

# Loanwords and Narrative Historical Linguistics

Rasmus G. Bjørn

*Archaeo-Linguistic Research Group (DA)*

MPI-SHH  
SUMMER SCHOOL  
2021

Doorway  
to Human History

You're in for a ride!

We'll travel the globe, (mainly Eurasia)

trade, a lot (tea, ivory, silk)

count, a bit, (just 7)

and try to *make a name* for ourselves.



# Structure of this lecture

## Loanwords

Relevance to the study of human history

What is a loanword?

How do we detect them?

Other linguistic contact phenomena

## Narrative Historical Linguistics

Connecting the dots

Posing hypotheses

Putting loanwords to use

## Perspectives

Testing hypotheses



# Quick background

- Rasmus G. Bjørn
- Archaeolinguistic Research Group (Dept. of Arch.)
- Indo-European studies, MA from Copenhagen, Roots of Europe
- „Old School“ qualitative linguist
  - So no numbers (just one numeral)



# Why loanwords?



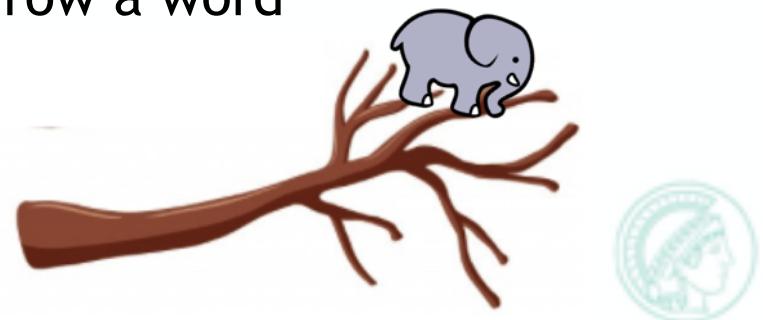
Language families usually reconstruct into prehistory

Loanwords provide evidence of contact

- that in turn helps us infer dating from one family to another

Words attest to phenomena both visible (e.g. livestock, agriculture, implements) and hidden to archaeology and genes, (e.g. numerals, social structure)

You don't need to make a child to borrow a word



# What is a loanword?

Loanword = borrowing = copying

Donor/source language provides a (native) word to a recipient language

Loanwords carry a *meaning* that tells something about the exchange



# All words have a history



Eng. *dandelion*,

- Influence from French
- French perceived ‘lion tooth’



# Terminology

Comparandum (pl. comparanda)

= word(s) being compared

item = technology, concept, crop, animal, or implement of interest

- ~ “can be compared to”
- > “regular” descended to”
- < “developed from”
- “was borrowed into”



# They're *very* common

<b>They</b>	<b>'re</b>	<b>very</b>	<b>common</b>
Old Norse <i>beir</i>		Old French <i>verai</i>	Old French <i>comun</i>
Norse influence		Norman rule	Norman rule
800 -1066 AD		1066-1200	1066-1200

especially in English

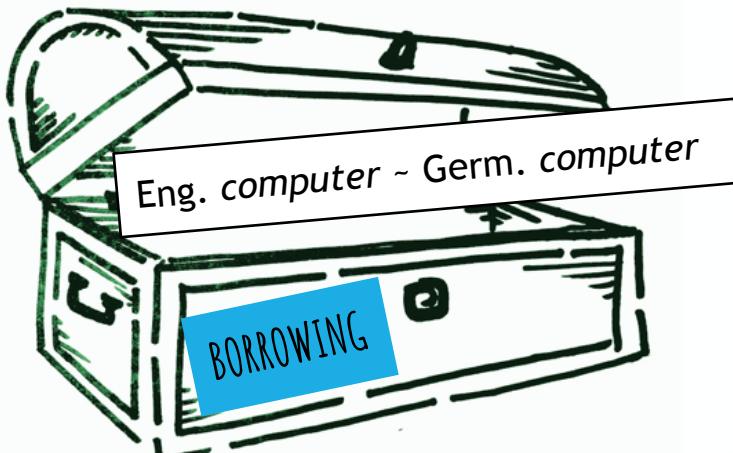
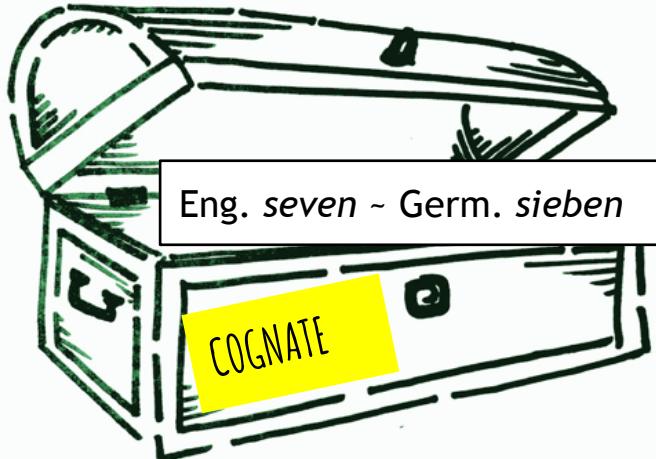
And they tell of different layers of history



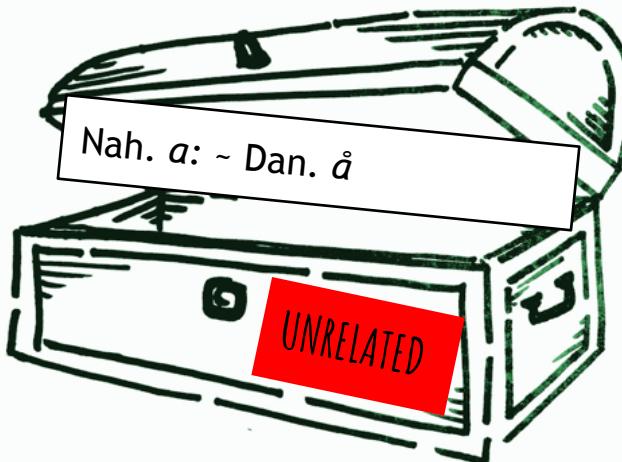
# Sorting words

Known divergent histories?

Regular sound changes?



Contact and similarity?



# Receiving loanwords

Phenomena important for deciding direction of borrowing

Subject to the rules of the receiving language

Internal chronologies (all of a sudden appear)

Internal etymologies (but note folk etymologies)

Circumstantial evidence



# Receiving loanwords

Subject to the rules of the receiving language



## Finnish

- no voice distinction (p vs. b)
- No initial consonant clusters (“street”):

Swedish *franska* ‘French’ → *ranska*

Eng. *humbug* → *humpuuki*



## Sayula Populuca (S. Mexico)

- No liquids (r/l)
- No consonant clusters

Spanish *plato* ‘plate, dish’ → *puná:tu*



# The plight of the linguist

Detecting a borrowing is an exercise in diachrony (Haugen 1950)

