

Competing motion-cum-purpose strategies in Northern Selkup

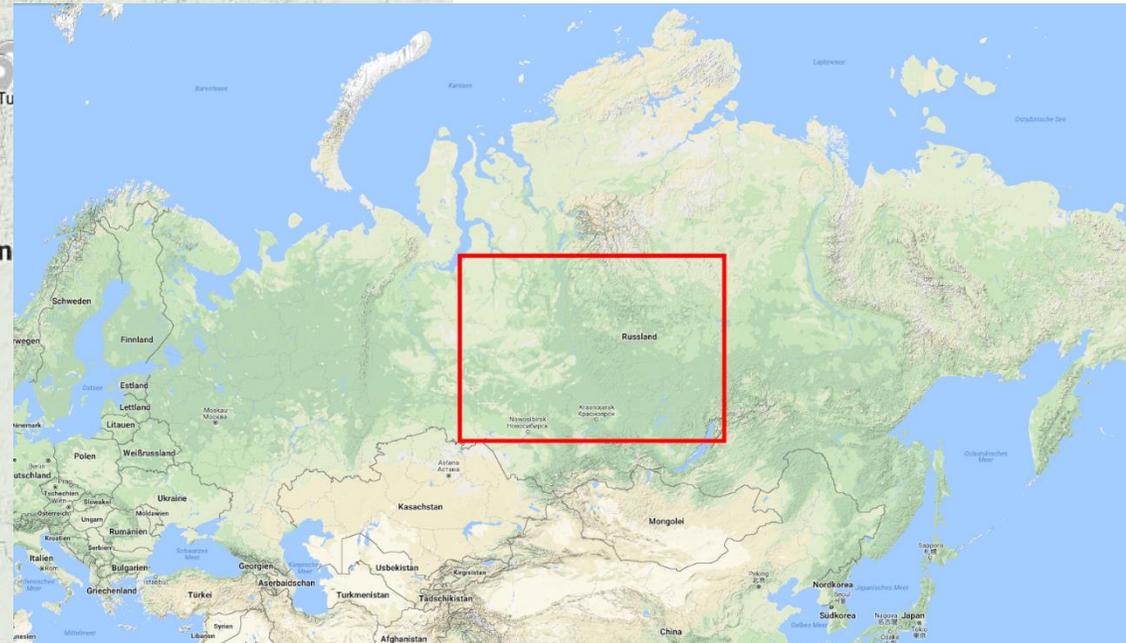
S. Toldova (NRU "Higher school of economics", Moscow)

M. Brykina (University of Hamburg; Lomonosov Moscow State University)

Outline

- The Selkup language
- INEL project >> corpus data
- Two types of purpose clauses in Northern Selkup
- Distribution of motion-cum-purpose constructions
- Explanation for this distribution
- Dialectal variation

Selkup





Irina Korobeynikova, one of the last speakers of Central Selkup

- Southern Samoyedic < Uralic
- Region: Western Siberia between Ob and Yenisey
- 3649 Selkups (census 2010)
- Sparse population, inaccessible locations
Great dialectal divergence
- Northern dialects:
severely endangered (< 100 speakers)
- Central and Southern dialects:
almost extinct (< 10 speakers)
- Intense contacts with different groups:
 - Ket (Yeniseian), Khanty (Finno-Ugric), Nenets (Samoyedic), Evenki (Tungusic) in the North
 - Siberian Tatar, Chulym (Turkic), Ket (Yeniseian) in the South



The INEL Project

The project „Grammatical Descriptions, Corpora and Language Technology for Indigenous Northern Eurasian Languages“ is a long-term project, aimed, among other, at digitalization and creating corpora from materials, preserved in various archives and collections.

