



KONEEN SÄÄTIÖ
KONE FOUNDATION

NORTHWARD SPREAD OF SAMOYEDIC

Olesya Khanina & Kaisla Kaheinen
(University of Helsinki)

1. INTRODUCTION



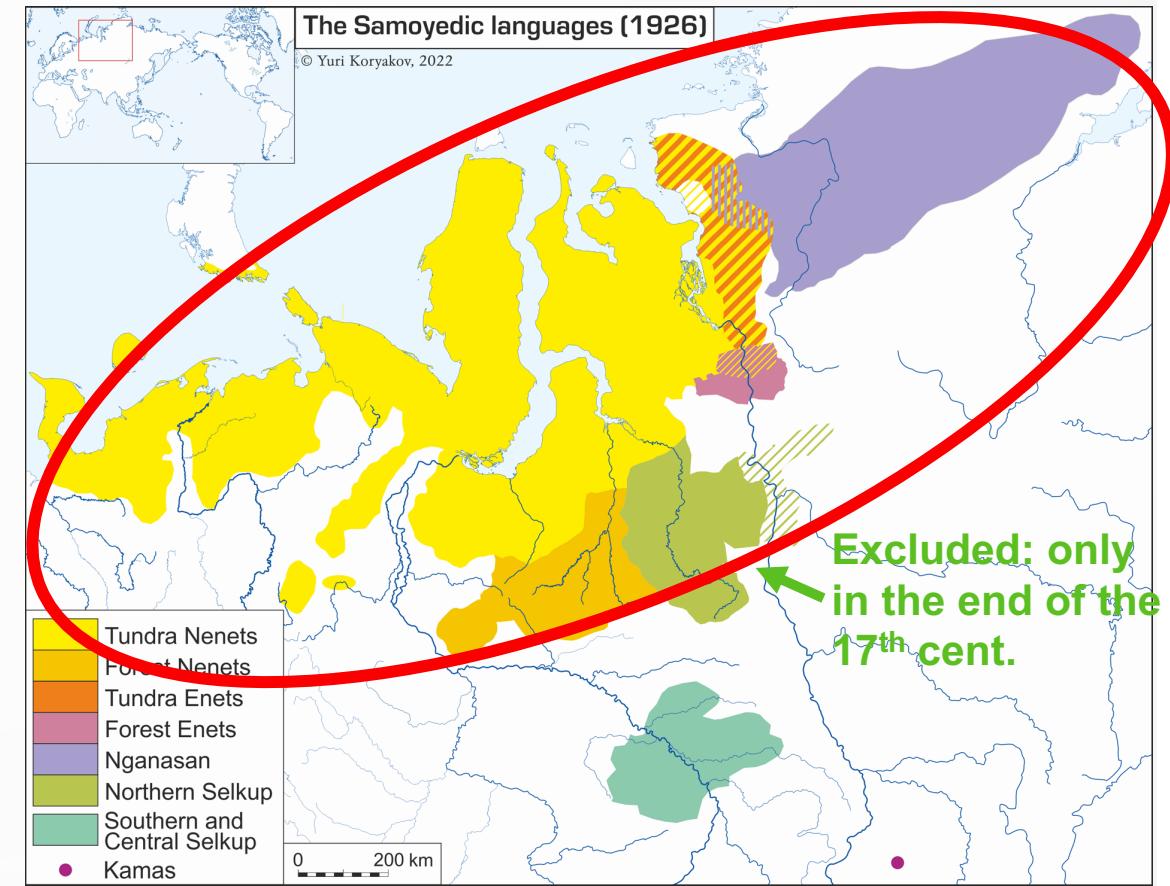
NORTHERN SAMOYEDIC LANGUAGES

Samoyedic

- Nenets (Tundra, Forest)
- Enets (Tundra, Forest)
- Nganasan
- Selkup (Northern, Central-Southern, Ket')
- Kamas (in the Sayans)
- Mator (in the Sayans)

5 languages
spoken in the
north

Proto-Samoyedic: Southern Siberia, ca. 2000 y. a.





SAMOYEDIC: INTERNAL STRUCTURE

- 6 possible genealogies!
- we assume a rake model with only 3 subgroups:
 - the two Nenets definitely form a group,
 - the two Enets definitely form a group,
 - Nenets and Enets most probably form a group.



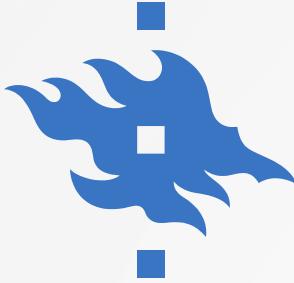
NORTHERN SAMOYEDIC LANGUAGES

Samoyedic

- Nenets (**Tundra, Forest**)
- Enets (**Tundra, Forest**)
- Nganasan
- Selkup (Northern, Central-Southern, Ket')
- Kamas (in the Sayans)
- Mator (in the Sayans)

**5 languages
spoken in the
north**

**No common phonetic
innovations, but numerous
lexical and morphological
innovations**



OUTLINE

1. Introduction
2. NS stock in the lexicon: common retentions and borrowings from non-Uralic
(Kaheinen, Accepted)
3. Changes in NS phonology (Gusev & Khanina, In prep)
4. NS morphological isoglosses (Gusev & Khanina, In prep)
5. Summary: what we now know and what we do not know
6. A scenario of the NS diversification and of the northward spread

2. LEXICON

Northern Samoyedic stock in the lexicon

Common retentions and borrowings
from Non-Uralic

Based on the doctoral dissertation *Etymologia ex silentio:
nganasanin äännehistoria ja kielikontaktit* (Kaheinen, Accepted)

Methods & data

- **Historical-comparative method** applied to lexical data
- Nganasan lexemes gathered from all major attainable sources, as well as archival sources
- Nganasan data was compared to that of other Samoyed languages, as well as etymological sources from Proto-Samoyed, its contact languages, and further Uralic
- Special focus on Nganasan-Enets-Nenets correspondences, which were the greatest in number

Nganasan sources

Abbr.	Source	Details
СНг	Костеркина, Момде & Жданова 2001: <i>Словарь нганасанко-русский и русско-нганасанский</i> .	School dictionary containing approx. 6000 words, though a large number of them derivates, with grammatical information.
Ng20	Kaheinen 2020: <i>Nganasanin korpuusalista</i> .	List of words contained in the NSL corpus (Brykina et al. 2018), compiled by Kaheinen (2020), with available grammatical information. Unpublished.
NgSI	Хелимский, Е. А.: <i>Нганасанский словарь</i> .	List of words compiled by Eugene Helimski, with some grammatical information. Unpublished.
KS	Kortt & Simčenko 1985: <i>Wörterverzeichnis der nganasanischen Sprache. Teil 1: nganasanisch-deutsch-russisches Glossar</i> .	List of words gathered as part of ethnographic study on Nganasan. Major inaccuracies but possibly valuable information on single lexemes not found elsewhere.
Pop.mskr.	Попов, А. А.: <i>Нганасанко-русский словарь</i> . Биб. карточки. Автограф Прокофьевой Е. Д. (Фонд 14 опись I: 68).	Collection of archival cards, containing approx. 4000 lexemes, located at the Kunstkamera archive in St. Petersburg. Studied by Kaheinen in 2021. Additional data from Попов 1940 and Popov 1966.
C	Castrén M. A. 1855: <i>Wörterverzeichnisse aus den samojedischen Sprachen</i> .	Data from the earliest linguistic expedition on Nganasan. Valuable for Old Nganasan forms and few obsolete lexemes.
G.mskr.	Valentin Gusev: <i>Nganasan dictionary</i> .	Word list for a dictionary in preparation, mostly containing lexemes attested in other sources, sometimes with additional grammatical data. Unpublished.