But we are not always afforded diachrony

- language isolates
- proto-languages

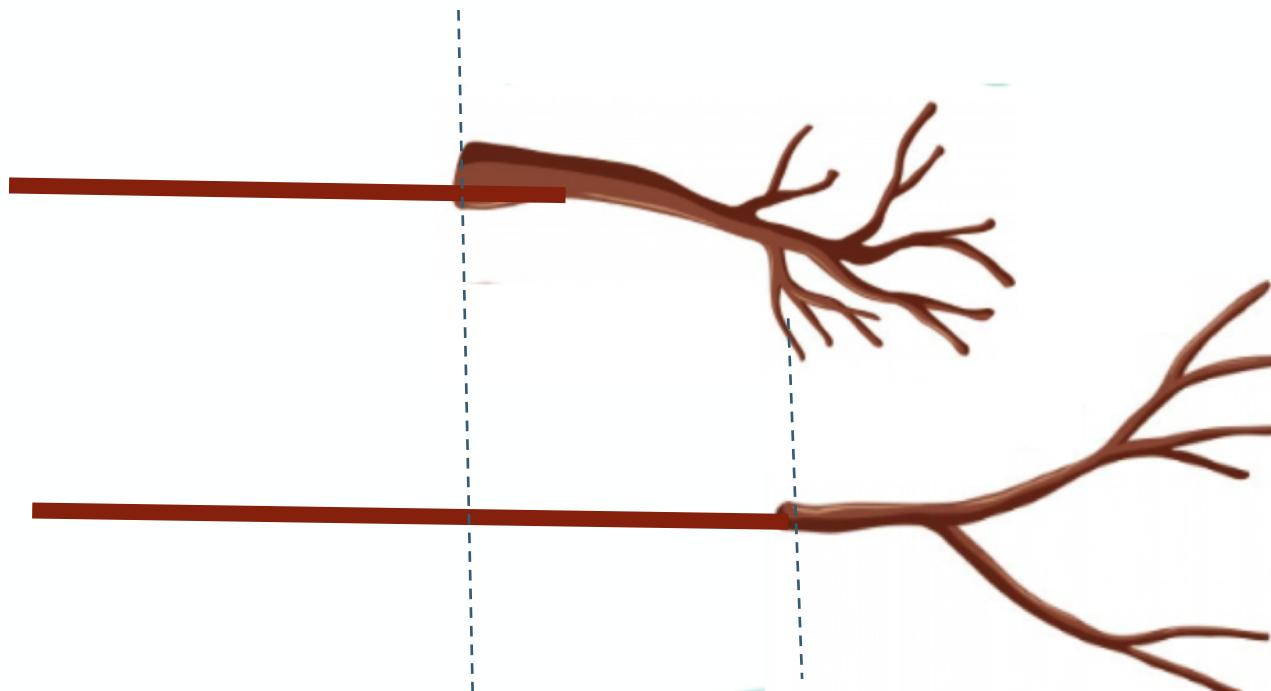
To this day, despite instant chat and recording, most speech is fleeting

There is A LOT that we don't know



# Cease and desist?

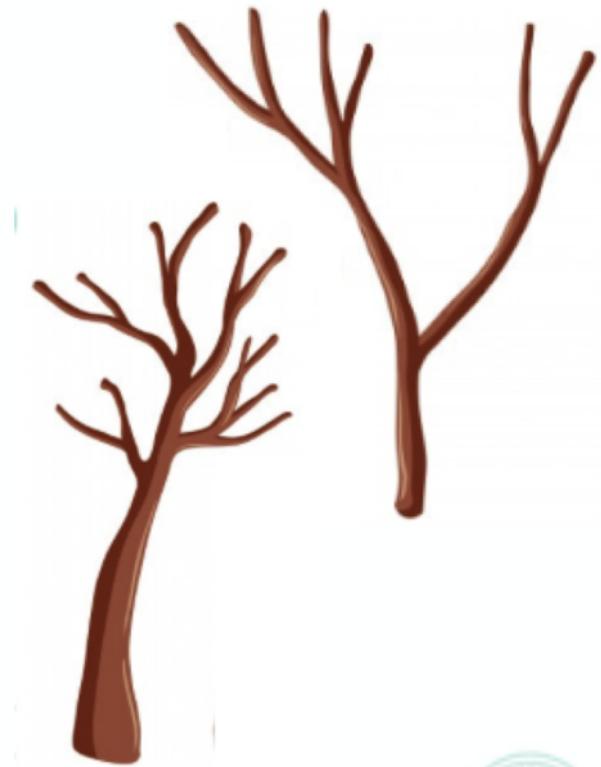
Cannot test with the usual rigour beyond the proto-languages



# Loanwords and language contact

In principle detached

Contacts calibrate neighboring families



# Other signals of language contact

Adstrate phenomena (substrate or superstrate)

Very important in understanding prehistoric contact

Difficult to tie to archaeology

Carries evidence for social dynamics

This pattern is not found in conservative Oceanic languages. Some examples of calqued formulae are listed in (9):

(9)	'literal' meaning	Takia	Waskia	
'the palm of my hand'	'my hand's liver'	<i>bani-g</i> hand-1SG	<i>ate-n</i> liver-3SG	<i>a-gitig</i> 1SG-hand
'I am dizzy'	'my eye goes round'	<i>mala-g</i> eye-1SG	<i>i-kilani</i> 3SG-go.round	<i>motam</i> eye
'I disobey him'	'I cut his mouth'	<i>awa-n</i> mouth-3SG	<i>gu-tale</i> 1SG-cut	<i>kurig</i> 3SG.mouth
'I am angry'	'my guts are bad'	<i>ilo-g</i> inside-1SG	<i>saen</i> bad	<i>a-gemay</i> 1SG-liver
'I am waiting'	'I am putting my eye'	<i>mala-g</i> eye-1SG	<i>gi-ga</i> 1SG-put	<i>motam</i> eye

## Calques

- Ross 2003: 187

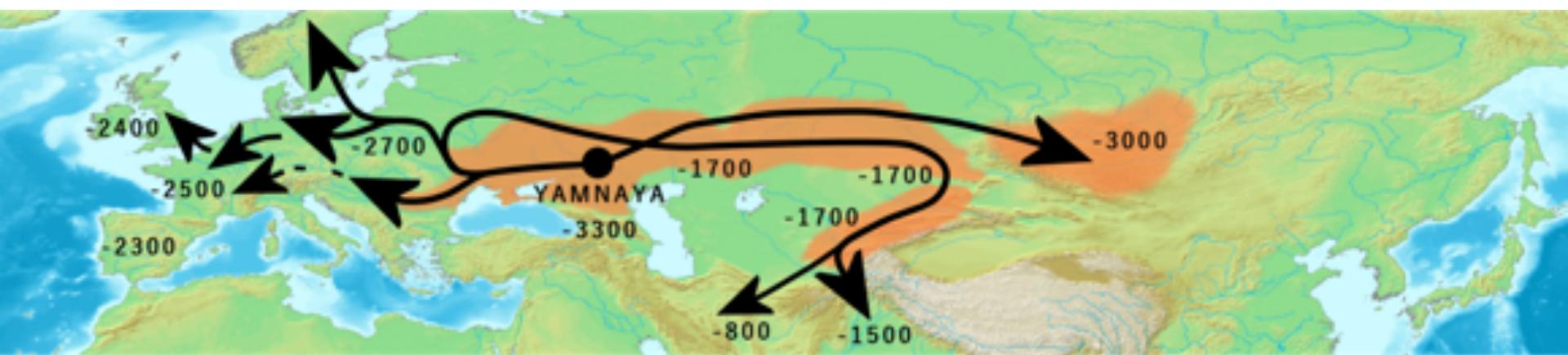


# Case: Typology in historical linguistics

The “deviant” typology of Tocharian

Peyrot (2019) and Warries (2019)

Proved sustained contacts with a “Samoyedic-like” language (Uralic) - without loanwords



# Deviant typology of Tocharian, stops

TABLE 2 Typological comparison of PIE, PToch., PU and PSam. obstruent systems

Proto-Indo-European			Proto-Tocharian	Proto-Uralic	Proto-Samoyedic
*k	*g	*g <sup>h</sup>	*k	*k	*k
*kʷ	*gʷ	*gʷʰ	*kʷ		
*ǵ	*ǵ	*ǵʰ	*ts	*c	*c
*t	*d	*dʰ	*t	*t	*t
*p		*bʰ	*p	*p *δ, *δ'	*p
*s			*s	*s *ś	(*ś) *s



# Deviant typology of Tocharian, vowels

TABLE 9      Typological comparison of Pre-Proto-Samoyedic and  
Pre-Proto-Tocharian vowel systems

Pre-Proto-Samoyedic			Pre-Proto-Tocharian			
<i>i</i>	<i>i̯ (= i̮)</i>	<i>u</i>	<i>i</i>	<i>ə</i>	<i>u</i>	
<i>e</i>	<i>ɛ (= ē)</i>	<i>o</i>	<i>(*ē &gt;) e</i>	<i>ē (&lt; *o)</i>	<i>o (&lt; *ā)</i>	
	<i>a</i>			<i>a</i>		



# Deviant typology of Tocharian, cases

Table 11. The Proto-Samoyedic and Proto-Tocharian case systems compared; the first function listed is considered to be the more original (local) function(s), with other (grammatical) functions developing later listed second.