- headed by Prof. Dr. Beáta Wagner-Nagy
- located at the University of Hamburg
- carried out by the Academy of Sciences of Hamburg

- current languages: Selkup, Dolgan, Ewenki
- already available corpora: Selkup, Kamas



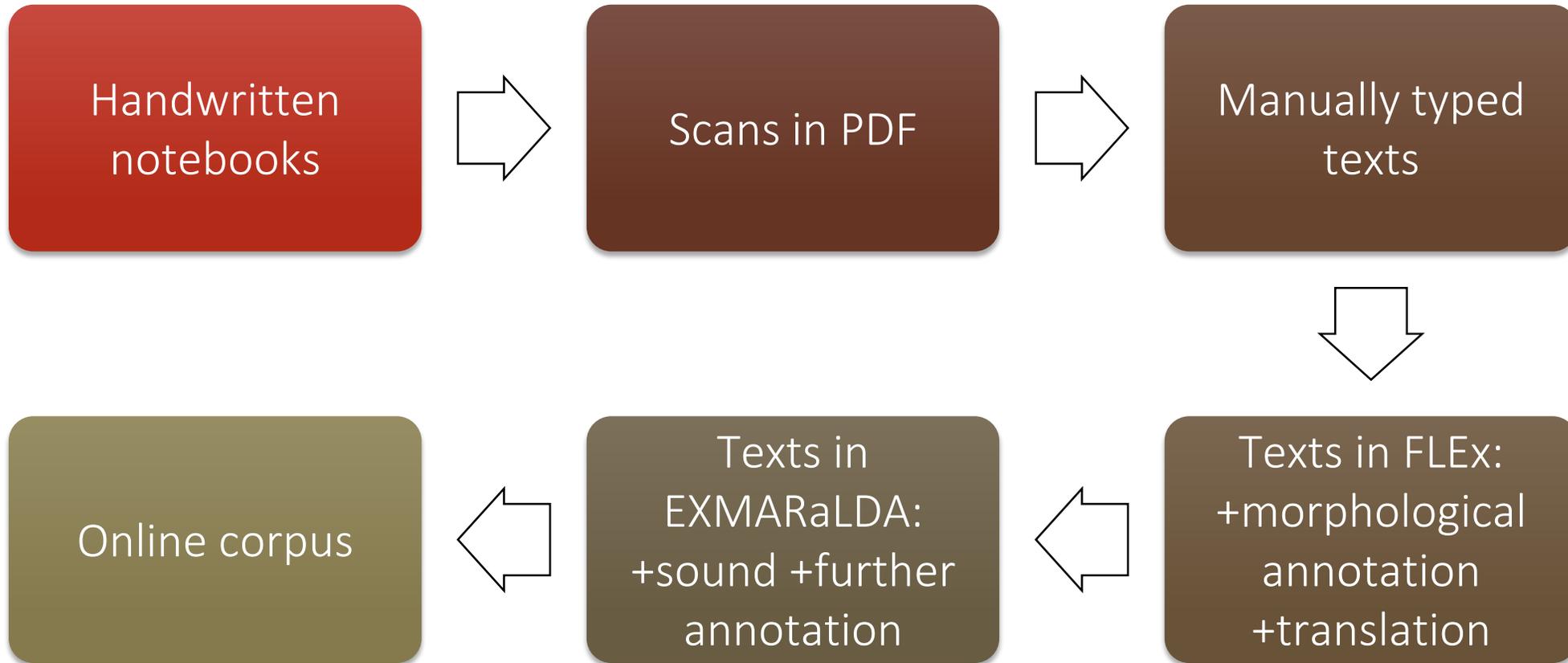
The INEL Project: Selkup corpus

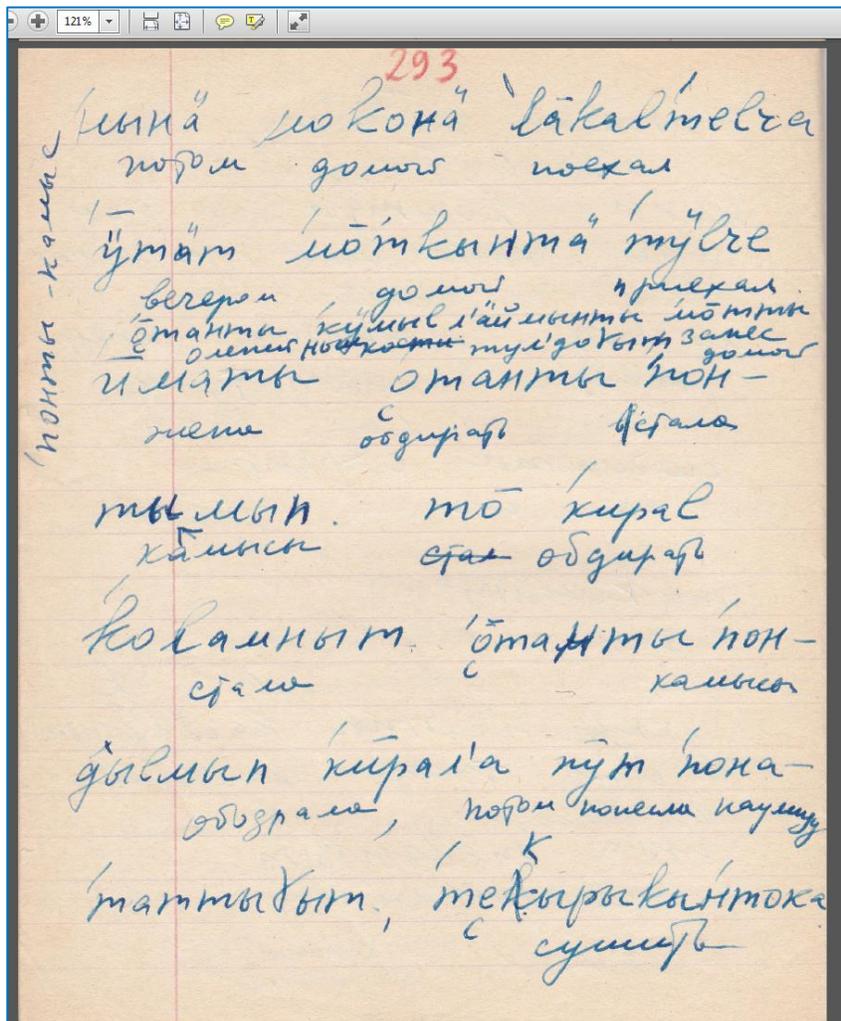


*Angelina Kuzmina
(1924-2002)*

- Data: manuscripts & audio recordings from the archive of **Angelina Kuzmina** preserved at IFUU
 - Data collected in 1960s–1970s in all dialectal groups
 - **35 volumes of handwritten notes**, incl. 295 texts, most of them previously unpublished
 - **Reel-to-reel tapes** (ca. 25 hrs, incl. **7 hrs** with 85 texts)
- Transcribing with the help of native speakers

Workflow for Selkup corpus





- Unreadable characters
- Corrections
- Insertions
- Comments
- Selkup and Russian interference
- Phrase segmentation is not always clear
- ...
- No OCR or handwritten text recognition possible
- Manual text typing



Workflow: FLEx

1.1 Word	Čäptä	los-ira	čö:ti	.
Morphemes	čäptä - ⁰	los-ira - ⁰	čö:ti	
Lex. Entries	čäptä - ⁰ ₂	lo:s-ira - ⁰ ₂	čö:ti ₂	
Lex. Gloss Eng	tale [SG.NOM]	bear [SG.NOM]	about	
Lex. Gloss Rus	сказка [SG.NOM]	медведь [SG.NOM]	о	
