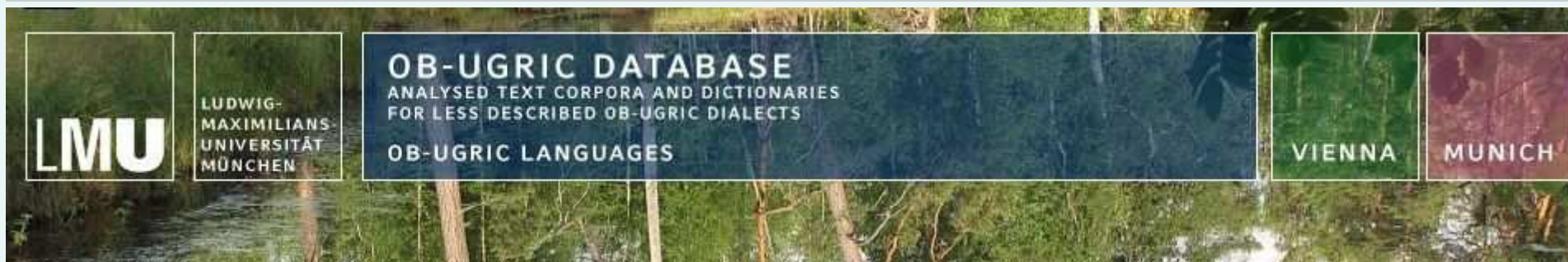


# Zsófia Schön & Axel Wisiorek



3rd International Winter School for Finno-Ugric Studies  
Munich, 02/13/2015



# Outline

1. About the Project(s)
2. Target Languages
3. Data Sources
4. Workflow
5. Corpora



# 1. About the Project(s)

**EUROBABEL**  
BETTER ANALYSES BASED ON ENDANGERED LANGUAGES

- August 2009-July 2012
- Universities of Munich  
Vienna  
Szeged  
Helsinki
- 29 Members

**OB-UGRIC DATABASE**  
ANALYSED TEXT CORPORA AND DICTIONARIES  
FOR LESS DESCRIBED OB-UGRIC DIALECTS

- July 2014-June 2017
- Universities of Munich  
Vienna
- The Language Archive
- 13 Members

## 2. Target Languages

**EUROBABEL**  
BETTER ANALYSES BASED ON ENDANGERED LANGUAGES

**OB-UGRIC DATABASE**  
ANALYSED TEXT CORPORA AND DICTIONARIES  
FOR LESS DESCRIBED OB-UGRIC DIALECTS

Northern Mansi

+

Eastern Mansi

Pelym Mansi

Kazym Khanty

North Vagilsk Mansi

Surgut Khanty

Yugan Khanty



## 3. Data Sources

- Based on written material only:

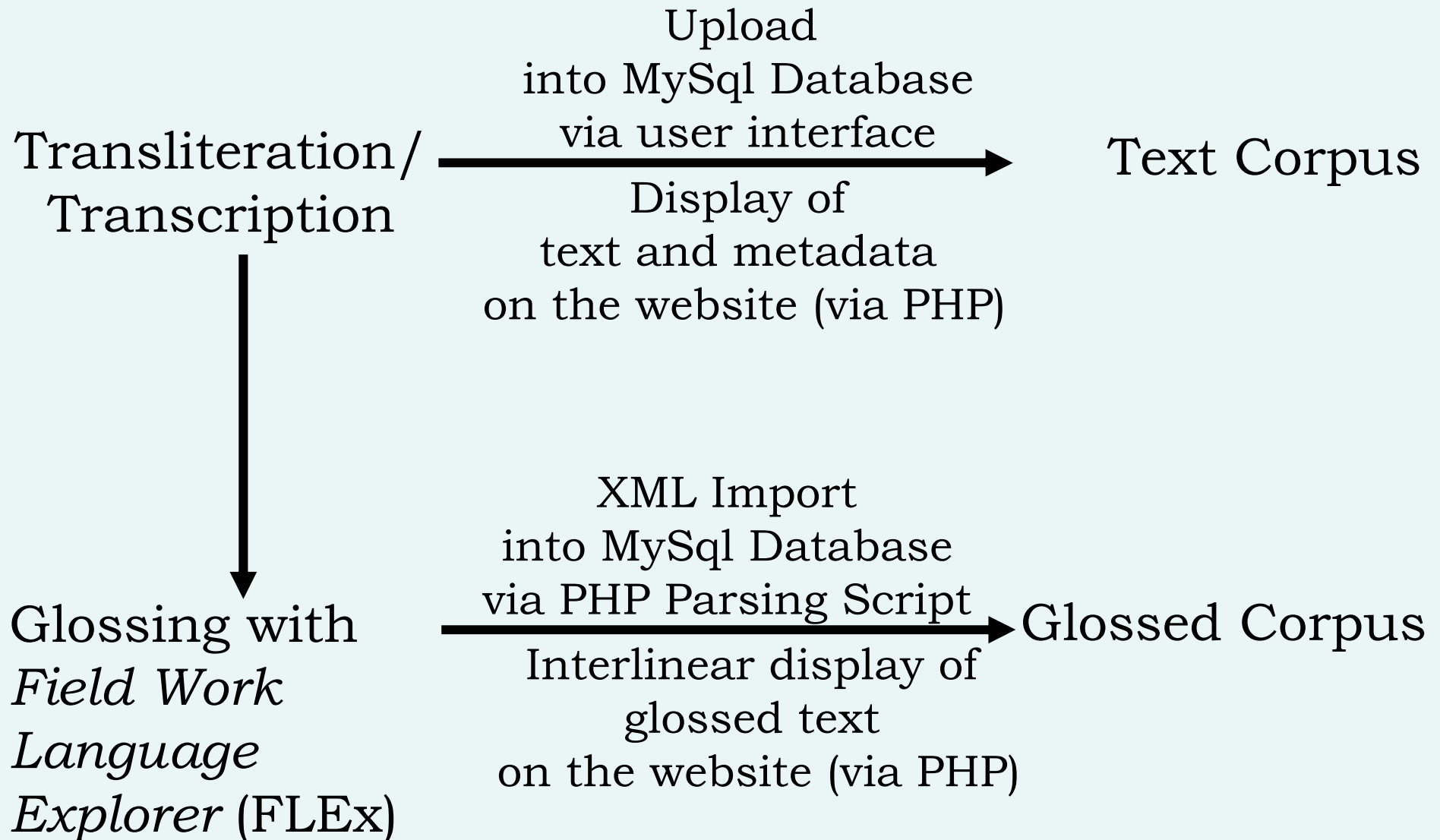
Northern Mansi, Eastern Mansi, Pelym Mansi,  
North Vagilsk Mansi, Kazym Khanty,  
Surgut Khanty

- Based on written and spoken material:

Yugan Khanty



## 4. Workflow



# 5. Corpora

INTRODUCTION

ABOUT EUROBABEL

ABOUT OUDB

ENCYCLOPEDIA SECTION

LINGUISTIC SECTION

**Text Corpus**

All Corpora

All Mansi Corpora

All Khanty Corpora

Northern Mansi

Eastern Mansi

North Vagilsk Mansi

Pelym Mansi

Kazym Khanty

Surgut Khanty

Yugan Khanty

Search Glossed Corpus

Transcription, Transliteration and Orthography

e-Dictionaries

e-Grammar

LINKS

INTERNAL SECTION

all dialects optional dialect all genres all texts all texts engl. translai Search