Case / functions in:	Proto-Tocharian	Proto-Samoyedic
(al)lative	direction of movement	movement into; direction of movement; dative?
locative	internal location; movement into	(internal) location; instrument?
ablative	source; starting point; standard of comparison?	source; starting point; standard of comparison
periative	path or direction of movement; external location; cause; instrument? standard of comparison?	path of movement
comitative	accompaniment	—
predestinative	—	X{PRED} for ... [Y{GEN/POSS}]; future reference? (translative?)

# Obfuscating factors

Lack of data

Folk etymologies

Dialectal borrowings

Reasons to keep an open mind



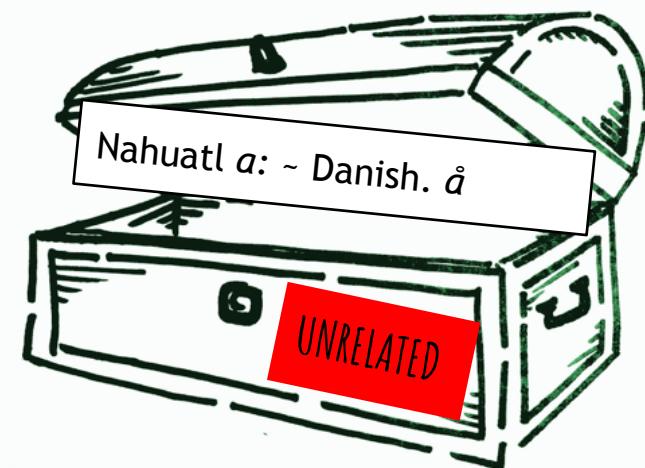
# Testing a (wrong) hypothesis

Nahuatl *a*: ‘water’ looks like Danish *å* ‘river, stream’

Are they related?

Perhaps as loanwords? Vikings

Uto-Aztecán \**pa:-* =/= Proto-Germanic \**ahw*



# Data impasses

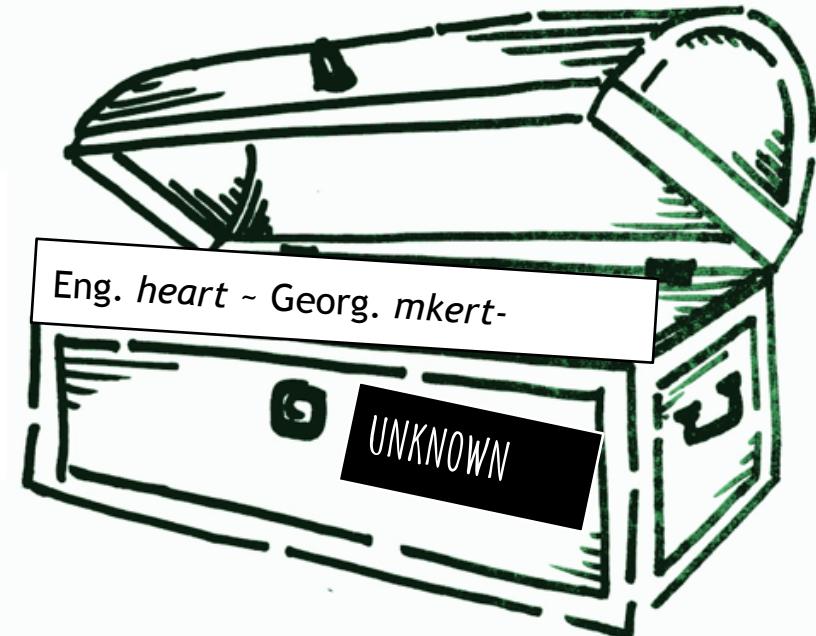
Several comparanda without a convincing explanation, e.g.

PIE \*k'erd- 'heart' ~ Proto-Kartvelian \*m-kerd 'chest'

We still don't have the evidence to place it in one of three chests, so we'll opt for the fourth

Doesn't mean it can't be done!  
(cf. Róna-Tas 1983)

under favourable circumstances. Among the undoubtedly Turkic loanwords in Hungarian we find *gyomor* 'stomach', *kar* 'arm', *koldök* 'navel', *térd* 'knee' to mention only those which denote parts of the body. They were borrowed not because the ancient Hungarians did not have stomachs, arms etc. but as denominations of the parts of body of the animals while taking over a new type of animal husbandry. Later on they were generalized (see Róna-Tas 1981). The two exam-



# Folk etymologies



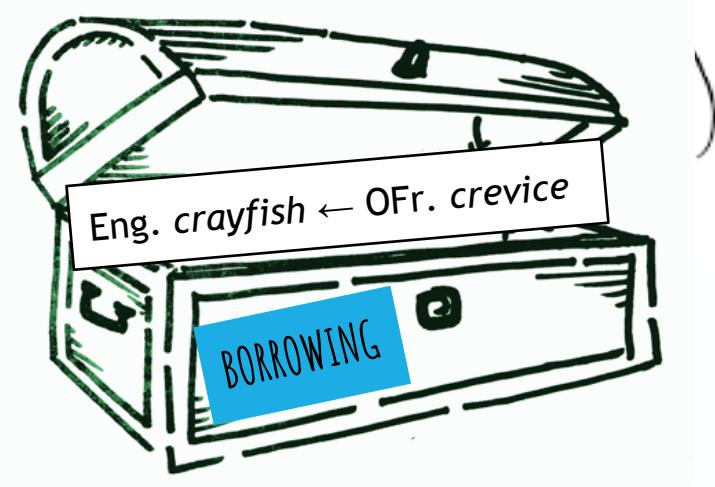
We like to make sense of the wor(l)d

English *crayfish* or *crawfish*



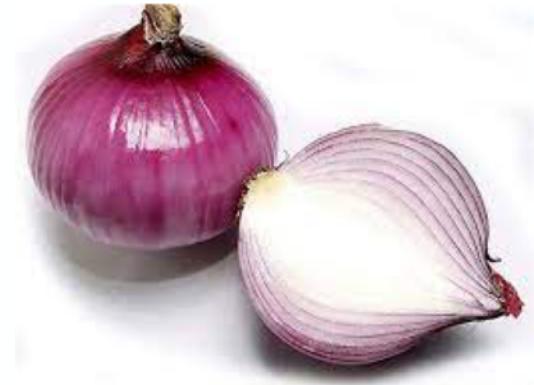
but is it a fish?

No, a folk etymology of Old French *crevic*



# Another folk etymology

German *Zwiebel* ‘onion’ = ‘two ball?’  
no ← Lat. *caepolla*



Or forced academic etymologies?  
Wool, from ‘to pluck’ ... or a borrowing

or both



# When all else fails -- loanword?

No, chance, or rather, undecided.

I perceive a tendency to choose chance over unknown.

But evidence is needed to prove chance



# Chance galore

Linguistic coincidences and surprises

🔒 Privat gruppe · 17,5 tusind medlemmer



 Yongky Utama  
14. august kl. 03.38 · ⓘ ...

- Finnish: pohjoinen "north"
- Shanghainese: 北 poq /poʔ/ "north"

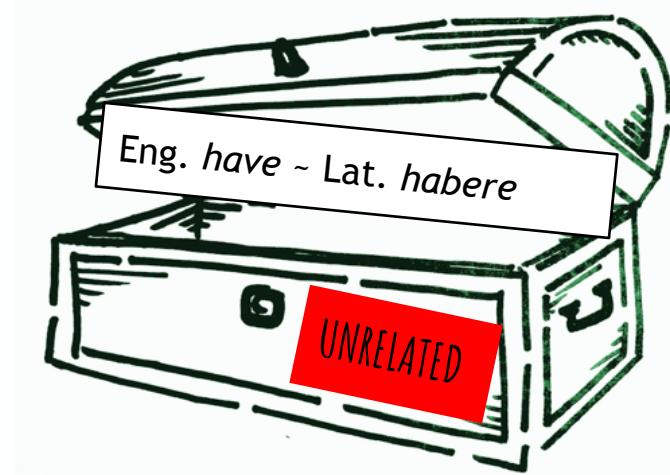
pohjoinen: derived from pohja "bottom"  
北 poq: derived from Old Chinese \*pw:gs "back (of body)", later written as 背

