

# Impulsreferat (M. Rießler)



documentary linguistics | descriptive linguistics  
| theoretical linguistics | computer linguistics |  
archives | corpora | language endangerment |  
revitalization | ...

Obs! Linguistics is a science but not a means for  
revitalizing languages

- A linguistic sub-discipline
  - Own subject: language documentation
  - Own methods, theories, scientific journal, etc.
- What is a language documentation?
  - A long lasting, available, multi-faceted archive of spoken language
  - It ideally includes annotated audio and/or video recordings from the widest range of speech genres as well as a sketch grammar and a dictionary explaining the used annotation conventions

# Documentary linguistics



- “Traditional” annotated text editions are not language documentation!
- “Traditional” descriptive grammarology is not language documentation!
- Theoretical descriptions of linguistic structure (empirically based or not) are not language documentation!

The presented analyses (however plausible they might be) cannot be repeated and proven on basis of original data.

# Archives for endangered languages



- “Traditional” text editions, “traditional” descriptive grammars and theoretical (empirically based!) descriptions can still be useful for endangered language maintenance and revitalization.
- Linguists working with endangered languages have the ethical responsibility for preserving and making available their original data.
- Archiving original data and annotations seems to be the most appropriate way for achieving this goal.

# Archives for endangered languages



## What makes an archive?

- Original field notes, DVDs, DV-tapes stored in a safe?
- Recordings and annotations stored in file directories on hard drives or university servers?
- Web-based media presentations, on-line text editions and dictionaries?

# Archives for endangered languages

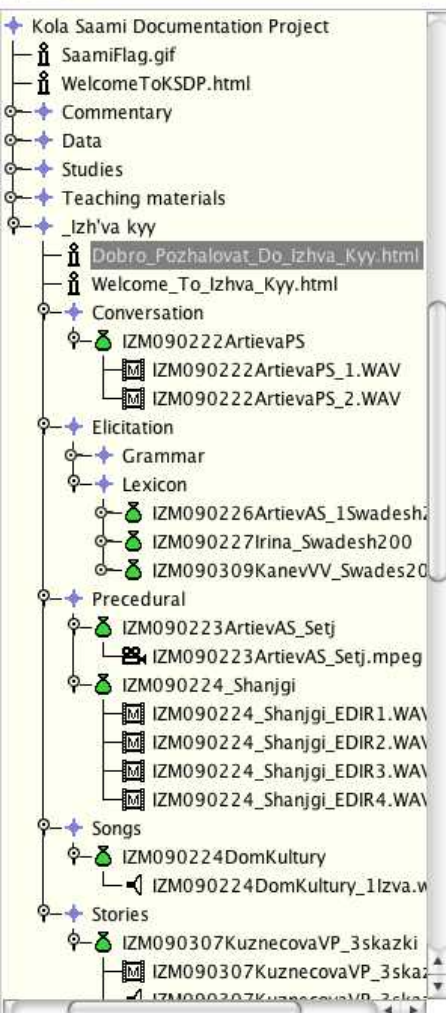


What makes an archive?

- ~~Original field notes, DVDs, DV-tapes stored in a safe?~~
- ~~Recordings and annotations stored in file directories on hard drives or university servers?~~
- ~~Web-based media presentations, on-line text editions and dictionaries?~~

Ideally (cf., e.g., DoBeS, HRELP): Open formats (.xml not .doc, etc.); Free access (at least) to all metadata; Chronic backup and format conversion

# Archives... (Izva kyy – intro)



## Izva Kyy - Архив ижемского языка Российской Федерации

Архив Izva Kyy создается в рамках проекта Kola Saami Documentation Project (KSDP) при сотрудничестве с Рохиром Блокландом, (Универстет Гамбург), Еленой Карвовой (Университет Потсдам) и представителями ижемского сообщества из Краснощелья и Ловозера.

Izva kyy 'ижемский язык' это вариант ижемского диалекта языка коми распространенный на Кольском полуострове на северо-востоке России. Ижемский диалект является диалектом коми-зырянского языка и, таким образом, относится к пермской ветви уральской языковой семьи. Носители языка называют себя ижватас. С момента прибытия первых ижемцев (ижватас) на Кольский полуостров в конце 19 века, носители Izva kyy находились в тесном культурном и языковом взаимодействии с носителями саамских языков.

Архив Izva Kyy открыт для пополнения любыми материалами, связанными с языком, не зависимо от того, кто является их создателем или владельцем. Если Вы хотели бы разместить материалы в нашем архиве, обращайтесь, пожалуйста, к [Рохиру Блокланду](#).

В данный момент над созданием архива работают Рохир Блокланд (Гамбург), Михаэль Рисслер (Фрайбург), Елена Карвовская (Потсдам), Ирина Канева (Краснощелье) и Валентина Кузнецова (Ловозеро).

### Авторские права

Материалы, находящиеся в архиве, были предоставлены нам носителями ижемского языка при условии, что они не будут использоваться в коммерческих целях. Распространение материалов недопустимо без согласия заведующих архивом. Подробнее смотрите [Этические и правовые аспекты](#) на сайте DoBeS.

Если Вам необходима процитировать материалы данного корпуса в вашей публикации, указывайте, пожалуйста следующий источник:  
*Рохир Блокланд, Ирина Канева, Елена Карвовская, Валентина Кузнецова & Михаэль Рисслер, 2009, Izva Kyy Archive, MPI Nijmegen, www.mpi.nl/DOBES/.*

Если вы цитируете или ссылаетесь на неопубликованные документы, собранные здесь (описания и анализ), отправляйтесь, пожалуйста на автора(ов) и название документа, а также обращайтесь на ссылки, которые приводятся в начале каждого документа.