Lex. Gram. Info.	n n:num-case	n n:num-case	pp	
Free Rus	Сказка про медведя.			
	silk-s сказка про медведя.			
	Eng A tale about a bear.			
Lit.	čäptä los ira čoty.			
Note	silk-s			
	Rus 1) Разбивка на предложения: в большинстве случаев стоят точки после селькупского и/или русского текста, кроме предложений 2, 3 Кузьминой. 3) Многие слова не соответствуют словарным формам, но они оставлены как есть, за исключением тех случаев, когда соответ...			
	Eng 1) The sentences have been segmented according to the punctuation in original Selkup and/or Russian text, except for the sentences...			
	Kuzmina's signature.			
1.2 Word	Ijal'an	sə:ntirpə:tin	poqin	.
Morphemes	ija -l'a -n - ⁰	sə:ntir -pə: -tin	po - ⁰ -qin	
Lex. Entries	ija -l'a -t ₁ - ⁰ ₃	sə:ntir -mpi ₂ -tīt ₁ +fr. var.	po:₃ - ⁰ ₄ -qin ₁	
Lex. Gloss Eng	child DIM PL [NOM]	play PST.NAR 3PL	space.outside [SG] LOC	
Lex. Gloss Rus	ребенок DIM PL [NOM]	играть PST.NAR 3PL	нство.снаружи [SG] LOC	
Lex. Gram. Info.	n n>n n:num n:case	v v:tense v:pn	n n:num n:case	
Free Rus	Мальчики играли на улице.			
	silk-s мальчики играли на улице.			
	Eng The boys were playing in the street.			
Lit.	'ijalan 'sondyrpodyn 'poqyn.			