Original Title	Dialect	English Title	Genre Form	Genre Content	ID	glossed	Audio	Translation
a j mo:s-nenije	kazym khanty (KK)		prose (pro)	Tales (tal)	885	-	-	engl., ger
nie ni βerta βer	surgut khanty (SK)		prose (pro)	Ethnographic Texts (eth)	1076	-	Audio	engl., ger, ru
"wa psxum! jani ne o:nipen, o:nipen kantli: [...]	northern mansi (NM)	The Mother-In-Law Wants to Have New Clothes from the Son-In-Law.	poetry/song (poe)	Performances at Bear Ceremonies (bep)	1164	-	-	engl., ger
7 maj - radio xotal	northern mansi (NM)	May 7th - day of radio	prose (pro)	Journalism (jou)	1230	glossed	-	engl., ger
ak* jo:ran o:li, a:xyiy piyiy onisi: [...]	northern mansi (NM)	Samoyeds rob a Samoyed Family but Are Destroyed Themselves.	prose (pro)	War Songs - Heroic Songs (her)	1136	-	-	engl., ger
ak* ojk o:li. to ramel pats, takke:t, ma:nal pats, takke:t.	northern mansi (NM)	A Lonesome Man Lovers his Son to the Earth, the Fate of the Son.	(mix)	War Songs - Heroic Songs (her)	1143	-	-	engl., ger
ak* ojk o:li. xu:ram piy onisi: [...]	northern mansi (NM)	Legend of a Man's Youngest and Lazy Son who at the End Becomes a Protective Spirit.	prose (pro)	War Songs - Heroic Songs (her)	1144	-	-	engl., ger
ak* xum o:li takke:t pa:welt, takke:t u:set. [...]	northern mansi (NM)	Tab Collection of the Protective Spirits	prose (pro)	Mythological Texts (myt)	745	-	-	engl., ger
am e:pej sa:n u:nttasam [...]	northern mansi (NM)	Prayer while Offering a Food Sacrifice.	poetry/song (poe)	Mythological Texts (myt)	1034	-	-	engl., ger
am ogmea:šiam #, # am	pelym mansi (PM)		prose (pro)	Riddles (rid)	1293	-	-	engl.

= 274 texts



# 5. Corpora

INTRODUCTION

ABOUT EUROABEL

ABOUT OUDB

ENCYCLOPEDIA SECTION

**LINGUISTIC SECTION**

**Text Corpus**

- All Corpora
- All Mansi Corpora
- All Khanty Corpora
- Northern Mansi
- Eastern Mansi
- North Vagilsk Mansi
- Pelym Mansi
- Kazym Khanty
- Surgut Khanty
- Yugan Khanty
- Search Glossed Corpus
- Transcription, Transliteration and Orthography
- e-Dictionaries
- e-Grammar

LINKS

INTERNAL SECTION

Surgut Khanty optional dialect all genres all texts all texts russian trans tal Search

Original Title	Dialect	English Title	Genre	Form	Genre Content ID	glossed	Audio	Translation
i ki sesəy ɸer (1996/A)	surgut khanty (SK)	A man made a trap (1996/A)	prose (pro)	Tales (tal)	1081	-	Audio	engl., ru
i ki sesəy ɸer (2008/A)	surgut khanty (SK)	A man made a trap (2008/A)	prose (pro)	Tales (tal)	1085	-	Audio	engl., ru
kəjəp	surgut khanty (SK)	The Shaman's Drum	prose (pro)	Tales (tal)	738	-	Audio	engl., ru, hun
pi tən̄kəliyən-o pi sɛyən A	surgut khanty (SK)	The Little Bird and his Sister	prose (pro)	Tales (tal)	732	glossed	Audio	engl., ru, hun
təɸət i mi (1996/B)	surgut khanty (SK)	The fire goddess (1996/B)	prose (pro)	Tales (tal)	1083	-	Audio	engl., ru
əj jəyset	surgut khanty (SK)	The Three Brothers	prose (pro)	Tales (tal)	734	glossed	Audio	engl., ru, hun
oɸəs ort pe nə ro pəttətə qo. pərili	surgut khanty (SK)		prose (pro)	Tales (tal)	737	-	Audio	engl., ru, hun
təŋkərəli	surgut khanty (SK)	The little mouse	prose (pro)	Tales (tal)	731	-	Audio	engl., ru, hun

= 8 texts



# 5. Corpora

INTRODUCTION

ABOUT EUROABEL

ABOUT OUDB

ENCYCLOPEDIC SECTION

LINGUISTIC SECTION

**Text Corpus**

All Corpora

All Mansi Corpora

All Khanty Corpora

Northern Mansi

Eastern Mansi

North Vagilsk Mansi

Pelym Mansi

Kazym Khanty

Surgut Khanty

Yugan Khanty

Search Glossed Corpus

Transcription, Transliteration  
and Orthography

e-Dictionaries

e-Grammar

LINKS

INTERNAL SECTION

Translation Audio Files **Glossed Text**

Original Title	Dialect	English Title	Genre Form	Genre Content	ID	glossed
pi:təŋkəliyən-o:piseyən A	surgut khanty (SK)	The Little Bird and his Sister	prose (pro)	Tales (tal)	732	glossed
German Title	Russian Title	Hungarian Title	Finnish Title			
	Братец-птичка и его старшая сестра A	Madárka és a nénye A				
Informant	Collector	Editor	Source Publication	Status	Audio File	
Pokačeva, Dar'á Stepanovna	Csepregi, Márta (CS)	Csepregi, Márta	Csepregi, Márta, 1998: 66-67, 136.	IPA transliteration	1	