Nganasan sources

Abbr.	Source	Details
СНг	Костеркина, Момде & Жданова 2001: <i>Словарь нганасанко-русский и русско-нганасанский</i> .	School dictionary containing approx. 6000 words, though a large number of them derivates, with grammatical information.
Ng20	Kaheinen 2020: <i>Nganasanin korpuusalista</i> .	List of words contained in the NSL corpus (Brykina et al. 2018), compiled by Kaheinen (2020), with available grammatical information. Unpublished.
NgSI	Хелимский, Е. А.: <i>Нганасанский словарь</i> .	List of words compiled by Eugene Helimski, with some grammatical information. Unpublished.
KS	Kortt & Simčenko 1985: <i>Wörterverzeichnis der nganasanischen Sprache. Teil 1: nganasanisch-deutsch-russisches Glossar</i> .	List of words gathered as part of ethnographic study on Nganasan. Major inaccuracies but possibly valuable information on single lexemes not found elsewhere.
Pop.mskr.	Попов, А. А.: <i>Нганасанко-русский словарь</i> . Биб. карточки. Автограф Прокофьевой Е. Д. (Фонд 14 опись I: 68).	Collection of archival cards, containing approx. 4000 lexemes, located at the Kunstkamera archive in St. Petersburg. Studied by Kaheinen in 2021. Additional data from Попов 1940 and Popov 1966.
C	Castrén M. A. 1855: <i>Wörterverzeichnisse aus den samojedischen Sprachen</i> .	Data from the earliest linguistic expedition on Nganasan. Valuable for Old Nganasan forms and few obsolete lexemes.
G.mskr.	Valentin Gusev: <i>Nganasan dictionary</i> .	Word list for a dictionary in preparation, mostly containing lexemes attested in other sources, sometimes with additional grammatical data. Unpublished.

Theoretical and terminological considerations

Contact situation	Type of contact influence	Results
Retention OR language shift	Influence by bilingual speakers	Translation loans and possibly PAT borrowing, sometimes leading to increased complexity Phonological change only in intense contact
Language shift by adult population	Influence by second language speakers	Phonological change early on Simplification of grammatical structures, both PAT and MAT borrowing (speciality vocabulary)

Classification of contact situations and likely outcomes by Ross (2013: 36).

Condition	Argument	Basis
Large number of lexemes without etymology	<i>Urschöpfung</i> hypothesis	Words do not just appear out of thin air.
Phonotaxis incompatible with proto language	Phonotactic argument (<i>terminus post quem</i>)	Phonotactically aberrant words cannot go back to proto language.
Irregular correspondences between related languages	Spread argument (<i>terminus post quem</i>)	No reconstructable proto form.
Limited or topical spread in daughter languages or dialects	Spread argument	Spread pointing to late local loans.
Certain semantic fields overrepresented	Semantic group argument	Words forming a semantic group lessens the chance of coincidence.
Semantically opaque toponyms	–	Likely to be borrowings from a previous population. Contacts locatable based on names.

Main arguments used in the study of lexical layers of unknown origin and language shift hypothesis (classification according to Junntila 2015: 136–151).

Condition	Argument	Basis
Large number of lexemes without etymology	<i>Urschöpfung</i> hypothesis	Words do not just appear out of thin air.
Phonotaxis incompatible with proto language	Phonotactic argument (<i>terminus post quem</i>)	Phonotactically aberrant words cannot go back to proto language.
Irregular correspondences between related languages	Spread argument (<i>terminus post quem</i>)	No reconstructable proto form.
Limited or topical spread in daughter languages or dialects	Spread argument	Spread pointing to late local loans.
Certain semantic fields overrepresented	Semantic group argument	Words forming a semantic group lessens the chance of coincidence.
Semantically opaque toponyms	–	Likely to be borrowings from a previous population. Contacts locatable based on names.

Main arguments used in the study of lexical layers of unknown origin and language shift hypothesis (classification according to Junntila 2015: 136–151).

Results

	SW (Katzschmann 2008: 30)	SW, Updated data
Lexemes, underived total	647	646
Nganasan cognates total	409	473
Percentage of items having Nganasan cognate	63,2%	73,2%

Retention rates for PS lexemes in Nganasan, old data compared with new

	Regular	Irregular	Nganasan spread only	Total
Tundra Nenets	159	117	0	276
Forest Nenets	126	87	0	213
Tundra Enets	178	153	0	331
Forest Enets	129	100	0	229
Nganasan	211	173	638	1023

Nganasan lexemes according to spread, data from Kaheinen (Accepted).

Examples: regular correspondences

PS	Ng	TE	FE	TN	FN
*åjwu 'suitor'	CHr ηojbi : 3SG.PX -dü	EnSI ai	ЛЭС ej	T65 нэв (ηæw°)	L56 21a ηäw°
*aŋərā 'flesh side of hide'	CHr náŋeru : -tu	EnSI eara	ЛЭС naara	T65 ня́ра (nya°ra)	L56 304b nya°la ~ nya°ra (Salminen 2007: 370)
*írwå 'loin'	KS ηírbu	EnSI ubadi	—	—	—
*járaptut 'pack ice'	CHr jorübtüq : GEN.SG jorübrüde : -tü	EnSI dìaratu?	--	-	—
*jarwz- 'blink (rapidly), frown'	NgSI jarbü̯s- (U)	—	--	T65 ярволабта(сь) (yarwoləbta-)	—
*jolsä (v. *jolsä-) 'whole' (v. 'to end')	CHr jülsï- : AOR.PRF.3SG jülsïqe (I)	-	ЛЭС diɔsi- ~ diɔzi-	T65 ёльце (yolcye)	БВ дёльш (yołsy°)
*puntə(-) 'belief, (to) trust'	CHr hunte : hunde : -du	EnSI pudiru-	—	T65 пун (pun°)	—
*pursoj 'aorta, carotid artery'	NgSI härṣue : -cü	EnSI pud'u	—	T65 пурцо (purco)	L56 366a pülsyu
*töntär 'back, side of sledge'	NgSI тундыр (I)	EnSI tôde? : PL tôderu?	EnSI tôde?	T65 тёндер" (tyondyer)	L56 518b tyontyer
*wíet- 'to cut oneself (of shaman, during a ritual)'	Ng20; NgSI bïqtə- : AOR.IPF.3SG -tī	EnSI DRV. buutode	—	T65 вы́г(д) (wíq)	—

Examples: irregular correspondences

- A few cases of recurring irregularities, e.g.

- a ~ a

Ng		TE	FE	TN	FN	comments
CH _r baruši (U)	'devil'	EnSI barut'i	ЛЭС bari	T65 варючи (waryucyi°)	–	a ~ a
CH _r batüqe (U)	'tail(bone)'	EnSI batu?o	ЛЭС bato?ɔ ~ bəto?ɔ	–	–	a ~ a
Ng20 laheru	'leaf'	EnSI laboro	EnSI loboro	–	–	a ~ a (~ FE o)

- a ~ (*)ə

Ng		TE	FE	TN	FN	comments
CH _r kasa- (U)	'almost' (v.aux)	EnSI kot'e-	ЛЭС ketʃi-	T65 хäця(сь) (хəсуа-)	L56 160b käqsyä-	a ~ (*)ə
CH _r lakaräräiq	'sudden(ly)'	EnSI lokorí	ЛЭС ləkuri?	T65 лäкри' (lək°ryih)	L56 215b läk°ryi°	a ~ (*)ə; alleged ~ tung ПТМ *lak- 'близко, скоро; кстати, как раз; сразу' (TMC I 487– 488; 515) (АнХел 2007: 112)
Ng20 tarar- (U)	'to jump, to dance (of shaman)'	EnSI toraro-	ЛЭС təra-	T65 тäра(сь) (təra-)	L56 462b täla-	a ~ (*)ə

Examples: irregular correspondences

(Cont.)