- High degree of variability and inconsistency in manuscripts

kəti- 'SAY': 19 variants

possible allomorph,: kətɔ:-, kət-

real variation: kəti-, kəti-/kət-, kəti-, kətə-, kətä-, kəttä-, kətti-, kätti-, qäti-, qəti-, ?kəč'-

Workflow: EXMaRALDA Partitur Editor

ref	NVA_1965_Reindeers_nar.024 (001.024)					
st	‘йматы ‘отанты ‘понтимып тө ‘кирал ‘qolamныт.					
stl	i:mati otanti pontimip to: kiral' qol'amnit.					
ts	Imati ɔ:tantɪ po:ntɪ mip to: kiralqolamnit.					
tx	Imati	ɔ:tantɪ	po:ntɪ	mip	to:	kiralqolamnit.
mb	ima-ti	ɔ:ta-n-ti	po:ntɪ	mi-p	to:	kɪr-al-qolam-ni-t
mp	ima-ti	ɔ:tä-n-ti	po:ntɪ	mi-m	to:	kɪri-ätɔ:l-qoolam-tʃi-ti
ge	wife.[NOM]-3SG	reindeer-GEN-3SG	fell.from.deer's.leg.[NOM]	something-ACC	away	skin-MOM-INF-begin-PRS-3SG.O
gr	жена.[NOM]-3SG	олень-GEN-3SG	камыс.[NOM]	нечто-ACC	прочь	ободрать-MOM-INF-начать-PRS-3SG.O
mc	n-n:case1-n:poss	n-n:case1-n:poss	n-n:case3	n-n:case3	preverb	v-v > v-v:inf-v-v:tense-v:pn
ps	n	n	n	n	preverb	v
SeR	np,h:A	np:Poss		np:P		
SyF	np,h:S			np:O		v:pred
IST	new			accs-inf		
BOR						
BOR-Phon						
BOR-Morph						
CS						
fr	Жена стала олений камыс обдирать.					
fe	His wife began to skin the fell from the reindeer's legs.					
fg	Seine Frau fing an das Fell von den Beinen des Rentiers abzuziehen.					
ltr	жена обдирать стала камысы обдирать стала					
nt						
nto	‘понты – камыс					

- Syntactic functions
- Semantic roles
- Informational structure
- Borrowings and code-switching

Workflow: Metadata (EXMaRALDA Coma)

EXMaRALDA Coma | selkup.coma | Corpus: INEL Selkup Corpus

File Edit View Tools Analysis Maintenance Help

Corpus Data Basket (0)

Communications (187)

Filter
remove all filters

S	Name	1G...
	KR_1969_Witch1_flk	flk
	KR_196X_RavensAndHares2_flk	flk
	KR_196X_Witch2_flk	flk
	KR_IF_196X_RavensAndHares_flk	flk
	KR_IF_KTP_KEP_196X_PlansForTomorrow...	conv
	KTP_196X_Berry_flk	flk
	KuLP_1976_Hospital_nar	nar
	KuMI_196X_Mosquitos_nar	nar
	KuV_196X_Questions_nar	nar
	KVS_1965_Hunt_nar	nar
	MIYA_196X_ILivedInMineevka_nar	nar
	MK_1965_Lgov_transl	transl
	MMS_196X_Lifestory_nar	nar
	NAI_1965_HareParka_flk	flk
	NEP_1965_FoolInSackCoat_flk	flk
	NEP_1965_HareParka1_flk	flk
	NEP_1965_NenetsAndWhiteBear_flk	flk
	NEP_1965_OrphanBoyAndPanOldMan1_flk	flk
	NEP_1965_ThreeBrothers1_flk	flk
	NEP_196X_AjakKoja_flk	flk
	NEP_196X_FirstFairytale_flk	flk
	NEP_196X_HareParka2_flk	flk
	NEP_196X_NenetsAndWhiteBear2_flk	flk
	NEP_196X_OrphanBoyAndPanOldMan2_flk	flk
	NEP_196X_SecondFairytale_flk	flk
	NEP_196X_ThirdFairytale_flk	flk
	NEP_196X_ThreeBrothers2_flk	flk
	NN1_1976_HareParka_flk	flk
	NN2_196X_TrialOfTwoFriends_nar	nar
	NN3_196X_SpeechOnVillageCouncilMeetin...	nar
	NN4_196X_VillageCouncilMeeting_conv	conv
	NST_1965_Tyrshaqo1_flk	flk
	NST_196X_Tyrshaqo2_flk	flk
	NVA_1965_Reindeers_nar	nar

