkaa i niitä näyttäviä painutossihoja ono paljo. Kaik painutossihat sen peräst möö voimma jakkaa peenii gruppii, ja joka mokoma gruppa näyttää otno-senjoja, kummat keskenää ollaa sihoilu, ollaa likimäisiä. Kaikkus hyväst painutossihoiin merkitös näkky silloin, ku joka painutossi-halle annetaa kysymys, kuhu painutossiha vastajaa. Painutossihoiin nimet möö otamma internatsionalnoist terminologiast, kumpaa pittiissä lingvistises literaturas (literatures, kumpa tutkii keeltä). Literaturaa neet terminat tali sanat joutuit vanhast latinan keelest, kumpa satoja **voosia oli kansainväliseen** internatsionalnoin tiitokeeleen. **Katsomma**, mitä painutossihoja ono izaran keelees ja mil viisu neet grupittuut. Joka painutossihalle paatnma kysymyksen, mille se vastajaa.

Nominativa. Ken? Mikä? Ket? Mit?

Nominativa näyttää predmetan, nissan nimmiä eikä mit-tää otno-senjoia, mis predmetta ono **toisu predmettoihe**. Sen-



В рамках проекта для лингвистов также будет разработан **OCR-редактор**, основанный на открытом исходном коде, который позволит редактировать OCR-тексты на финно-угорских языках. Право редактировать ресурсы данной коллекции будет предоставлено, прежде всего, финноугроведам.

Разработка этого инструмента будет проводиться НБФ в сотрудничестве с исследователями.





ПРОДОЛЖЕНИЕ ПРОЕКТА

Во время осуществления пилотного проекта будут разработаны способы и техническая реализация оцифровки других материалов, что будет необходимо при осуществлении следующего проекта в рамках языковой программы в 2014-2016 гг.

Целью проекта-продолжения станет как создание новых условий для международного сотрудничества, так и расширение, по мере возможностей, существующей кооперации на другие российские и финские институты коллективной памяти, хранящие и переводящие в цифровую форму различные финно-угорские языковые материалы.



КОНТАКТНЫЕ ДАННЫЕ И ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Проект по оцифровке родственных (финно-угорских) языков
www.nationallibrary.fi/services/digitaalisetkokoelmat/finnougric_en_ru.html

Коллекция Фенно-угрика
fennougrica.kansalliskirjasto.fi/

Фонд Конне
www.koneensaatio.fi/en/

Национальная библиотека Финляндии
www.nationallibrary.fi/



КОНТАКТНЫЕ ДАННЫЕ И ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Jussi-Pekka Hakkarainen

Руководитель проекта,

Сотрудник исследовательской библиотеки

Национальной библиотеки Финляндии

The National Library of Finland

P.O. Box 26 (Teollisuuskatu 23)

00014 University of Helsinki

+358 9 191 40793

+358 50 363 9223

jussi-pekka.hakkarainen@helsinki.fi