 10      1 kommentar



# Fooled by similarity

Engl. <i>to have</i>	=/=	Lat. <i>habere</i> 'to have'
PIE * <i>kap-</i>		PIE * <i>gʰabʰ-</i>
cf. Lat. <i>capere</i>		cf. Engl. <i>to give</i>



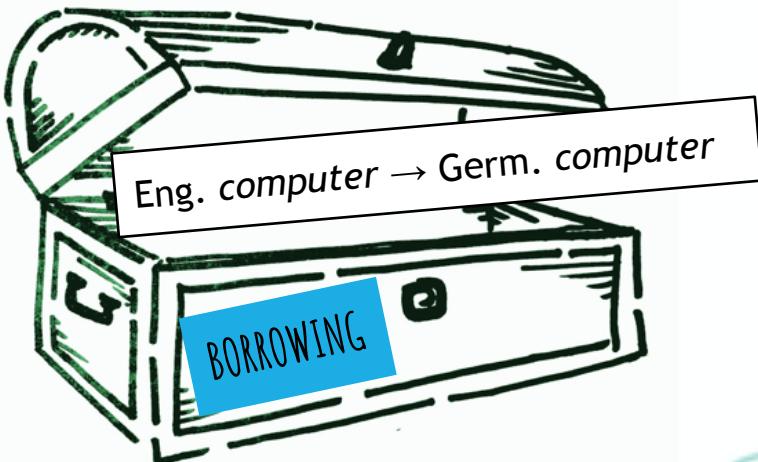
# It's happening now

English *computer* = Danish *computer* = German *computer*

But looks can deceive:

Faroese *telda* ‘computer’, lit. “counter”

This is a calque or loan translation



# It's happening now

But also *software* = *software* = *Software*

and Eng. *mouse* = Dan. *mus* = Germ. *Maus* (calques!)

A vocabulary event



# Vocabulary events

- Certain stable technologies or institutions
- entered the language at a given time as innovations

The occasional and one-off expansion of the vocabulary due to new technology (cf. computers worldwide) or social organization (e.g. religion).

- Metallurgy
- Agriculture
- Pastoralism



# TASK !

Write in the chat items unique or special to the culture or region your studying

Vocabulary events

- three new items
- region
- time range

Example: “bicycles, co-operative movement, pork (Denmark, 1800-2000)”



# just as it has happened historically

Historically, direct contacts are needed

One word, many languages

Massive spread

A sequence of borrowings, usually where the origin of the word is unknown, but this does not need to be the case.

## Wanderwort

(dialectal borrowing: Wander = wander, Wort = word)



# Chai tea

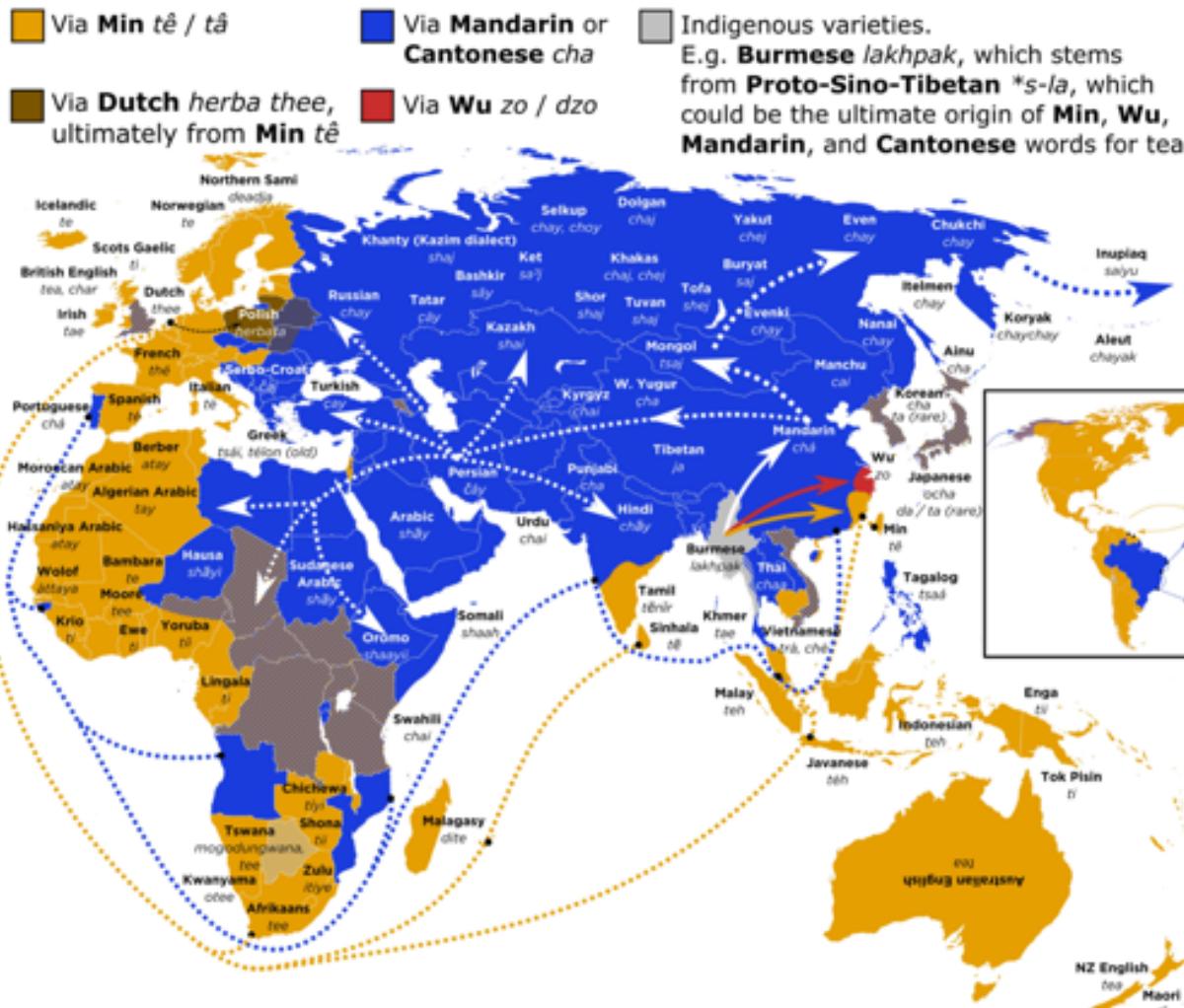
# Etymology of Tea

Reddit user Nuzarkhem

Write in the chat the word for tea in your language, specify language in parenthesis behind. Thus “te (Danish)”.  
Land (Central) vs. sea route (Min)