Metadata

Search

Search reset

Communication NVA_1965_Reindeers_nar

Description (Communication)

1 Genre	nar
2a Recorded by	Kuz`mina, Angelina (KuAI)
2b Date of recording	1965
3a Dialect group	Northern
3b Dialect	Taz
3c Subdialect	...
4 Speakers	NVA
5a Transcribed by	KuAI
5b Date of transcribing	...
5c Typed by	BrM
5d Time-aligned by	...
6a Processed by	...
6b Date of processing	...
7a Translation into Russian	KuAI, OSV
7b Translation into English	DO
7c Translation into German	DO
8a Glossed by	OSV
8b Glosses checked	...
9a Annotation SeR	BJ
9b Annotation SyF	BJ
9c Annotation IST	WH
9d Annotation BOR CS	...

Speakers (108)

Filter
remove all filters

S	Sigle	Var
	SML	Sargaeva, Mariya M...
	KPG	Kunina, Polina Gavrilovna
	PL	Polina, Lyuda
	TVP	Ton`kin, Viktor Petrovich
	KAI	Kalin, Aleksej Ivanovich
	KGE	Kalin, Grigorij Egorovich
	KND	Kargacheva, Nina Dmitr...
	KNS	Karsavin, Nikolaj S.
	KNK	Kargacheva, Nina Kuz`...
	KuLP	Kusamin, Leonid Petrovich
	TAV	Tamel`kin, Andrej Vasil...
	NVA	Nerin, Vladimir Aleksand...
	SAIAn	Sargaev, Aleksandr An...
	KIA	Karsavin, Ivan Antonov...
	KMI	Kusamin, Maksim Ivano...
	KMG	Kusamin, Mixail Gavrilov...
	?: MK	Mundakov, Kolya
	NAI	Nerin, Andrej Ivanovich
	NST	Nerina, Svetlana Trofim...
	AZF	Andreeva, Zoya Fyodo...
	KaMN	Katkelev, Makar Nikifor...
	KEP	Kunina, Elizaveta Petro...
	KMA	Kunin, Mixail Alekseevich
	KNM	Kunin, Nikita Mixajlovich
	KPA	Kunin, Pavel Alekseevich
	KPM	Kusamina, Polina Mixajl...
	KTP	Kunina, Tat`yana Petro...
	KuMN	Kuboleeva, Mariya Nikiti...
	MN	Mundakov, Nikolaj Alek...
	MMS	Magichev, Mixail Semyo...
	MVI	Mal`cev, Vasilij Ivanovich
	_MUSTER_201710	...
	NEP	Nerina, Elizaveta Pavlo...
	SAG	Sajgotin, Afanasij Grigo...

Selkup Corpus

Basic data from INEL-corpus

- 60 folklore, narrative and translated texts
- 35 speakers
- 3321 sentences
- 19511 tokens

Additional data from Selkup Spoken Language Corpus (SELSLC)

	texts	sent	tokens
SELSLC northern	26	1410	7814
SELSLC central	48	3549	22130
SELSLC southern	66	4018	23422

Grammar and corpus data

- There exist several description of Selkup, including a thorough traditional grammar [Kuznetsova et al. 1980].
- Focus of the grammar: morphology and phonology, much less information on syntax, including complex clauses.
- Restricted access to native speakers → we can make use of corpus data to fill in the gaps.
- Typology offers parameters to deepen the existing description, which can confirm the predictions of typological classifications or single out language specific features.
- Case study: purpose constructions

Purpose clauses: infinitive

(1) Täpi-t nɔ:kir timn'ä-si-k mora-t tə:
he-GEN three brother-CRC-PL sea-GEN behind

na säri qorqi-p **qät-qa** qən-nɔ:-tit.
this white **bear-ACC** **kill-INF** go.away-CO-3PL

Three of his brothers came across the sea to kill this white bear.