- (non-alternating) s (< *t³s) ~ s (< *s)

Ng	TE	FE	TN	FN	comments
CHr susu (U) 'small hill'	EnSI sôse?o	ЛЭС sɔʃi	-	-	No gradation: *t3s ~ *s
CHr l'išü (U) 'lazy'	EnSI lisane?o	ЛЭС leʃo	T65 лэсё (læsyo)	БВ лышу (l̥syu)	No gradation; vocalism

- word-internal l(') ~ l' (Nganasan-Enets)

Ng	TE	FE	TN	FN	comments
NgSI tel'ibti- (l) 'to ask for sth.'	EnSI tolíte-	(?) ЛЭС tɔ?ir-	(?) T65 тэла(сь) ям. (tæla-) 'повторить за кем-л.'	-	l' ~ l'
Ng20 torul'a 'reindeer calf'	EnSI torúlea	(?) ЛЭС tɔrgu 'внутреннее высушенное оленье сало'	-	-	o ~ o; l' ~ l'; TE < Ng

Wanderwörter

Ng		TE	FE	TN	FN	comments
CHr auku (U)	'pet reindeer'	EnSI auku	ЛЭС auka	T65 ӈавка (ӈawka)	БВ ӈавка (ӈawka)	Хелимский [1994] 2000: 321; a ~ a
CHr ba(a)ӈka(da)qa (U)	'reindeer that cannot give birth'	EnSI bagoði	ЛЭС bagizi	T65 ванѓды (waŋk°di)	БВ ванҝат (waŋkat°)	АнХел 2007: 120; a ~ a
NgSI turuse	'plane (tool)'	EnSI turuso	СБ турус	T65 түрос, түрас (turos° ~ turas°)	L56 497b tūlus°	vocalism

Possible internal loans

Ng		TE	FE	TN	FN	comments
Ng20 śunedi	'strong reindeer'	EnSI śunoði	--	T65 сюнды (syun°di°)	L56 455b syuj°ti°	Ng < TE; *-təjə > Ng -Tee
CHr śüsüede	'finger, joint'	EnSI śuśuoðo	-	-	-	TE < Ng
Ng20 baade	'until' (adv.)	EnSI baaðo	EnSI baaða	-	-	TE < Ng; a ~ a

Internal borrowing unlikely

Ng		TE	FE	TN	FN	comments
CHr ledi- (I)	'not able' (v.aux)	EnSI leʔi-	ЛЭС ləzis-	-	-	vocalism
Ng20 kaaje	'bellows'	EnSI siajo	EnSI kejo	-	-	k ~ s
CHr lunsï (I)	'spleen'	EnSI lísiko	EnSI l'iš'iku	T65 люси (lyusyi)	? Поп78 łỹsyu	voc.and cons.

Alleged Tungus cognates

Ng		TE	FE	TN	FN	comments
KS cil'i	'root or branch of reindeer horn'	EnSl t'ire	EnSl nado t'ire	T65 тилюй (tyilyuy°)	L56 512a tyilye	alleged. Sam. < Evk. дил 'голова' (TMC I: 205–206) (АнХел 2007: 101) 'основание рогов' (cf. TMC I: 206); semantic connection apparently only to Manchu (?)
Ng20 śimka ~ śimkej	'tent pole'	–	--	T65 сымзы (símcio°); L56 446b syimci°	L56 446b syimsu	k ~ s; > Evk. чымка, кымка (TMC II: 394; Bac. 1969: 109); cf. АнХел 2007: 137: "PNS" *kimk-(suff.) 'шест'; *kimk8'id. – probably anachronistic
CHr hiadee (U/I)	'forehead'	EnSl pea	ЛЭС рёja	T65 пэ́я (peya)	Поп78 100 rёja	*räjå ("PNS") 'лоб' (АнХел 2007: 126) ~ Tung. *rēje (TMS II: 361); Sam. vocalism irregular
Ng20 lebsue (U)	'tail (of fish)'	–	–	T65 ябцо (yabco)	L56 102a japso : acc.pl jipsu	e ~ a; ~ (< ?) Tung.; Evk. лайчача ~ Ev. н'апча~(TMC I: 494) (АнХел 2007: 124) – based on consonantism Sam. > Evk. more likely; *l ~ *j
NgSl l'ibtiq : l'ibtide (I)	'fur sock'	EnSl íutu?	–	T65 либт"(д) (lyibt°q)	–	first syll. TE *ü ~ TN *i – АнХел (2007: 82) "PNS" *l'iptət ~ ? Tung. Nan ли́птоко/u ~ ли́фтоко/u (TMC I: 499), more likely < Rus. dial. липты мн. 'исподние самоедские сапоги или чулки, пыжиковые, шерстью внутрь; липты и др.' (Аникин 2000: 362)

Words that occur only in Nganasan

No.	Word	Source	Meaning	Comments
87	hüqä	Ng20	'gills'	
110	hond'ir : honśire	CHr	'(bone) ledge'	
111	hone	CHr	'hunger'	
112	honsu	CHr	'bed'	comp. PS *påjnå
200	jonsie	NgSI	'reindeer with one horn shorter'	
201	joraku(-)	Ng20	'bent, crooked'	
202	jose-	NgSI	'(to be) opaque (of water)'	
279	kurasigümü (?)	KS	'frame of drum or net made of wood or mammoth bone'	
280	kusegee ~ küsegee	KS	'reindeer with horns close together'	
281	kutarutue (?)	KS	'punishment'	
282	küeje	CHr	'lead'	
283	künjal'e-btu-	Ng20	'to give birth'	
364	monti	Ng20	'reindeer unable to give birth'	
365	mońši	NgSI	'highland; mountain far away from water'	
454	ŋejbeu-	CHr	'to work'	
455	ŋeli-	CHr	'to win'	
456	ŋeq	CHr	'shaman'	
457	ŋ'ia	CHr	'spit'	
601	tome-	NgSI	'to chew'	comp. PS (SW) *tət ³ wå- 'kauen'; comp. TE EnSI таδи? 'то, что торчит в зубе'; таδиид'e? 'зубочистка' ~ TN T65 тадэ'(н) (tadeh) 'застрявший в зубах кусочек пищи' < *tåtэн
602	tond'ükü	CHr	'lemming'	
603	tonśümue	Ng20	'area between two mountains'	

Words that occur only in Nganasan

Only 29 out of 638 clearly phonotactically aberrant lexemes:

berkīti- ‘cry out’; *halŋī koli* ‘small fish’ (~ *halmī*); *horkī-* ‘to bubble (sound)’; *hübīki* ‘curious’; *jabkude-* ‘to slap sth. loudly’; *karkītu-* ‘be jumpy’; *kebkude-* ‘to jump’; *kalhuj-* ‘be covered with fat’; *kejki-* ‘lean against sth.’; *kelhe-* ‘condensate’; *koagel-* ‘to mix, to change subject’; *kīabtu-* ‘to encrust’; *kilmuru-* ‘pretend to be sth.’; *kirugu* ‘avalanche’; *kiesie* ‘hand bail’; *käbtu* ‘short-legged dog’ (? → *käbtusi* ‘flatness’); *känterue-* ‘laugh out loud’; *kärku-* ‘to jump around’; *lebku-* ‘to agree’; *lebkur* ‘canopy’; *l'ebku-* ‘to choke’; *mobki-* ‘to crumble’; *narkikue* ‘ripped apart’; *nubku-* ‘to run (of animal)’; *selŋī* ‘pillow’ (? DRV. → *selŋibti-* ‘to hide sth.’); *śibkü* ‘girth’; *tebkel'i-* ‘to be tight or suffocating’; *tukie* (*tukie håa*) ‘thicker spot on a runner’; *celhe-* ‘to collapse’

Words of unknown origin: what can we know?