End up together

Attest to a  
of relation



# Silk on the road

Sea route not found yet



North and south of the Caspian sea

Old Chinese \*sir

Mongolic sirgek

Greek sērikon → Latin sericum

Sarmatian \*silika

Old Russian šelkū

English silk



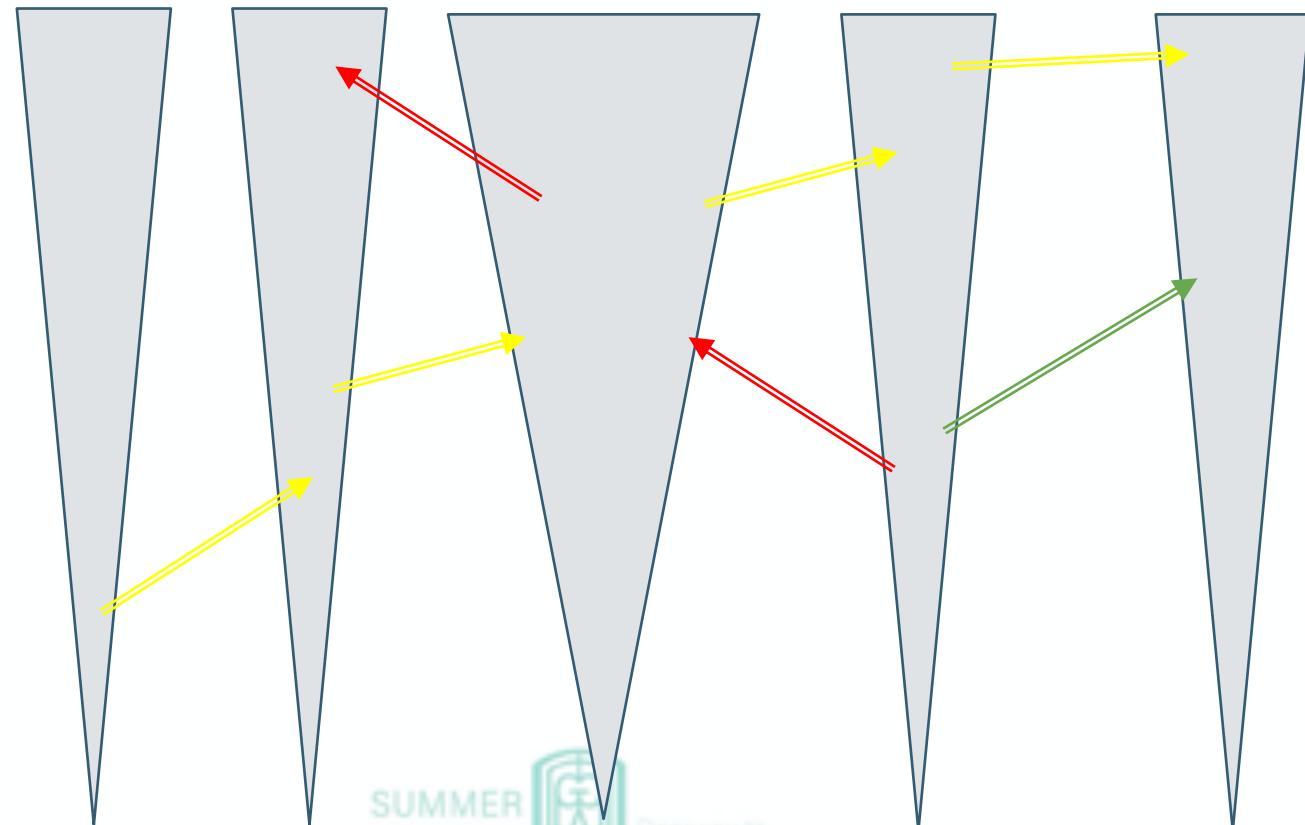
Mediated through Iranian dialects r/l alternations



# Methodology

## Loanwords intersect language families

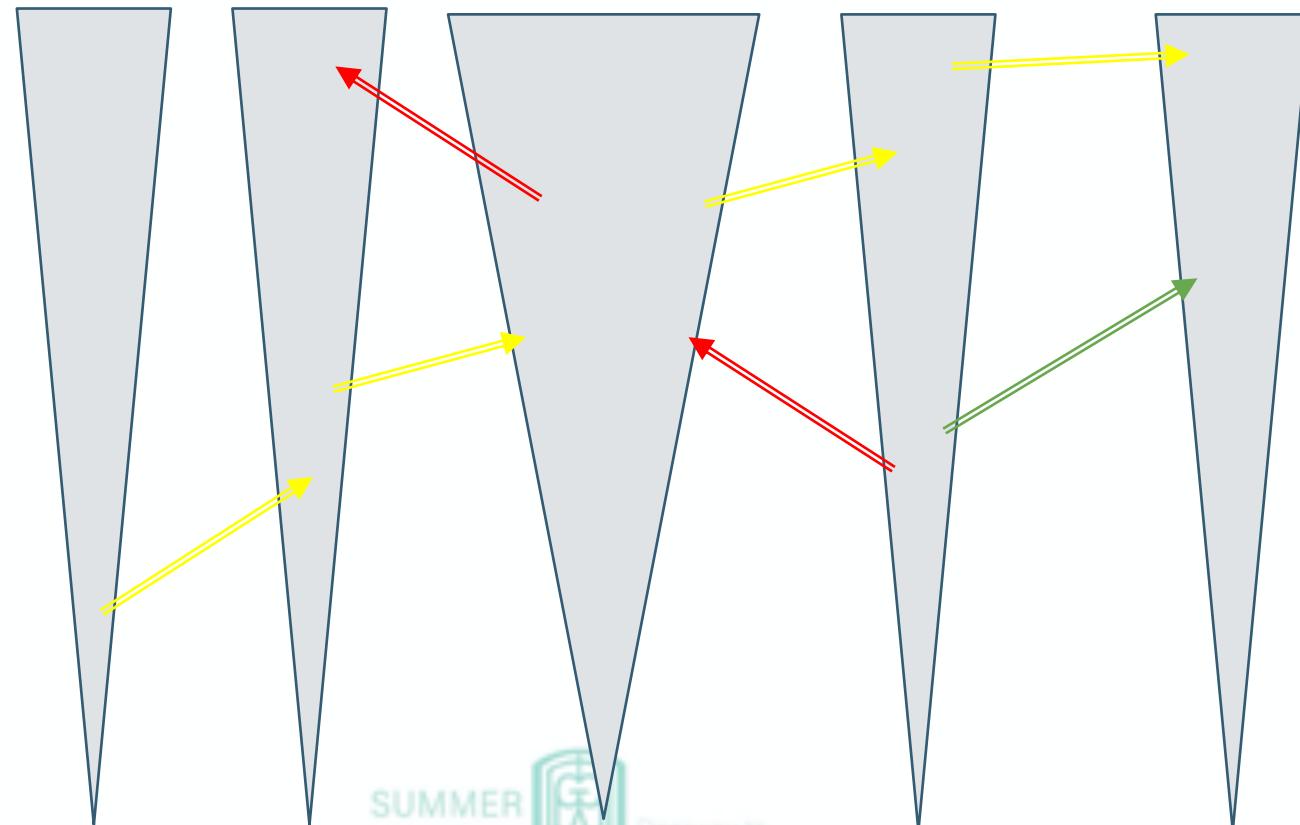
...and help date events - each family/branch has its own internal developments