(2) **Am-i-r-qa** sola-p me:-qil-li-mit.
eat-EP-FRQ-INF spoon-ACC make-DRV-OPT-1PL

We will make a spoon to eat.

Purpose clauses: converb

(3) Mat mač'ä qon'i-šä-k **su:ri-l'-l'ä.**
 I to.the.forest go.away-PST-1SG.S **wild.animal-VBLZ-CVB**
I went to the forest to hunt.

(4) Markin'č'a tɔ:ptij č'e:li nɔ:t
 Markyncha.[NOM] next day.[NOM] then
 na qən-mi-nti imaqota-m **pe:lä.**
 INFER leave-PST.NAR-INFER.[3SG.S] **old.woman-ACC look.for-CVB**
On the next day Markyncha went to look for the old woman.

Distribution of infinitives vs. converbs in purpose constructions?

Distribution of infinitives and converbs

One of the main questions concerning subordinate constructions, including purpose clauses, is the question of syntactic boundedness [Van Valin 2005, Schmidtke-Bode 2009]. Though both infinitival and converb constructions demonstrate quite high degrees of unity, their behavior differs concerning two main parameters of linkage: semantic unity and expression of arguments.

Parameters to explore:

- Semantics of matrix verb
- Actionality of subordinate clause
- Presence of verb dependents (arguments or adjuncts)
- Position of purpose clause

1. Semantics of matrix verb

The infinitives have a wide distribution in purpose constructions co-occurring with matrix verbs of different semantic classes:

- (2) **Am-i-r-qo** sola-p meɜ-qil-li-mit.
 eat-EP-FRQ-INF spoon-ACC make-DRV-OPT-1PL

We will make a spoon to eat.

Usage of converbs is restricted to the context of motion verbs, mostly intransitive: *go, leave, come, sit*. Only in one example a transitive predicate occurred:

- (5) **Nami-m-ti** **karn'-al-ti-la** ɔ:ta-m-ti-p
 this-ACC-3SG **scrape-MOM-TR-CVB** reindeer-ACC-3SG-ACC
 nił'č'i-k moqona tɔ:qqi-q-ol'im-i-mpa-tij.
 such-ADVZ home shepherd-INF-be.going.to-EP-PST.NAR-3PL

To scrape them they were going to shepherd the reindeers home.

2. Actionality of subordinate verb

The infinitives can occur with verbs of different actional classes, the converb form is used mainly with verbs of activity.

- (1) Täpɪ-t nɔ:kɪr timn'ä-si-k mora-t tə:
 he-GEN three brother-CRC-PL sea-GEN behind
na säri qorqi-p qät-qa qən-nɔ:-tit.
this white bear-ACC kill-INF go.away-CO-3PL

Three of his brothers came across the sea to kill this white bear.

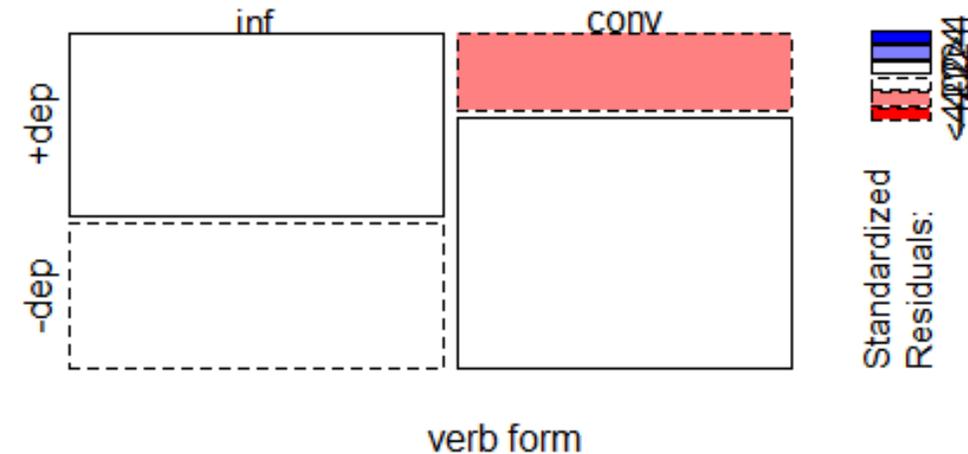
- (3) Mat mač'ä qon'i-šä-k **su:ri-l'-l'ä.**
 I to.the.forest go.away-PST-1SG.S **wild.animal-VBLZ-CVB**

I went to the forest to hunt.