- **638** uniquely Nganasan lexemes – A fairly large number but less than expected
- Only a fraction with aberrant phonotaxis
- Divided evenly between semantic fields
- No clearly identifiable "substrate" – that is, a **loanword stratum** separate from words that *could* go back to Proto-Uralic but only *happen* to not have known cognates.
- Words of unknown origin better to be left **unclassified for now**.

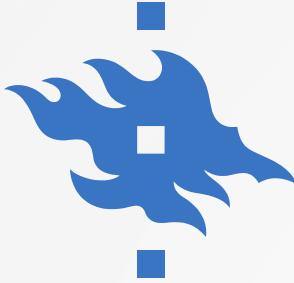
(Kind of a) Conclusion

- **1023** Nganasan lexemes with spread restricted to northern Samoyed or recent contact languages (Evenki, Dolgan)
- Out of these **638** unique to Nganasan
- **384** with at least one cognate elsewhere in northern Samoyed
 - 211 with regular correspondences → probably Proto-Samoyed
 - 173 with irregular correspondences → post-Proto-Samoyed borrowings
- Probable internal loans (often direction unknown) and *Wanderwörter* (ca. 20?)
- In some cases internal borrowing seems impossible, however
- Unclear cognates esp. in Tungusic languages (**more research is needed**)
- **No substantial identifiable strata!**

References

- Castrén M. A. 1855: *Wörterverzeichnisse aus den samoqedischen Sprachen*.
EnSI = Е. А. Хелимский: Энецкий словарь. Manuscript.
- Gusev, Valentin: *Nganasan dictionary*. Manuscript.
- Junttila, Santeri 2015: *Tiedon kumuloituminen ja trendit lainasanatutkimuksessa. Kanta-suomen ba-ttilaislaijien tutkimushistoria*. Doctoral dissertation. University of Helsinki.
- Kaheinen, Kaisla 2020: *Nganasanin korpu-sanalista*. Manuscript.
(Accepted): *Etymologia ex silentio: nganasanin äännehistoria ja kielikontaktit*. Doctoral dissertation. University of Helsinki.
- Kortt, Ivan R. & Simčenko, Jurij B. 1985: *Wörterverzeichnis der nganasanischen Sprache. Teil 1: nganasanisch-deutsch-russisches Glossar*.
- Ross, Malcolm 2013: 2013: Diagnosing contact processes from their outcomes: the importance of life stages. *Journal of Language Contact* 6 (2013). 5–47.
- SW = Janhunen, Juha 1977: *Samoqedischer Wortschatz. Gemeinsamoqedische Etymologien*. Helsinki.
- Аникин, А. Е. 2000: *Этимологический словарь русских диалектов Сибири. Заимствования из уральских, алтайских и палеоазиатских языков*. 2-е издание, исправленное и дополненное. Москва: Наука.
- АнХел = Аникин, А. Е. & Хелимский, Е. А. 2007: *Самодийско-тунгусо-маньчжурские лексические связи*. Москва: Языки славянской культуры.
- БВ = Бармич, М. Я. & Вэлло, И. А. 2002: *Словарь ненецко-русский и русско-ненецкий (лесной диалект)*. СПб: Просвещение.
- Вас. = Василевич, Г. М. 1969: *Эвенки. Историко-этнографические очерки (XVIII — начало XX в.)*. Ленинград: Наука.
- Костеркина, Момде & Жданова 2001: *Словарь нганасанско-русский и русско-нганасанский*. Под ред. В. Ю. Гусев. СПб: Просвещение.
- ЛЭС = Ханина, Олеся & Шлуйнский, Андрей: *Лесной энецкий словарь*. Manuscript.
- Попов, А. А.: *Нганасанско-русский словарь*. Биб. карточки. Автограф Прокофьевой Е. Д. (Фонд 14 опись I: 68).
- СБ = Сорокина, И. П. & Болина, Д. С. 2001: *Словарь энецко-русский и русско-энецкий*. СПб: Просвещение.
- ТМС I-II = Цинциус, В. И. (отв. ред.) 1975: *Сравнительный словарь тунгусо-маньчжурских языков*. Том. I-II. Ленинград: Наука.
- Т65 = Терещенко, Н. М. 1965: *Ненецко-русский словарь*. Москва: Советская энциклопедия.
- Хелимский, Е. А.: *Нганасанский словарь*. Manuscript.

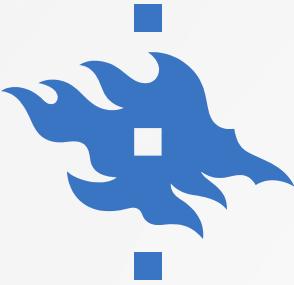
3. PHONOLOGY



3. CHANGES IN NS PHONOLOGY

Proto-Uralic and Proto-Samoyedic phonologies are reasonably well reconstructed.

- List of changes 'phonological segments of PS > modern NS languages' (64 items)
 - Gusev 2022, Helimski 1993, 1997, 2005, Janhunen 1975-6, 1977, Mikola 2004, Salminen 2012;
 - Constellations of NS languages where a change have occurred.
 - Some reservations:
 - what counts as a change depends on a particular reconstruction of PU, PS;
 - different cross-linguistic frequencies, different probabilities of independent innovations;
 - different weight in the phonological system > morpheme-recognizability;
 - we omit recent changes in one language only (the last 200 years);
- ⇒ the contrasts in numbers below are only *rough estimates* of what is going on**



3. CHANGES IN PHONOLOGY: DATA

Some snapshots from the table

	FN	TN	FE	TE	Ng	
11 *k > x before back vowels intervocally	+	+	+	+	-	FN+TN+FE+TE
12 *k > ś before front vowels	+	+	+	+	+	all
13 *w > b word-initially	-	-	+	+	+	FE+TE+Ng
14 *w > j before front vowels word-initially	-	+	-	-	-	TN
15 *w > 0 intervocally	-	-	+	+	+	FE+TE+Ng
16 *w > b' before front vowels intervocally	-	+	-	-	-	TN
17 *j > 0 in some contexts/words intervocally	-	-	+	+	+	FE+TE+Ng
30 *NT > ND	-	+	-	-	-	TN
31 *NT > D	-	-	+	+	-	FE+TE
32 *nt > n	+	+	-	-	-	FN+TN
33 *ns, *ms, *rs > s, z (both variants)	-	-	+	-	-	FE
34 *rs > rc	-	+	-	-	-	TN
35 *ns, *ms, *rs > d'	-	-	-	+	-	TE
50 *ə > ɔ	-	-	+	+	-	FE+TE
51 *å > a	+	+	+	+	-	FN+TN+FE+TE
52 *å > o in the 1st syllable	-	-	-	-	+	Ng
53 *å > u in non-1st syllables	-	-	-	-	+	Ng
54 *e > i after non-labial consonants	-	-	+	+	-	FE+TE
55 *j > i after non-labial consonants	-	-	-	-	+	Ng
56 *e, j > u after labial consonants	-	-	+	+	-	FE+TE

FN+TN+FE+TE	8
FN+TN	5
FE+TE	10
FE+TE+Ng	5
TN+FE+TE	3
FN+TN+Ng	2
FN+TN+FE+Ng	1
FN+FE+TE	1
all	7
Ng	10
TN	8
FN	2
FE	1
TE	1
	64