# Methodology

But the reality is likely more complex

...and we can never expect to have all the pertinent data points



SUMMER  
SCHOOL  
2021

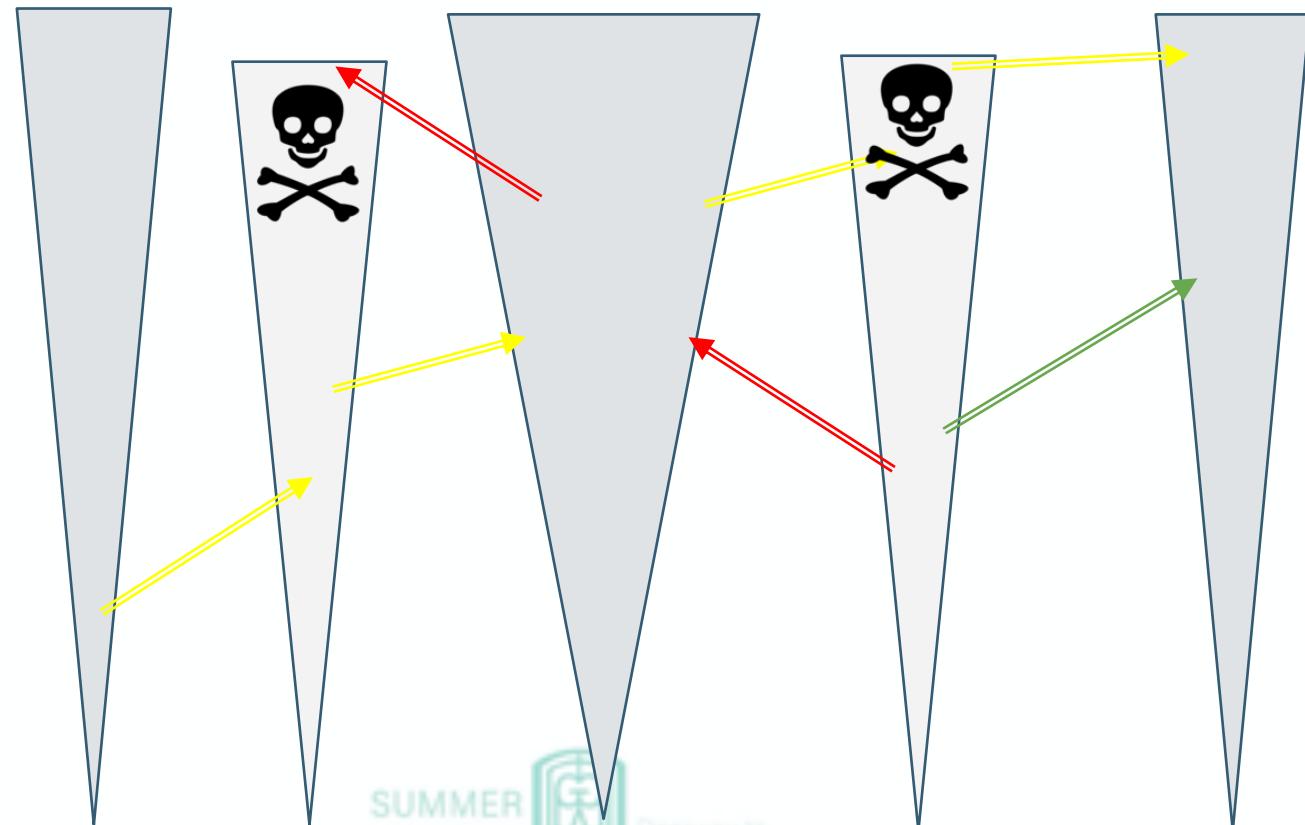
Doorway to  
Human History



# Methodology

...and we have to posit *hypotheses*

...to account for the observed phenomena



# Sustained contacts

Opposite of wanderwort phenomena

Many words in few languages

Altaic/Transeurasian problem

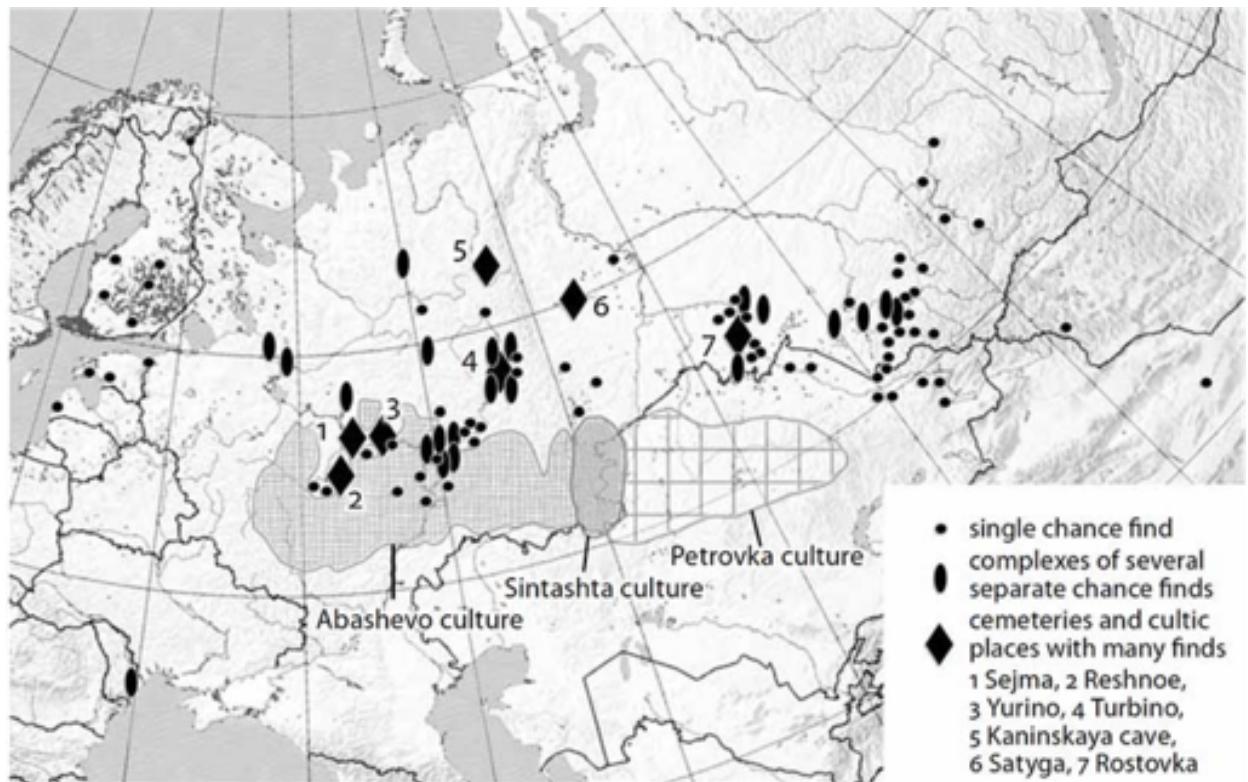
Germanic ~ Balto-Finnic (Kallio 2012)



# Indo-Iranian and Fennougric

Indo-Iranian → Fennougric (Koivulehto 2001)

‘lord’  
‘bee’  
‘awl’  
‘spindle’  
‘hundred’



# Narrative Historical Linguistics

Testing hypotheses



# Narrative Historical Linguistics (NHL)

Term coined by Ross (2021)



Aims to make a coherent history for the language family.

Not confined to loanword studies, but a strong device for hypotheses

Two basic approaches:

- Macro-hypotheses - new correlations, to be tested
- Micro-hypotheses - old correlations, new aspects



# *Every language family tells a story*

Ross (2021) on Austronesian

Ehret (2015) on Nilo-Saharan

Nichols (2021) on Uralic

Darden (2001) on Indo-European

Sagart et al. (2019) on Sino-Tibetan



# Vocabulary events in NHL

Vocabulary events central to language families:

Crossing the Wallace line for Austronesian (Ross 2021)

The horse and the chariot for Indo-European (Anthony 2007)

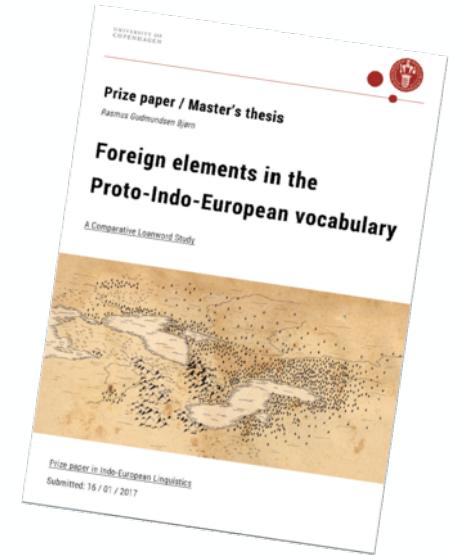
You may consider such vocabulary events for your own region



# Avoiding fiction

Working with incomplete data - making a good working hypothesis/narrative

- a. Be explicit about the assumptions
  - i. (e.g. Steppe homeland, Uralic on the Volga)
- b. Be explicit about things that still don't add up
  - i. (numerals, vocabulary)
- c. Be ready to revise
  - i. (Uralic NOT on the Volga!)
- d. As with any good hypothesis, it should respond to and, if viable, stand the weight of additional data points



# Possible ways forward

## Further linguistic evidence

- data
- analysis

## Circumstantial evidence

- Location and dating of speech communities
- If tied, then further archaeological or genetic perspectives



# EXCITING NEW PERSPECTIVES

Getting into the heads of pre-historic people



SUMMER  
SCHOOL  
2021



Doorway to  
Human  
History



# New hypotheses and old ideas

Working principle:

- Nobody publishes without reason
- Tweaks to a hypothesis make it stronger
- Combining different perspectives



# Case: Language and the Indus Valley Civilization

Ansumali 2021, Nature (found it via Facebook two days ago!)

Items emanating from the IVC

- i. Historical and linguistic evidence indicates that IVC was the most likely origin of those proto-words.
- ii. Archaeological evidence indicates that the objects signified by those proto-words were prevalently produced and used in IVC.
- iii. Those proto-words' etymologies can be traced back to one of the language groups present in the Indian sub-continent using the most stringent criteria prescribed in historical linguistics.
- iv. Historical and linguistic evidence should prove that the identified language group was present in the Indian sub-continent in the IVC era.
- v. Archaeogenetic data should prove that present speakers of this language group are significantly genetically related to the IVC population.
- vi. The proto-words should be etymologically related to the stable non-borrowable basic vocabulary of a language, which can trace out a language's ancestry.