3. Dependents

While infinitival constructions may have overtly expressed arguments or adjuncts, there is a tendency for converb clauses to have no arguments or adjuncts, except for DOs.

	+dep	-dep
infinitive	35	28
converb	13	43

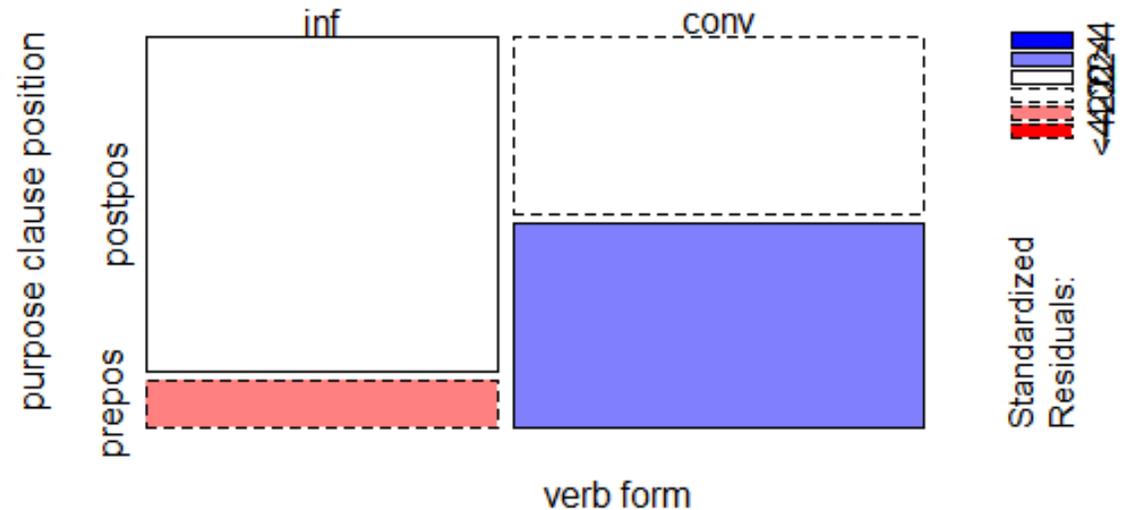


X-squared = 11.576, df = 1, p-value = 0.000668

4. Position

While infinitival clauses tend to occur in postposition to the main clause, converb constructions quite often occur in the preposition as well.

	post	pre
infinitive	42	6
converb	26	30



X-squared = 17.491, df = 1, p-value = 2.886e-05

Distribution of infinitives and converbs: tendencies

- Infinitives are used with matrix verbs of various semantics, can have various arguments and adjuncts, tend to be posed after the verb.
- The usage of converbs is restricted to motion-cum-purpose contexts, they tend to have no adjuncts and are more frequently used in clause-internal position.
 - Converb clauses demonstrate higher degree of unity in motion-cum-purpose constructions than the infinitival ones. Besides, the infinitive form presupposes a potential non-realised event, which goes in hand with the cross-linguistic definition of purpose in [Schmidtke-Bode 2009]: the purpose constructions contain an intention component. Whereas converb is a specific form for motion-cum-purpose context and builds a non-compositional construction which enforces the purpose interpretation.

What makes converbs appear in motion-cum-purpose construction?

Prototypical functions of INF and CVB in Selkup

Infinitives are used in complement clauses with all kind of matrix verbs:
phasal, modal, mental...

Converbs are prototypically used for coordination and **secondary predications of manner**:

(6) Ni:nɨ konna **kuri-l'ä** na tü-nta.
 then upwards **go-CVB** INFER come-INFER.[3SG.S]

Then he came upwards walking.

(3) Mat mač'ä qon'i-šä-k **su:ri-l'-l'ä.**
 I to.the.forest go.away-PST-1SG.S **wild.animal-VBLZ-CVB**

I went to the forest to hunt.

This last meaning could be the source for purpose-usages of converbs.