3. CHANGES IN PHONOLOGY: DATA

Setting up the genealogical tree: 23

Setting up the genealogical tree: 23

Some snapshots from the ~~Unexp~~ but expe

	Lexical Outputs from the CVC					
	*	*	*	*	*	
11 *k > x before back vowels intervocally	*	*	*	*	-	FN+TN+FE+TE
12 *k > s before front vowels	*	*	*	*	*	all
13 *w > b word-initially	-	-	*	*	*	FE+TE+Ng
14 *w > j before front vowels word-initially	-	*	-	-	-	TN
15 *w > 0 intervocally	-	-	*	*	*	FE+TE+Ng
16 *w > b' before front vowels intervocally	-	*	-	-	-	TN
17 *j > 0 in some contexts/words intervocally	-	-	*	*	*	FE+TE+Ng
30 *NT > ND	-	*	-	-	-	TN
31 *NT > D	-	-	*	*	-	FE+TE
32 *nt > n	*	*	-	-	-	FN+TN
33 *ns, *ms, *rs > s, z (both variants)	-	-	*	-	-	FE
34 *rs > rc	-	*	-	-	-	TN
35 *ns, *ms, *rs > d'	-	-	-	*	-	TE
50 *a > ɔ	-	-	*	*	-	FE+TE
51 *å > a	*	*	*	*	-	FN+TN+FE+TE
52 *å > o in the 1st syllable	-	-	-	-	*	Ng
53 *å > u in non-1st syllables	-	-	-	-	*	Ng
54 *g > i after non-labial consonants	-	-	*	*	-	FE+TE
55 *j > i after non-labial consonants	-	-	-	-	*	Ng
56 *g, j > u after labial consonants	-	-	*	*	-	FE+TE

Unexpected from the genealogical tree, but expected from the areal contacts: 8

Archaisms? 4

FN+TN+FE+TE	8
FN+TN	5
FE+TE	10
FE+TE+Ng	5
TN+FE+TE	3
FN+TN+Ng	2
FN+TN+FE+Ng	1
FN+FE+TE	1
all	7
Ng	10
TN	8
FN	2
FE	1
TE	1



3. CHANGES IN PHONOLOGY: DATA

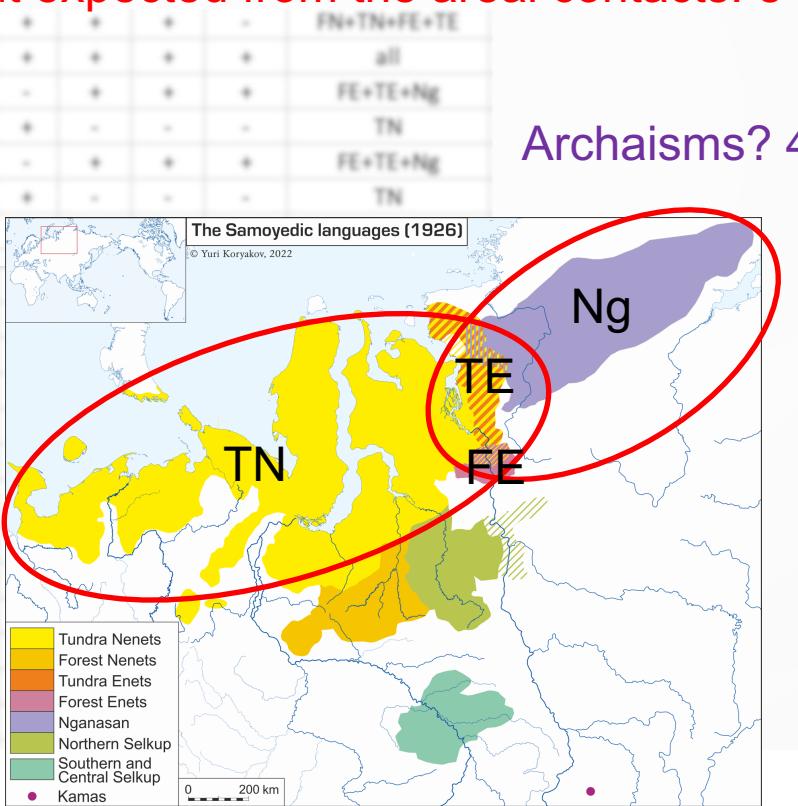
Setting up the genealogical tree: 23

Setting up the genealogical tree: 23

Some snapshots from the ~~Uhexp~~ but sync

- 11 *k > x before back vowels intervocally
 12 *k > i before front vowels
 13 *w > b word-initially
 14 *w > j before front vowels word-initially
 15 *w > 0 intervocally
 16 *w > b' before front vowels intervocally
 17 *j > 0 in some contexts/words intervocally

30 *NT > ND
 31 *NT > D
 32 *nt > n
 33 *ns, *ms, *rs > s, z (both variants)
 34 *rs > rc
 35 *ns, *ms, *rs > d'
 50 *a > ɔ
 51 *å > a
 52 *å > o in the 1st syllable
 53 *å > u in non-1st syllables
 54 *ę > i after non-labial consonants
 55 *j > i after non-labial consonants
 56 *ę, j > u after labial consonants



Archaisms? 4

Unexpected from the genealogical tree, but expected from the areal contacts: 8

FN+TN+FE+TE	8
FN+TN	5
FE+TE	10
FE+TE+Ng	5
TN+FE+TE	3
FN+TN+Ng	2
FN+TN+FE+Ng	1
FN+FE+TE	1
all	7
Ng	10
TN	8
FN	2
FE	1
TE	1



3. CHANGES IN PHONOLOGY: UNEXPECTED SPREAD

FE+TE+Ng

1. *p > f:
 - En: only as a variation, later the more archaic variant was restored.
 - Ng: later f > h (intervocally h in strong grade, b in weak grade)
2. *s > s, θ (free variation)
3. *w > zero intervocally
4. *j > zero intervocally (in some contexts/words only)
5. *w > b word-initially



3. CHANGES IN PHONOLOGY: UNEXPECTED SPREAD

FE+TE+Ng

1. *p > f:
 - En: only as a variation, later the more archaic variant was restored.
 - Ng: later f > h (intervocally h in strong grade, b in weak grade)
2. *s > s, θ (free variation)
3. *w > zero intervocally
4. *j > zero intervocally (in some contexts/words only)
5. *w > b word-initially

lenition,
very common
cross-linguistically
(Blevins 2004: 144-
147, Bybee &
Shelece 2019), i.e.
**easy to copy in a
contact situation**



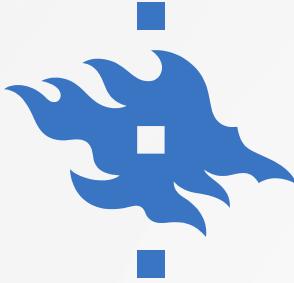
3. CHANGES IN PHONOLOGY: UNEXPECTED SPREAD

FE+TE+Ng

1. $*p > f:$
 - En: only a λ variation, later the more archaic variant was restored.
 - Ng: later $f > h$ (intervocally h in strong grade, b in weak grade)
2. $*s > s, \theta$ (free variation)
3. $*w >$ zero intervocally
4. $*j >$ zero intervocally (in some contexts/words only)
5. $*w > b$ word-initially

the change is phonetic,
not phonological!