# The elephant in the room

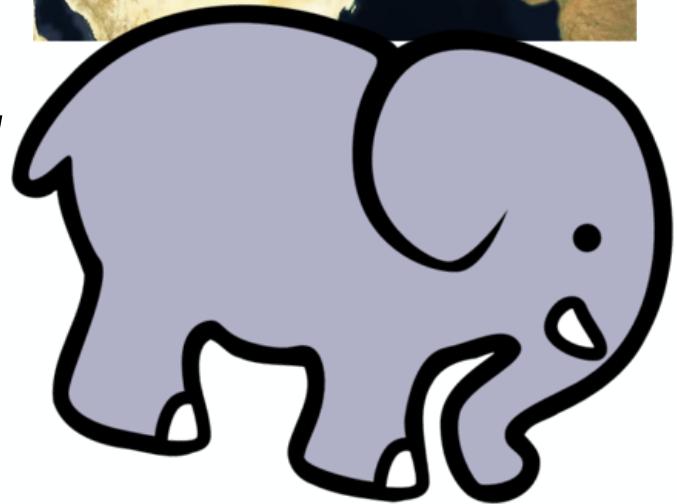
And she found just that:

Ivory in Akkadian and Old Persian \*pi(:)Ru

Archaeological evidence

Dravidian root ‘tooth’

Dravidian spoken in the Indus Valley Civilization



# Case: Old Balkanic

Old Balkanic - the language of the first farmers in SE Europe

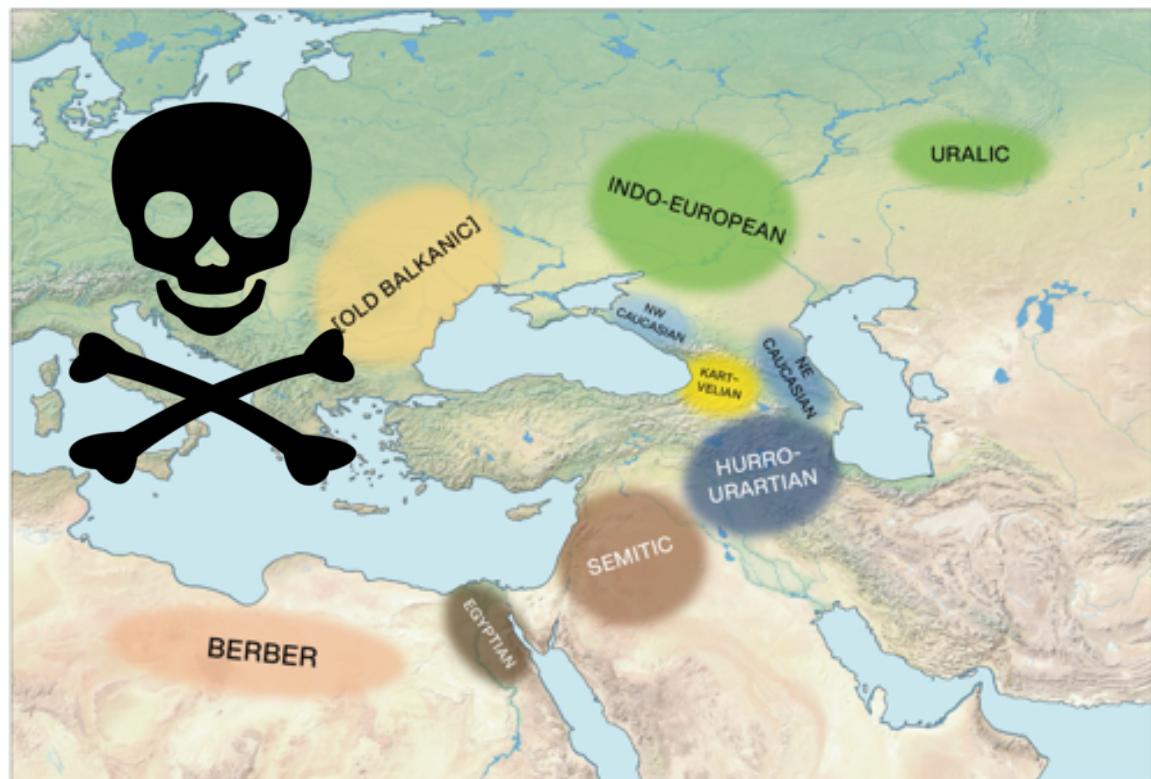
Semitic loanwords in Indo-European?

- domestic animals
- crops
- rituals
- numerals

Arch. match →

Substrate phenomena

- Semitic?
- Egyptian?
- Berber?



# Case: 7

Write, in one message and Latin letters, the numeral ‘seven’ in all the languages you know.

e.g. “syv (Danish)”

In the literature, ‘seven’ was often mentioned as a potential borrowing in separate clades.

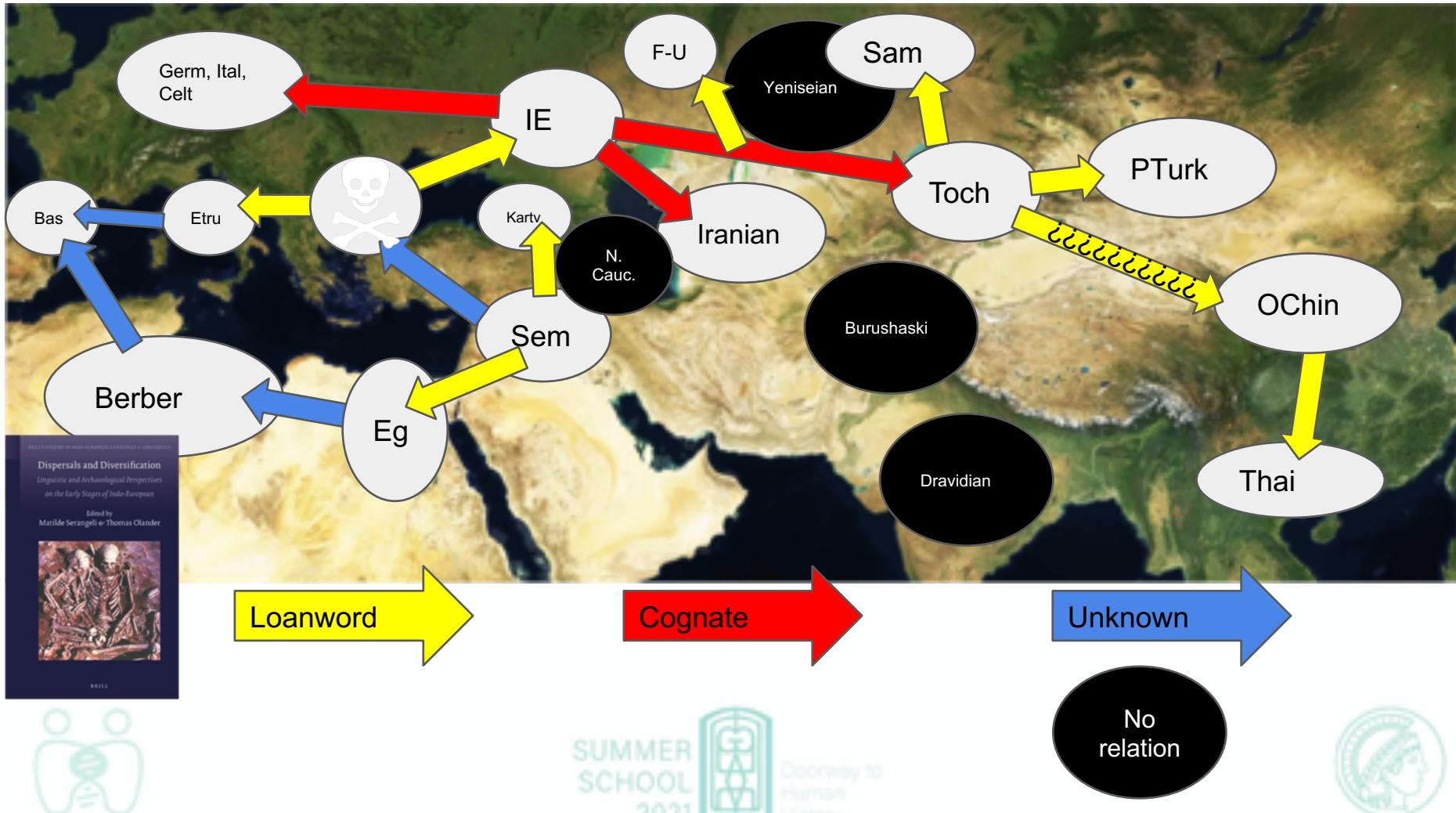
Elsewhere, the idea of a “numeral spread event” was mentioned.