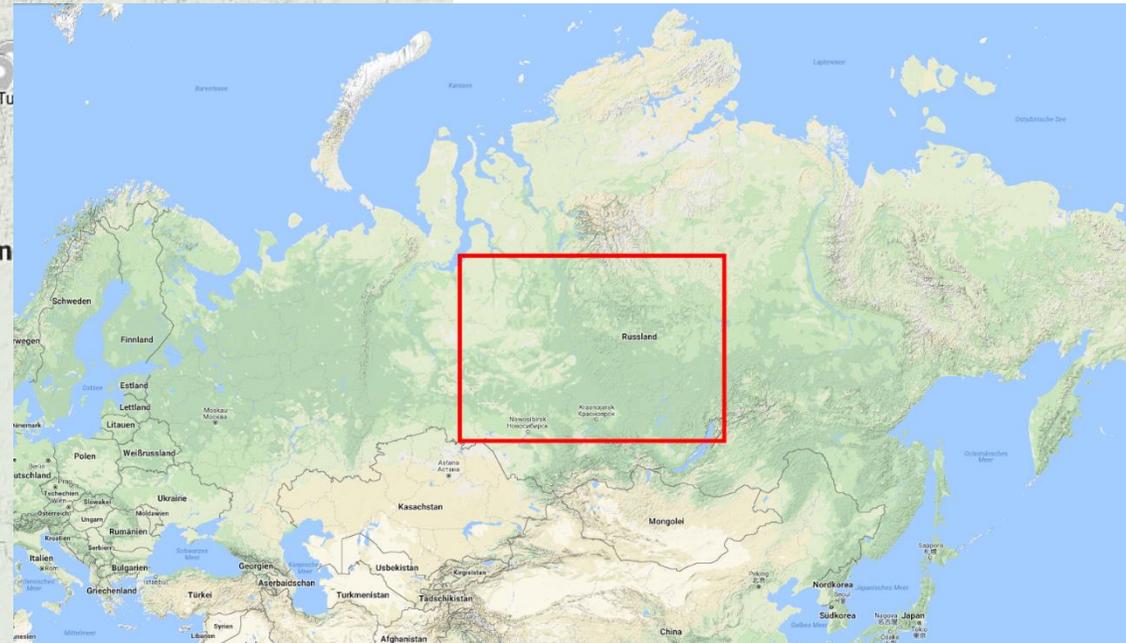
Dialectal variation

The distribution of infinitive vs. converb constructions in purpose clauses proves to be stable for **Northern Selkup** across various text collections and in time from 1930-s to 2000-s. We have checked our results on the text collections from SELSC, [Kuznetsova et al. 1993], [Tuchkova, Wagner-Nagy 2015], 4 texts from Varkovitskaya archive, SIBERIAN LANG (project supervised by O.Kazakevich).

☒ According to the data of SELSC, converb constructions are almost not present at all in **Central Selkup** dialects: we have found only 1 example with converb in 25 contexts, where we would expect converb construction, according to Northern Selkup data.

? In **Southern Selkup** dialects we observed variation of infinitive and converb constructions in “northern mode”.

Selkup





INEL Project „Grammatical Descriptions, Corpora and Language Technology for Indigenous Northern Eurasian Languages“: <https://inel.corpora.uni-hamburg.de>

SELSLC Project “Selkup Spoken Language Corpus“: <https://www.slm.uni-hamburg.de/selslc.html>

SIBERIAN LANG Project “Minority languages of Siberia as our cultural heritage“: <http://siberian-lang.srcc.msu.ru/en/about>

Kuznetsova A.I., Helimsky, E.A., Grushkina, E.V. (1980). *Ocherki po seljkupskomu jazyku. Tazovskij dialekt. Tom 1.* Moskva. (in Russian)

Kuznetsova A.I., Kazakevitch O.A., Ioffe L.Ju., Helimsky, E.A., (1993). *Ocherki po seljkupskomu jazyku. Tom 2. Teksty. Slovarj.* Moskva. (in Russian)

Schmidtke-Bode, K. (2009). *A typology of purpose clauses* (Vol. 88). John Benjamins Publishing.

Tuchkova, N.A., Wagner-Nagy, B. (2015). «Семь Богов мудростью обладающий Итте...»: *тексты с героем Итя в селькупском фольклоре.* Tomsk. (in Russian)

Van Valin, R. D. (2005). *Exploring the syntax-semantics interface.* Cambridge University Press.

Thank you!

<https://inel.corpora.uni-hamburg.de/>