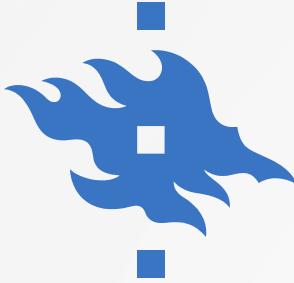
lenition,
very common
cross-linguistically
(Blevins 2004: 144-
147, Bybee &
Shelece 2019), i.e.
**easy to copy in a
contact situation**



3. CHANGES IN PHONOLOGY: UNEXPECTED SPREAD

TN+FE+TE

1. *p > b intervocally
2. *t, *č > δ intervocally,
3. n, η > ? word-finally

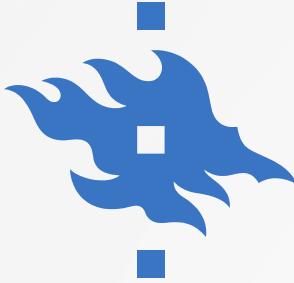


3. CHANGES IN PHONOLOGY: UNEXPECTED SPREAD

TN+FE+TE

1. *p > b intervocally
2. *t, *č > δ intervocally,
3. n, η > ? word-finally

lenition,
very common
cross-linguistically
(Blevins 2004: 144-
147a, Bybee & Shelece
2019), i.e. **easy to copy**
in a contact situation

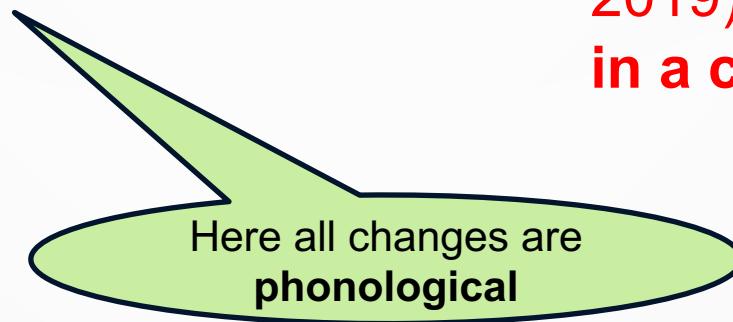


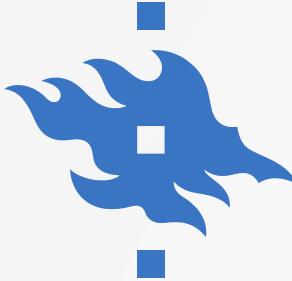
3. CHANGES IN PHONOLOGY: UNEXPECTED SPREAD

TN+FE+TE

1. *p > b intervocally
2. *t, *č > δ intervocally,
3. n, η > ? word-finally

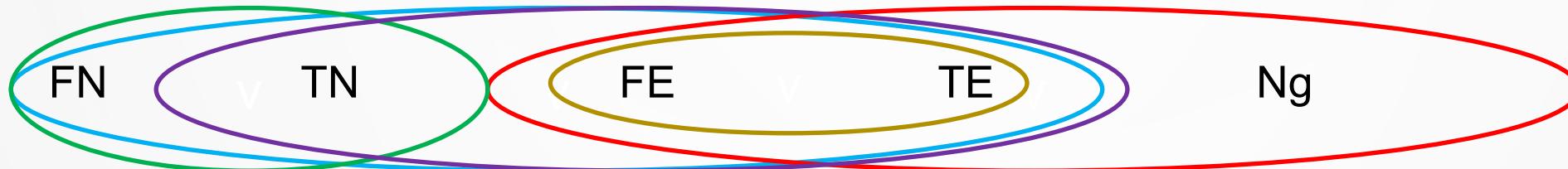
lenition,
very common
cross-linguistically
(Blevins 2004: 144-
147a, Bybee & Shelece
2019), i.e. **easy to copy**
in a contact situation

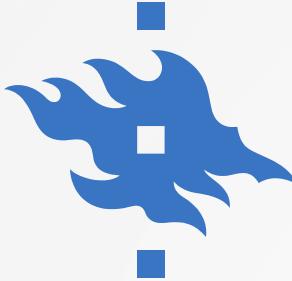




3. CHANGES IN PHONOLOGY: SUMMARY

1. Some sound changes happened in Proto-Enets-Nenets (8), in Proto-Nenets (5), in Proto-Enets (10): branch-defining.
2. However, other, less numerous, sound changes spread **horizontally, not vertically**: within (Proto-) Enets and (Proto-)Nganasan (5), or Tundra Nenets and (Proto-)Enets (3).





3. CHANGES IN PHONOLOGY: TIMING

3. The horizontal spread of sound changes happened **before the Enets split**, though (also continued) **after the Nenets split**
 - TN has shared some sound changes with Enets, but FN did not
 - neither FE, nor TE show any sound change shared with a NS language, but not with the other Enets;
 - the earliest documentation of Enets as distinct belongs to the 18th cent. by G.F. Müller:
 - by then, their individual changes, not replicated in any other NS, already happened

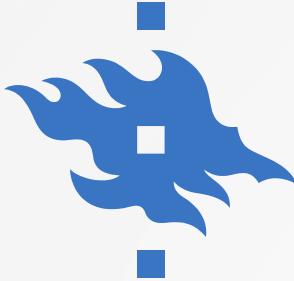
separate sound changes in FN, TN
(~ after 1000 AD)

separate sound changes in FE, TE
(~ after 1200 AD)

Time line →

areal spread of sound changes

4. MORPHOLOGY



4. NS MORPHOLOGICAL ISOGLOSSES

- We know much less about Proto-Uralic and Proto-Samoyedic morphological systems!
 - difficult to compile a list of changes ‘PS suffixes > modern NS suffixes’;
 - list of isoglosses in inflectional morphology attested in 2+ NS languages (77 items);
 - primarily isoglosses based on innovations; clear archaisms not included;
 - *not done at this stage*: Samoyedic, but not NS > PS morphology
 - isoglosses include both related and unrelated morphemes:
 1. related morphemes with identical/similar functions,
 2. absence (=loss) of a cognate morpheme or of one of its functions;
 3. unrelated morphemes with equivalent functions: cross-linguistically or areally non-trivial,
 4. formal parallelism: reanalysis, co-occurrence, use in a construction, restrictions, etc.,
 5. co-occurrence of functions regardless the cognacy.

⇒ as in phonology, the contrasts in numbers are only *rough estimates* of what is going on

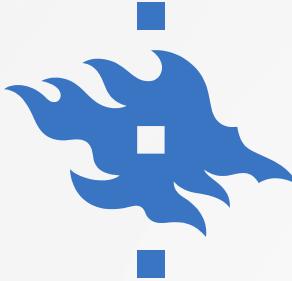


4. NS MORPHOLOGICAL ISOGLOSSES

- We know much less about Proto-Uralic and Proto-Samoyedic morphological systems!
 - difficult to compile a list of changes 'PS suffixes > modern NS suffixes';
 - list of isoglosses in inflectional morphology attested in 2+ NS languages (77 items);
 - primarily isoglosses based on innovations; clear archæological traces not included;
 - *not done at this stage*: Samoyedic, but not NS > PS morphemes
 - isoglosses include both related and unrelated morphemes:
 1. related morphemes with identical/similar functions,
 2. absence (=loss) of a cognate morpheme or of one of its variants;
 3. unrelated morphemes with equivalent functions: cross-linguistical analogies;
 4. formal parallelism: reanalysis, co-occurrence, use in a construction;
 5. co-occurrence of functions regardless the cognacy.

The list is mainly based on our own research

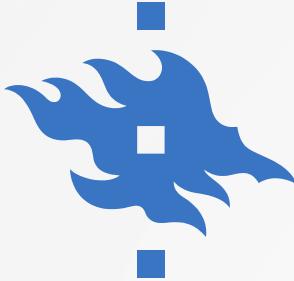
⇒ as in phonology, the contrasts in numbers are only rough estimates of what is going on



4. NS MORPHOLOGICAL ISOGLOSSES: DATA

1		FN	TN	FE	TE	Ng	
59	Posterior participle *-mäntå	+	+	+	-	-	FN+TN+FE
60	Posterior participle *-CsVNtå > TE -tSedo	-	-	-	+	-	TE
61	Anterior participle *-mäjå	+	+	+	+	+	all
62	Anterior participle *-mäjå passive only	-	-	-	(+)	+	TE+Ng
63	Anterior participle based on past tense affix *-så-			-	+	+	TE+Ng
64	Negative anterior participle ("not yet") *-mätämaj	+	+	+	+	(+)	FN+TN+FE+TE
65	Passive anterior participle -du?uj			+	-	-	FE
66	Loss of the productive simultaneous nominalization *-må	-	-	-	-	+	Ng
67	Simultaneous nominalization *-mon	-	(+)	-	-	+	TN+Ng
68	Anterior nominalization *-kma-	+	+	+	+	+	all
69	Decline of the anterior nominalization *-kma-	-	-	+	+	-	FE+TE