But do the observations belong to a common phenomenon? or are they merely due to chance?



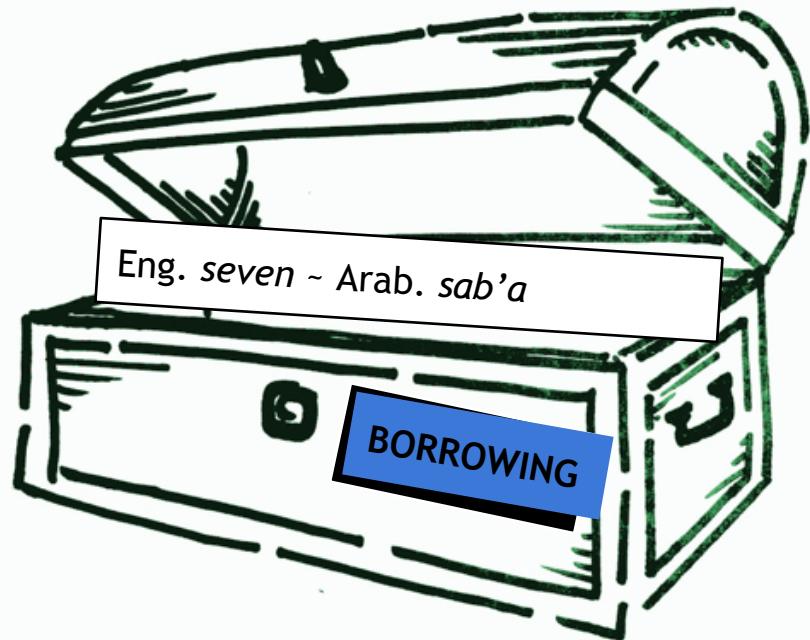
# The spread of seven

PSem \**tsab̥a-(t-Vm)*, ~ PIE \**septm̄*, ~ Egyptian *śfh-w*, ~ PBerb. \*sa, ~ PKartv. \*świd-, ~ Hurrian *śind(i)*, ~ Basque *zazpi*, ~ Etruscan *semph*, ~ (P)Uralic \**sejt̊ew*, ~ PTurkic \*(s)yet(t)i, ~ Old Chinese *sjet* → Thai ced



# TSEPTM WORLD TOUR

SEMITIC 5500/BC  
OLD BALKANIC 5000/BC  
INDO-EUROPEAN 4700/BC  
EGYPTIAN 4500/BC  
BERBER 4000/BC  
BASQUE 3700/BC  
KARTVELIAN 3500/BC  
ETRUSCAN 3300/BC  
URALIC+TURKIC 3000/BC  
OLD CHINESE 2700/BC  
THAI 1000/AD



# NOW IN A LANGUAGE NEAR YOU!



SUMMER  
SCHOOL  
2021

Doorway to  
Human  
History



# Social structures

Persistent and socially important concept

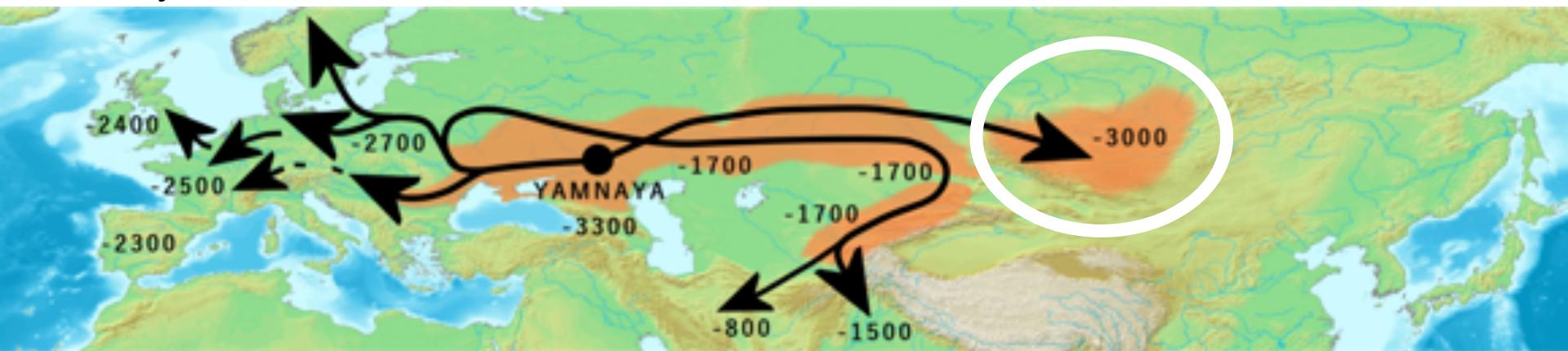
We saw ‘lord’ borrowed from Indo-Iranian into Fenno-Ugric

Similarly Proto-Germanic \**kuningaz* ‘king’ → Finnish *kuningas*



# Making a name for yourself?

Early Proto-Tocharian = Afanasievo



Relevant forms, (1) ‘name-famous’; (2) ‘name’; (3) ‘famous’

Proto-Indo-European	* <i>h<sub>3</sub>neh<sub>3</sub>-mn-ḱleūos</i> (1)	Proto-Uralic	* <i>nime</i> (2)	Proto-Turkic	*( <i>at</i> ) <i>kü</i> (1)
> Proto-Tocharian	* <i>ñæm-klyäwæ</i> (1)	Proto-Yukaghir	* <i>nime</i> (2)	> Old Turkic	<i>kü</i> (3),
> Tocharian A	<i>ñom-klyu</i> (1),	Chukchi	<i>ninn</i> (2)	Uighur	( <i>at</i> ) <i>kü</i> (1)
Tocharian B	<i>ñem-kälywe</i> (1)	Old Japanese	<i>na-</i> (2)	Ainu	* <i>namup</i> (2)



# Coping with comparanda

- My personal experience with an ancient loanword:

(P)IE \**b<sup>h</sup>ei-*(*t/k-*) 'bee' > OIr. *bech*; ON *bý*; Lith. *bitė*; OCS *bičela*; (?)Lat. *fūcus* 'drone; hornet, wasp';  
~ Egyptian *bj.t*  
(cf. Gamkrelidze & Ivanov 1995: 773)

- 2016: Laughable
- 2017: Tentative analysis
  - "further circumstantial evidence need be provided before a more favorable scenario for their relation can be envisaged, tentatively through the proposed Neolithic substrate (2017: 48)
  - "...unique contacts extremely improbable" (2017: 138)
- 2019: Application on numerals
  - "...an Afro-Asiatic language closely related to Semitic in contact with Proto-Indo-European..." (2020: 69)
- 2020: Fully formed hypothesis
  - "...a logical chronology first importing the produce 'honey', before adopting full-scale apiculture later..."



MAX-PLANCK-GESELLSCHAFT



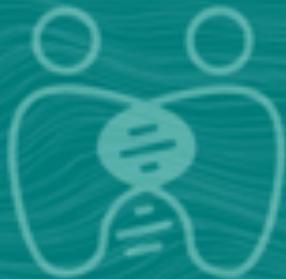
# Loanwords in your research

Identify cultural interfaces

- Regions and times of potential exchange
- Vocabulary events

Have a look at the words you took down at the middle of the lecture.





Thank you for your  
attention!

Q&A away!

MPI-SHH  
SUMMER SCHOOL  
2021

Doorway  
to Human History