## pi:təŋkəliyən-o:piseyən A

1. pi:təŋkəliyən-o:piseyən baŋəyən. t̪i baŋtin əj məte ʎe:tnə pi:təŋkəli jə:stə: "ser mə: ʎorəp ni:k məʎəttəyə jaŋqəŋəm." pə:nə t̪u: pu:r:nə, pu:r qatəlnə mən, ənəl sər:t ʎe: t̪u: su:rtə:ŋ... o:pəliŋnem jə:stə: "nəŋ n̄ə:rəq qu:ʎ ʎərə! potʃ qu:ʎəliŋ mə:ntem, qon qu:ʎəliŋ mə:ntem. mə: ur:ʎ əŋintəm." pə:n əŋintəy. noq ʎərəyʎəmə:ŋə: "potʃ qu:ʎəlijet mə:nt ku:r:təptən qon qu:ʎəlijet?" o:pil jə:stə: "əntə!" pə:nə pi:təŋkəli pu:r:təmtəy: "məŋk i:kije məŋtəm, pə:n nəŋet məŋk i:kije məŋtəm."

2. t̪i mən. məŋk i:ki qv:tə jəʎət, qv:t qoni jaqənəm ʎe:jə:ŋ, məŋk i:ki jaqən əntem. ke:r oʃtinə poʃt̪i pu:təli v:məʃ. ur:ʎə toyə mən, poʃt̪i t̪i ʎi:ʎ, ʎo:jə t̪i ʎi:ʎ, ʎi:ʎ, pə:n əj məte ʎe:tnə ke:mən ... kə:k jəʎo ... məŋk i:ki lu:mp seŋkiŋtə səj. t̪u: ʎe:tnə pi:təŋkəli qan'təqintəy. ke:r potʃə toyə qan'imtəy. məŋk i:ki jaqə ʎəŋmə:ŋ ʎe:tnə n̄ə:ʎəm: "mə: ʎo:jə pu:tələm qojənə ʎi:ʎi? sər mə: ke:r noq əŋtəm, ke:r t̪e:m ləkəmtələm." pə:nə ke:r əŋtə ju:ʎe kəntʃt̪əyə t̪oʎəlnem t̪i kət nərtəmtəy, pi:təŋkəli t̪ə ke:m n̄ərimtəy. "əj, v̄ŋkenoʃliŋki, mə: nəŋet sər ke:r ləkəmtələm." t̪u: ʎe:tnə pi:təŋkəli n̄ə:ʎəm: "qanəm i:ki, mə:nt ke:r v̄:ʎ ləkəmtə, mə: nəŋeti o:pem mələm." "jə:, ʎəs, toyənə qu:ntə, toyənə. nəŋ məntem mə: nəŋet qo:ŋə o:jəyətəŋ?" "mə: məntem sər:yə:ŋ əj pajtəŋ pelkəm ur:ʎə jeryəntəŋ, ki:mət pajtəŋəm ur:ʎə jeryəntəŋ, nəŋ t̪u: le:ki jəʎə!" pə:nə pi:təŋkəli mən. pi:təŋkəli jaqə jəʎətmə:ŋ jə:stəy: "mə: nəŋet məŋk i:kije məjəm." o:pəli j̄i:sə:ŋ. pi:təŋkəli jə:stə: "v̄:ʎ j̄i:sə:ŋ pət̪kəyʎəmin o:ʎpi qv:nəŋə ʎoŋq ʎəŋtəmən, pə:nə ke:ʎrəm ʎəy toyə pan'əmən." pə:nə o:pəliŋ əntə j̄i:stəyə jəy. toyə tərəm.