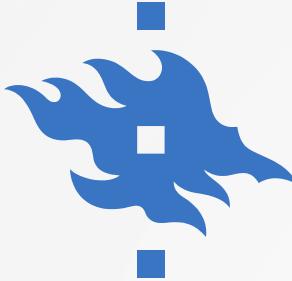
1		FN	TN	FE	TE	Ng	
3	Pl: Plural marker -nV in the possessive forms for kinship terms and words like 'friend'	+	-	+	+	-	FN+FE+TE
4	Core cases (non-possessed): Non-phonetic loss of the GEN.SG and ACC.SG markers	-	-	+	+	-	FE+TE
5	Core cases (non-possessed) have merged in the Dual (Nom vs. Obl in TE and Ng)	+	+	+	-	-	FN+TN+FE
6	Core cases: GenPl and AccPl have been ousted by NomPl forms	-	-	+	+	-	FE+TE
7	LAT.SG has the same co-affix for poss and non-poss forms	-	-	-	-	+	Ng
8	LOC Sg and Pl: -ntå instead of -kä in	-	-	-	-	+	Ng
9	ABL.SG Reflex of the single *t (vs. or of a cluster)	+	+	+	+	-	FN+TN+FE+TE



4. NS MORPHOLOGICAL ISOGL

	0%	1%	2%
1. Plural marker - <i>ta</i> in the possessive forms for singular terms and words like 'friend'	0	0	0
2. Some cases (non possessive). Non-phonetic loss of the GEN.SG and ACC.SG markers.	0	0	0
3. Some cases (non possessive) have merged in the Dual (from vs. GEN in TS and RG).	0	0	0
4. Some cases. GenSF and AccSF have been ousted by NeutSF forms.	0	0	0
5. Latv.SG has the same suffix for gen and non-gen forms.	0	0	0
6. Latv.SG and PL. - <i>ta</i> instead of - <i>ta</i> - <i>ta</i>	0	0	0
7. Latv.SG suffix of the single 't' (i.e. on off a cluster)	0	0	0

FN+TN+FE+TE	17
FN+TN	11
FE+TE	14
FE+TE+Ng	8
TE+Ng	9
FN+TN+FE	10
TN+FE+TE	3
TN+FE	1
TN+Ng	1
TN+FE+Ng	1
FE+Ng	2
FN	2
TN	4
TNwest	1
TNwest+TNcentre	1
TNcentre	1
FE	6
TE	2
Ng	6



Expected from the genealogical tree: 42

4. NS MORPHOLOGICAL ISOGLYPHS

Unexpected from the genealogical tree,
but expected from the areal contacts: 31

Archaisms? 4

FN+TN+FE+TE	17
FN+TN	11
FE+TE	14
FE+TE+Ng	8
TE+Ng	9
FN+TN+FE	10
TN+FE+TE	3
TN+FE	1
TN+Ng	1
TN+FE+Ng	1
FE+Ng	2
FN	2
TN	4
TNwest	1
TNwest+TNcentre	1
TNcentre	1
FE	6
TE	2
Ng	6



4. UNEXPECTED MORPHOLOGICAL ISOGLOSSES

FE+TE+Ng (verb only)

1. Probabilitive *-btV (FE, TE -ta, Ng -bta,)
2. Habitual *-mumpV (FE, TE -obi/-mobi, Ng -muŋha)
3. Simultaneous converb *pu[?]θ- (FE, TE -bu[?]o/-pu[?]o, Ng -hü[?]θ)
4. Generic state FE, TE -be, Ng -min/mün
5. Anterior converb *kaj-Inf (FE, TE -xaj-Inf/-kaj-INF, Ng -kaj-INF)
6. Soft/Future imperative *-kua (FE -kuri, TE -kua, Ng -ku(ə))
7. 3Sg.R *-tən (FE, TE -zo[?], Ng -δə)
8. Auditive is formally identical to plural possessive forms of noun (FE, TE -munu-, Ng -munu-jt'ü)



4. UNEXPECTED MORPHOLOGICAL ISOGLOSSES

TE+Ng (7 verb, 1 pronoun, 1 transcategorial)

1. Interrogative *-på (TE -ba/-pa, Ng -hu)
2. Soft/future imperative marker *-kua (TE -kua, Ng -kuə)
3. Anterior participle based on past tense affix *-så- (TE -si(j), Ng -suəd'əə)
4. Futuritive *-?sutə (TE -t'uzo as analytical debititive, Ng -?sutə as plain future)
5. Adjectivizer *-je used with simultaneous participle (TE -de, optional; Ng -ntuə obligatory)
6. Anterior participle *-məjə is restricted to passive (Castrén's TE -j, not in modern TE; Ng -məə)
7. (+ loss of Subjunctive *-ní (very rare in TE, absent from Ng))
8. Reflexive pronoun (=body): *p?(tə)tə (TE pu(zo)zo, Ng hütəδə)
9. Exclamatives inside other word-final suffixes before their last consonant (TE -ou, -ej, Ng -əu, -əi)



4. UNEXPECTED MORPHOLOGICAL ISOGLOSSES

FN+TN+FE (4 verb, 2 noun, 4 pronoun)

1. Interrogative *så (FN, TN -sa, FE -sa)
2. Posterior participle *-məntå (FN -manta, TN -mənta, FE -muda/-uda)
3. Suppositional prospective PTCP.PROSP-COMP *məntå-rəkå (FN -manta-łaxa, TN -mənta-rəxa, FE -uda-raxa)
4. Loss of the Dubitative *-li
5. Core cases (non-possessed) have merged in the Dual (Nom vs. Obl in TE and Ng)
6. Loss of the number distinction in Essive/Translative
7. Innovative 2 person pronouns
8. Genitive and Accusative pronouns use the same suppletive stem
9. Demonstrative/interrogative pronouns with *-rsä (FN tal-s'a, TN tər-ća, FE el-se, tor-se, kur-se)
10. Reflexive pronoun (=body): *p?(kə)tə (TN pi(x)əd°, FE pu(xu)zu)



4. UNEXPECTED MORPHOLOGICAL ISOGLOSSES

TN+FE+TE (2 verb, 1 pronoun)

1. Supine *-mənsi (TN -məncyo^o, FE, TE -mod'i/-od'i)
2. *-təm Indicative for 1SG.S (TN -təm, FE, TE -zo?)
3. Agglutinative Du and Pl personal pronouns

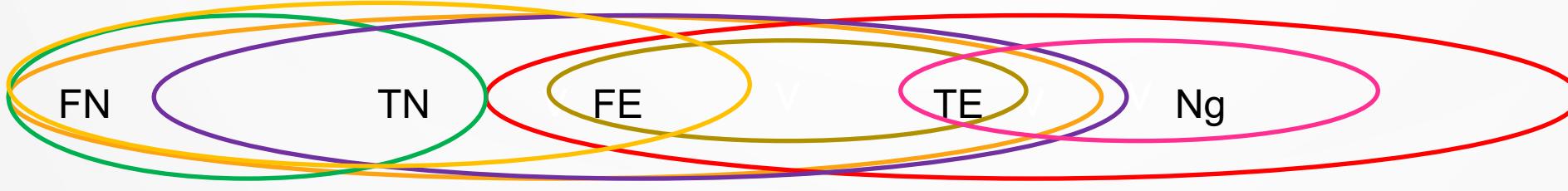
TN+FE (pronoun)