### Записи, находящиеся в архиве

Записи, которые на сегодняшний день находятся в архиве, очень неоднородны, также сильно различается степень их аннотации. В архиве представлены лексические и грамматические анкетирования, записи разговоров, рассказов и песен.

# Archives... (Izva kyy – metadata example)

The screenshot shows a web browser window titled "IMDI Browser" with the URL [http://corpus1.mpi.nl/ds/imdi\\_browser/?openpath=MPI77915%23](http://corpus1.mpi.nl/ds/imdi_browser/?openpath=MPI77915%23). The browser's address bar contains a search engine (Google) and navigation buttons. Below the browser window, the "IMDI-Browser" interface is visible, displaying a tree view on the left and a metadata record on the right.

**IMDI** ISLE Metadata Initiative

**Session**

<b>Name</b>	IZM090224_Shanjgi
<b>Title</b>	IZM090224_Shanjgi
<b>Date</b>	2009-02-24

**Description**

Baking of traditional Komi šan'gi. The whole process took about four ours and was recorded on video. The audio was additionally recorded using an external mic and an Edirol.

**Location**

**Project** [Izva Kyy Documentation](#)

**Content**

**Actors**

- [Actor](#) [Lena](#)
- [Actor](#) [Irina Viktorovna](#)
- [Actor](#) [Rogier](#)
- [Actor](#) [Vasilij Vasil'evič](#)

**MediaFile**

- [MediaFile](#)
- [MediaFile](#)
- [MediaFile](#)
- [MediaFile](#)

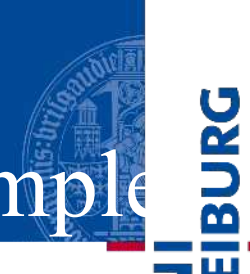
**Anonyms**

**References**

The left sidebar shows a tree view of the "Kola Saami Documentation Project" with various sub-items like "SaamiFlag.gif", "WelcomeToKSDP.html", "Commentary", "Data", "Studies", "Teaching materials", and "\_Izh'va kyy". Under "\_Izh'va kyy", there are sub-items for "Dobro\_Pozhalovat\_Do\_Izhva\_Kyy.html", "Welcome\_To\_Izhva\_Kyy.html", "Conversation", "Elicitation", "Precedural", "Songs", and "Stories". The "Conversation" item is expanded, showing a list of audio files including "IZM090222ArtievaPS", "IZM090224\_Shanjgi", and several "IZM090224\_Shanjgi\_EDIR" files.



# Archives... (Kola Saami – ELAN example)



Elan - KIL070421Kudckullj.eaf

File Edit Annotation Tier Type Search View Options Window Help

Grid Text Subtitles Audio Recognizer Metadata Controls

Volume: 100

Rate: 100

00:00:29.125 Selection: 00:00:01.889 - 00:00:07.930 6041

Selection Mode Loop Mode

	00:00:28.000	00:00:29.000	00:00:30.000	00:00:31.000	00:00:32.000	00:00:33.000	00:00:34.000
reference [76]	Kudckullj.009				Kudckullj.010		Kudckull
KSDP Cyr orthogra [63]	И оара сярна касът тэдт гәз-дәдт.				Гәз апс вьлшхәла тәйдта лөв.		Пург кәһ
morphemes [574]	le	i	oarr(wk)	a	särrn(wk)	a	kas't tedt gāz =dedt gāz app's(wk-npal) vil'šxel a tejdta ijij(supl-ie) pugk
part of speech [566]	le	conj	v	-v-ax	v	-v-ax	interrog dem n =dem n v v -a-ax dem v pro
gloss [571]	RO	and	sit	1SG.PR	say	1SG.PR	from_where PROX:NOM.SG gas PROX:N gas smell.3SG.PR look 1SG.P PROX:NOM.P be:3SG.PRS all
translation (Russia) [67]	и говорю: "откуда это газом пахнет"				смотрю откуда это		все закр

# Corpus- and computer linguistics for endangered languages



## E.g. Giellatekno (University Tromsø)

- Corpora of written languages: North and other Saami languages, Komi, Greenlandic, other
- Corpus linguistic tools (parsers, automators, etc.)
- Practical tools (dictionaries, spell checkers, translation tools, [T9, speech recognition], etc.)



- The documentary linguist...
  - needs programming skills anyway... :-)
- The C+C linguist...
  - needs larger corpora (language documentation) and better morphological and syntactic analyses (theoretical linguistics)
- The community...
  - needs dictionaries and other tools (language maintenance)
  - likes fancy gadgets and also wants the language visible (language revitalization)



- DoBeS Archive
  - [www.mpi.nl/dobes/](http://www.mpi.nl/dobes/)
- Giellatekno
  - [giellatekno.uit.no/](http://giellatekno.uit.no/)
- Kola Saami Documentation Project
  - [www2.hu-berlin.de/ksdp/](http://www2.hu-berlin.de/ksdp/)
- Rießler, Michael (Universität Freiburg)
  - [www.skandinavistik.uni-freiburg.de/institut/mitarbeiter/riessler/](http://www.skandinavistik.uni-freiburg.de/institut/mitarbeiter/riessler/)