3. pə:nə məte mərə baŋtin pu:r:nə məŋk i:ki t̪i jəʎət. o:ʎpi v̄:j, o:ʎpi j̄i:ŋ əntə ʎəpəl. pi:təŋkəli n̄ə:ʎəm: "qanəm i:ki potʃ pəryi, qanəm i:ki potʃ pəryi!" qanəm i:ki pə: pelkə:ŋ pə: kətʃ ʎəryi:ŋ, to:ŋ sər:yə:ŋ, potʃ pelkə:ŋ t̪i jaqənəm. jaqə əŋə kərəy. t̪u: ʎoŋk pətəyə, ke:ʎrəm ʎəy o:ʎt̪i:ŋə toyə kərəy. pə:nə ʎəŋtəm ʎəpəl tot tərəm. pi:təŋkəliyən-o:piseyən t̪u: məŋk i:ki t̪e:s-ʎəyŋet i:t̪pə ʎəŋtəyən.

# 5. Corpora

Translation Audio Files Glossed Text

Original Title	Dialect	English Title	Genre Form	Genre Content	ID	glossed
pi:təŋkəliyən-o:pisəyən A	surgut khanty (SK)	The Little Bird and his Sister	prose (pro)	Tales (tal)	732	glossed

German Title	Russian Title	Hungarian Title	Finnish Title
	Братец-птичка и его старшая сестра А	Madárka és a nényje A	

Informant	Collector	Editor	Source Publication	Status	Audio File
Pokačeva, Dar' à Stepanovna	Csepregi, Márta (CS)	Csepregi, Márta	Csepregi, Márta. 1998: 66-67, 136.	IPA transliteration	1

## pi:təŋkəliyən-o:pisəyən A (glossed version)

```
# 1 .# pi:təŋkəliyən-o:pisəyən      βaŋəyən      ##
      pi:təŋkəli-yən o:pi-sə-yən      βaŋ-l-əyən
      pi:təŋkəli-yən o:pi-sə-yən      βaŋ-l-əyən
      little_bird-DU older_sister-COLL-DU live-PRS-3DU
      subs-infl:n subs-deriv:n>n-infl:n      v-infl:v-infl:v
```

'There lived a little bird and his sister.'

'Es lebte einmal ein Vögelchen zusammen mit seiner älteren Schwester.'

'Жили-были братец-птичка со своей старшей сестрой.'

'Madárka a nényével él.'

```
tī      βaŋtin      əj mətə lə:tnə      pi:təŋkəli      jə:stəl      #: "#      ser      mɛ:      βorəp      ni:k      mɛβətəyə
tī      βaŋ-t-in      əj mətə lə:tnə      pi:təŋkəli      jə:stə-l      ser      mɛ:      βorəp      ni:k=      mɛβət-təyə
tī      βaŋ-tə-in      əj mətə lə:tnə      pi:təŋkəli      jə:stə-l      ser      mɛ:      βorəp      ni:k=      mɛβət-təyə
so      live-PTCP.PRS-3DU      some_day      little_bird      say-PRS[3SG]      now      1SG      trousers      down_to_the_bank      lower-INF
ptcl      v-deriv:v>ptcp-infl:ptcp      adv      subs      v-infl:v      adv      ppron      subs      prvb      v-deriv:v>inf
```

'Living so, at some point the little bird says: 'Well, I am leaving to lower my trousers down [into the river].''

'Während die beiden so leben, sagt das Vögelchen eines Tages: „Ich gehe jetzt zum Fluss, um mit meiner Hose zu fischen.“'

'Однажды он сказал своей сестре: «Я сейчас схожу и окуну свои штаны в речку».'

'Így élvén egyszer madárka mondja: "No, én elmegyek a nadrágomat lemeríteni."'



# 5. Corpora

INTRODUCTION

ABOUT EUROABEL

ABOUT OUDB

ENCYCLOPEDIA SECTION

LINGUISTIC SECTION

- Text Corpus
  - All Corpora
  - All Mansi Corpora
  - All Khanty Corpora
  - Northern Mansi
  - Eastern Mansi
  - North Vagilsk Mansi
  - Pelym Mansi
  - Kazym Khanty
  - Surgut Khanty
  - Yugan Khanty
  - Search Glossed Corpus**
  - Transcription, Transliteration and Orthography
  - e-Dictionaries
  - e-Grammar

**Search Glossed Corpus**

This concordance tool provides a search function for word forms, dictionary entries (stems and suffixes) and glosses in the glossed texts of the corpus and shows the keyword in context (KWIC).