1. Full stems of personal pronouns are used with emphatic suffixes (FE mod'i-ruu-n 'only I')



MORPHOLOGICAL ISOGLOSSES: SUMMARY

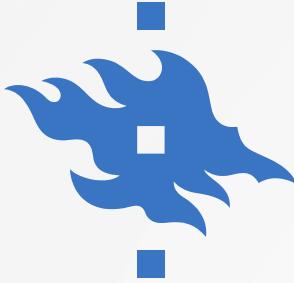
1. In morphology we see more unexpected shared features (31 to 42) than in phonology (8 to 23).
2. **Number of morphological isoglosses unexpected from the traditional genealogy (based on shared phonetic innovations) is very high.**
3. It is possible to conceptualise (most) morphological isoglosses as **pattern borrowing** (not matter borrowing):
 - an increase in frequency of an already existing, but marginal in its uses so far, Proto-Samoyedic pattern (which came out of use in the other languages),
 - morphosyntactic calquing: parallel constructions made up of own elements





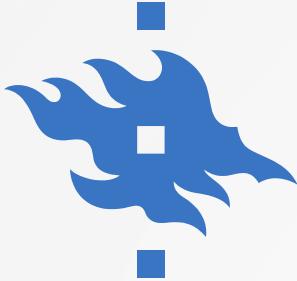
4. SUMMARY: WHAT WE NOW KNOW

- (not discussed here!) No phonetic innovation is unique to NS, but there are numerous lexical and morphological items unattested in the other Samoyedic.
- NS display ca. 300 **unique lexemes**, unattested in the other Sam/Uralic, or in the known neighbouring languages: half of them display **irregular** phonetic correspondences.
- Nganasan has further ca. 600-700 **unique lexemes** (unattested in the other NS).
- Neither NS, nor Ng lexemes form a clear strata, on phonetic or semantic grounds.
- Most **phonetic innovations** in NS languages spread **within the generally accepted subgroups** (Nen-En, or Nen, or En); some phonetic innovations spread within neighbouring languages from **distinct subgroups**: En & Ng or Tundra Nen & En.
- Some supposedly innovative **morphological isoglosses** between any 2+ NS languages are shared within the accepted subgroups; but many are shared within neighbouring languages from **distinct subgroups**: En & Ng, TEn & Ng, Nen & FEn, TNen & En.



4. SUMMARY: WHAT WE DO NOT KNOW

- The origin of the lexemes unique to NS or to Nganasan:
 - PS retentions? If so, were they inherited from PU or borrowed into PS?
 - Separate borrowings into individual NS languages from an unknown source?
 - One or several unknown languages? Were the unknown languages related to each other?
 - Were the unknown languages very similar to NS in their phonologies or did NS speakers were excellent in phonetic adaptation of the borrowings?
- The origin of the NS morphological isoglosses:
 - Was there any matter borrowing behind them, or were they all cases of pattern borrowing?
 - Did any of them form under the influence of an unknown language? Cf. (Gusev 2021)
 - Are they attested in the other Samoyedic?
- Were there more NS languages, intermediate between the ancestors of the modern NS?



4. SUMMARY: WHAT WE DO NOT KNOW

- The origin of the lexemes unique to NS or to Nganasan:
 - PS retentions? If so, were they inherited from PU or borrowed into PS?
 - Separate borrowings into individual NS languages from an unknown source?
 - One or several unknown languages? Were the unknown languages related to each other?
 - Were the unknown languages very similar to NS in their phonologies or did NS speakers were excellent in phonetic adaptation of the borrowings?
- The origin of the NS morphological isoglosses:
 - Was there any matter borrowing behind them, or were they all cases of pattern borrowing?
 - Did any of them form under the influence of an unknown language?
 - Are they attested in the other Samoyedic?
- Were there more NS languages, intermediate between the ancestors of the modern NS?

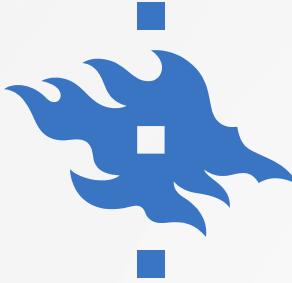
Possible to find out



5. PROPOSED SCENARIO

Or what could be hypothesized from what we know

- 1. Contacts with non-Uralic.**
 - 2. Prolonged interactions between speakers of NS at various times: early post-PS (1st millennium AD), before and after the Nenets split (ca. 1000 AD), before and after the Enets split (between 1200-1600 AD).**
- (1)-(2) could explain the NS puzzle: no phonetic innovations, but multiple morphological & lexical ones.
 - No Proto-Northern-Samoyedic, but Proto-Nenets, Proto-Enets, and Proto-Nganasan in intensive contacts with their immediate NS neighbours (more intermediate languages?).
 - Sociolinguistic reconstruction for the last 200 years provides a probable social-cultural-historic grounding.
 - (Khanina 2021, 2022): local migrations, multilingualism, language shifts, ideologies 'neighbourhood over ancestry'.



5. PROPOSED SCENARIO

3. Possible grounds for structural divergence between PNen, PEn, PNg:

- loss of intermediate NS languages,
- or/and a temporary geographic isolation, e.g. separate northward migrations from Proto-Samoyedic homeland;
- or/and an emergence of new social identities, e.g. speakers of non-Uralic languages who shifted to a NS language (substrate features, conscious differentiation efforts)

Archaeology: the modern NS area was inhabited long before the break-up of PS ca. 2000 years ago.

Genetics:

- (Karafet et al. 2018): some Y-chromosome haplogroups can be detected in modern speakers of Samoyedic languages, but not in their neighbours speaking other languages (Khanty, Ket, Evenkis, Yukaghirs, Koryaks),
- (Karafet et al. 2018): there is more similarity in the autosomal DNA (=full genome) between NS speakers and their eastern neighbours, than between NS speakers and the Selkups.
- (Kharkov 2021): modern Tundra Nenets speakers carry Y-chromosomes which represent distinct temporal layers of related haplogroups, implying that TN male ancestors could arrive from the south in several waves.



REFERENCES

- Blevins, Juliette. Evolutionary phonology. Cambridge: CUP.
- Bybee, Joan and Easterday, Shelece. "Consonant strengthening: A crosslinguistic survey and articulatory proposal" *Linguistic Typology*, vol. 23, no. 2, 2019, pp. 263-302.
- Gusev, Valentin. 2021. К типологической характеристике северосибирского субстрата, Вопросы языкоznания 5, 26-58.
- Gusev, Valentin. 2022. Рефлексы прасамодийских *а и *ä и нганасанский сингармонизм. In: Valentin Gusev, Anna Urmanchieva, Aleksandr Anikin (eds.) *Siberica et Uralica: In memoriam Eugen Helimski*. Szeged, 2022. P. 39–62.
- Helimski, Eugen 2005: The 13th Proto-Samoyedic vowel. – B. Wagner-Nagy (ed.): *Mikola-konferencia 2004*. Szeged: SzTE Department of finnougristics, 15–26.
- Helimski, Eugen. 1997. Die matorische Sprache. Wörterverzeichnis. Grundzüge der Grammatik. Sprachgeschichte. Unter Mitarbeit von Beáta Nagy. Szeged.
- Helimski, Eugen 1993: Прасамодийские *ê и *ë: прауральские источники и нганасанские рефлексы. – Hajdú Péter 70 éves. Budapest, 1993, 125–133).
- Janhunen, Juha 1975–1976: Adalékok az északi-szamojéd hangtörténethez: vokalizmus. Az első szótági magánhangzók. – Néprajz és nyelvtudomány, 19–20, 165–188.
- Janhunen, Juha 1977: Samojedischer Wortschatz: Gemeinsamojedische Etymologien. Helsinki.
- Mikola, Tibor. Studien zur Geschichte der samojedischen Sprachen. Aus dem Nachlass herausgegeben von Beáta Wagner-Nagy. (Studia Uralo-altaica 45.) Szeged, 2004.
- Salminen, Tapani 2012: Traces of Proto-Samoyed vowel contrasts in Nenets. – Per Urales ad Orientem: Iter polyphonicum multilingue. (SLIST 264.) Helsinki, 339–358.