SQL query: REGEXP search in token table/dictionary table with self join of token table for displaying the cotext of the keyword

# 5. Corpora

Search Corpus for

<b>Input:</b>	a	e	i	o	u	y	ü	c	j	k
<b>Match:</b>	a a: ə: ɔ ə	e e: ɛ ə ɐ: æ æ: ɛ: ɛ ea_	i i: u: w i	o o: ɔ ɔ: ɔa_ ɔ ɔæ_ ɔɔ	u u: ʉ ʉ:	y y: yø_	y y: yø_	tʃ tʃi	j	k k* kʰ

<b>Input:</b>	l	m	n	ng	b	d	f	g	h	p	q	r	s	t	v	w	x	z
<b>Match:</b>	l lʰ lʰʰ	m	n nʰ ŋ	ŋ	b	d	f	gɣ	x h	p	q	r	s sʰ j	t tʰ	v	β w	x x*	z

Use \_ for exactly one single character.  
 Use \* for an unspecified number of characters.  
 Use - in lemma/morpheme search for suffixes (-nə).

KK 886	βe:ɪpəsələtti . ja: . tom βa:ntafən : ij imi	<b>oxəl</b>	moitti βo:jən tixəla βerəm . ” siti potərtal sa:xat
KK 886	moitti βo:jən xol-pela si xoreməsijmən .” βuti joxətsəŋən .	<b>oləm</b>	xə nox-kiləmtəs , ləmətʰlēməs , niməlɛjəl βa:ttēməs
KK 886	pəʰsəŋ nʰoxilal nik-moitti pa:ɾsəʰle . put xə:nəm a:ɾŋən ləti-ja:nʰsi	<b>ɔ:məsti</b>	pa:ɾsajŋən . xə:təŋ xə lopijəl tonti-nʰa:nʰ-pox pəla :
KK 886	ja: si , xojat-kepa !” joxi-foŋəs , pasan xə:nəŋa	<b>ɔ:məs</b>	. fuβ xə:nŋələn sem-pələk por-nele ɔ:məsəl . ləsət-j
KK 886	, pasan xə:nəŋa ɔ:məs . fuβ xə:nŋələn sem-pələk por-nele	<b>ɔ:məsəl</b>	. ləsət-ja:səsət . xuβ muj βa:n βəsət , joxi-manti
KK 886	aft ta:sən-βuʰən jokanlajəmŋ .” manti pitsət . tapri-polxi taβret	<b>əxəlɛla</b>	lɛʰsajət . sem-pələk por-nele əxəl puʰija lələs . 5.
KK 886	pitsət . tapri-polxi taβret əxəlɛla lɛʰsajət . sem-pələk por-nele	<b>əxəl</b>	puʰija lələs . 5. joxi-joxtəs . aŋkəl-imi ij xə:t-suŋa

„Fuzzy search“ : IPA-matching with ASCII-characters



# 5. Corpora

Search Corpus for: **word form**

<b>Input:</b>	a	i							
<b>Match:</b>	a a: ə: d a:	i i: w: w i:							

<b>Input:</b>	l	m	n	ng	b	d	f	g	r	s	t	v	w	x	z
<b>Match:</b>	l l̥ l̥ʰ	m	n n̥ n̥ʰ	ŋ	b	d	f	g	r	s s̥ s̥ʰ	t t̥ t̥ʰ	v	β w	x xʷ	z

Use \_ for exactly one single character.  
Use \* for an unspecified number of characters.  
Use - in lemma/morpheme search for suffixes (-nə).

# 5. Corpora

Search Corpus for **word form** o:pi in all dialects sort by dialect Search Clear

<b>Input:</b>	a	e	i	o	u	y	ü	c	j	k
<b>Match:</b>	a a: e: o: e	e e: e ə ə: æ æ: ə: e ea_	i i: u: u i	o o: e o: o: oa_ o ø æ_ o_	u u: u u:	y y: yø_	y y: yø_	tʃ tʃi	j	k k* kʰ

<b>Input:</b>	l	m	n	ng	b	d	f	g	h	p	q	r	s	t	v	w	x	z
<b>Match:</b>	l l i f #	m	n n i ŋ	ŋ	b	d	f	g y	x h	p	q	r	s s i f	t t i	v	β w	x x*	z

**no result.**

Search Corpus for **lemma/morf** o:pi in all dialects sort by dialect Search Clear

<b>Input:</b>	a	e	i	o	u	y	ü	c	j	k
<b>Match:</b>	a a: e: o: e	e e: e ə ə: æ æ: ə: e ea_	i i: u: u i	o o: e o: o: oa_ o ø æ_ o_	u u: u u:	y y: yø_	y y: yø_	tʃ tʃi	j	k k* kʰ

<b>Input:</b>	l	m	n	ng	b	d	f	g	h	p	q	r	s	t	v	w	x	z
<b>Match:</b>	l l i f #	m	n n i ŋ	ŋ	b	d	f	g y	x h	p	q	r	s s i f	t t i	v	β w	x x*	z

SK 732 1. **pi:təŋkəliyən-o:piseyən** βəffəyən . tʃi βəftin əj məte fə:tnə pi:təŋkəli jə:stə: " "

SK 732 mən , ənəl sɔ:rt βeɪ . tʃu: su:rtə: ... **o:pəliɲem** jə:stə: " nəŋ nʲe:rəq qu:ɪ βəre ! potʃ

SK 732 : " potʃ qu:ɪləjjet mə:nt ku:təptən qon qu:ɪləjjet ? " **o:pil** jə:stə: " ənte ! " pə:nə pi:təŋkəli pu:təmtəy :

SK 732 qonəm i:ki , mə:nt ke:re v:ɪ ləkəmtə , mə: nəŋeti **o:pem** mələm . " " jə: , βəs , toyənə qu:ntə

SK 732 jə:stəy : " mə: nəŋet məŋk i:kije məjəm . " **o:pəlil** jɪ:sə: pi:təŋkəli jə:stə: " v:ɪ jɪ:sə !

SK 732 βerləmən , pə:nə ke:βrəm βəy toyə pənələmən . " pə:nə **o:pəlil** əntə jɪ:stəyə jəy . toyə tərəm . 3

SK 732 toyə kerəy . pə:nə βerləm βerəl tot tərəm . **pi:təŋkəliyən-o:piseyən** tʃu: məŋk i:ki tə:s-βəyɲet i:tə βəffəyən .



# 5. Corpora

Search Corpus for glosses **SG<1sg older\_sister** in all dialects sort by dialect Search Clear

<b>Input:</b>	a	e	i	o	u	y	ü	c	j	k
<b>Match:</b>	a a: e: ə e	e e: ɛ ə ə: æ æ: ɐ e ə	i i: w: w i	o o: ɔ o: ɔ ə ɔ ə ə ə ɔ	u u: ʉ ʉ:	y y: y ə	y y: y ə	tʃ tʃi	j	k kʷ kʰ

<b>Input:</b>	l	m	n	ng	b	d	f	g	h	p	q	r	s	t	v	w	x	z
<b>Match:</b>	l lʰ lʰ	m	n nʲ ŋ	ŋ	b	d	f	g ɣ	x h	p	q	r	s sʲ ʃ	t tʰ	v	β w	x xʷ	z

SK 732

qanəm i:ki , mənt ke:re v:f ləkəntə , mə: nəŋeti **o:pem** mələm .” “ je: , βəs , toyəne qunte

# 5. Corpora

Search Corpus for lemma/morf  in all dialects sort by dialect Search Clear

<b>Input:</b>	a	e	i	o	u	y	ü	c	j	k								
<b>Match:</b>	a a: e: d e	e e: e ə ə: æ æ: ə: ə ea_	i i: u: w i	o o: e o: o: oa_ o øæ_ o_	u u: u u:	y y: y_	y y: y_	tʃ tʃi	j	k kʷ kʰ								
<b>Input:</b>	l	m	n	ng	b	d	f	g	h	p	q	r	s	t	v	w	x	z
<b>Match:</b>	l lʰ lʲ	m	n nʲ ŋ	ŋ	b	d	f	g ɣ	x h	p	q	r	s sʲ ʃ	t tʰ	v	β w	x xʷ	z

NM 1229	akˀmat ɛ:rtɪn kali sunsi : ɛtəs .   59.	<b>sa:s</b>	fʊβ pa:jpəŋ nɛm ɛ:kˀa xɛn ti juw . tajʃ
NM 1229	la:sʰal , βɛs .   76. anʲi	<b>sa:s</b>	a:t pa: pa:jpənəl βɛs .   77.
NM 1229	naŋ xɪʃ patapasum ! – tuβa .   105.	<b>sap</b>	ma: la:sʰal , ka:mɪn sa:s ke:β pa:jpənəl βɛjtsəm .
NM 1229	tuβa .   105. sap ma: la:sʰal , ka:mɪn	<b>sa:s</b>	ke:β pa:jpənəl βɛjtsəm .   106. kˀa:lsən , tɛ:r
NM 1229	xa:jttawe . tiβ timos a:ʃ la:wi : – naŋ	<b>sa:s</b>	pa:jpənəl joxtiʃa .   127. ma: joxtiʃɛʃti kˀa:lsən ,
NM 1229	nɔ:wɪlum ɔ:ŋɛt a:j .   185. atəl misɛʃ βɛn	<b>pasi</b>	ɔ:ŋɛt . xasap tajʃ sa:jəŋ .   186. a:sɛm
NM 1229	ma: alay tax	<b>nas</b>	nʲa:ksʰata:lɛ:ɪn .
NM 1234	jorəy jɛmtalan ( jɛmtalanete ) porat ɛkˀa piyrisʲ mis kotlɛ	<b>pasnəl</b>	takˀi kole kiwɛrna xartiylite . ɛkˀa piyrisʲ ta xartəs ,
NM 1234	– matarəy alam naj aɣən jaltapɛjɪn ! ta kos paxwtasəy ,	<b>nas</b>	worna tatsəy , anʲ ɛkˀaəy worna totwesəy . usəŋ
NM 1234	permtaxtas . ojKate kat supəy juniytaste , akˀ katanen	<b>sas</b>	supəy junəytsanen . ojka lawi : – tarməl xujɛln
SK 732	o:βpi qɔ:nəŋə βoŋk βɛrləmən , pɛ:nə ke:βrəm βəy toɣə	<b>panləmən</b>	.” pɛ:nə o:pəlɪʃ ɛntə ji:stəyə jəy . toɣə tərəm



# 5. Corpora

## pi:təŋkəliyən-o:pi:səyən A (glossed version)

older\_sister-DIM.MEL-SG<3SG cry-PRS[3SG]  
 subs-deriv:n>n-infl:n v-infl:v

'His little elder sister cries.'  
 'Sein älteres Schwesterchen fängt an zu weinen.'  
 'И сестра заплакала.'  
 'A nővérkéje sir.'

pi:təŋkəli	jə:stəl	#:	"#	v:ɪ	ji:sə	#!#
pi:təŋkəli	jə:stə-ɪ			v:ɪ	ji:s-ə	
pi:təŋkəli	jə:stə-ɪ			v:ɪ	ji:s+imp. stem-ə	
little_bird	say-PRS[3SG]		PROH	cry+imp. stem-IMP.2SG		
subs	v-infl:v		ptcl	v-infl:v		

'The little bird says: 'Don't cry!'  
 'Das Vögelchen sagt: „Weine nicht!'  
 'И тогда он сказал: «Не плачь!'  
 'A madárka mondja: "Ne sirj!'

pətkeɣləmin	o:βpi	qɔ:nəŋe	βoŋk	βerləmən	#,#	pə:nə	ke:βrəm	βay	toɣə	<b>panləmən</b>	#."#
pətkeɣlə-min	o:βpi	qɔ:nəŋ-e	βoŋk	βer-ɪ-əmən		pə:nə	ke:βrəm	βay	toɣə	pan-ɪ-əmən	
pətkeɣlə-min	o:βpi	qɔ:nəŋ-e	βoŋk	βer-ɪ-əmən		pə:nə	ke:βrəm	βay	toɣə	pan-ɪ-əmən	
hurry-CVB	door	beside-DLAT	pit	do-PRS-1DU		and	hot	pick	there	put-PRS-1DU	
v-deriv:v>cvb	subs	pstp-infl:pstp	subs	v-infl:v-infl:v		cconj	adj	subs	adv	v-infl:v-infl:v	

'Let us quickly get a pit ready beside the door, and put a hot pick there.'  
 'Lass uns schnell neben der Tür eine Grube machen und heißes Eisen dort hinein tun.'  
 'По быстрому возле двери выроем яму и положим туда раскаленное железо.'  
 'Az ajtó mellett gyorsan egy gödröt csinálunk, és tüzes vasat teszünk oda.'





**Thank you!**