A diachronic linguistic geography for Uralic

and what archaeology and genetics might contribute

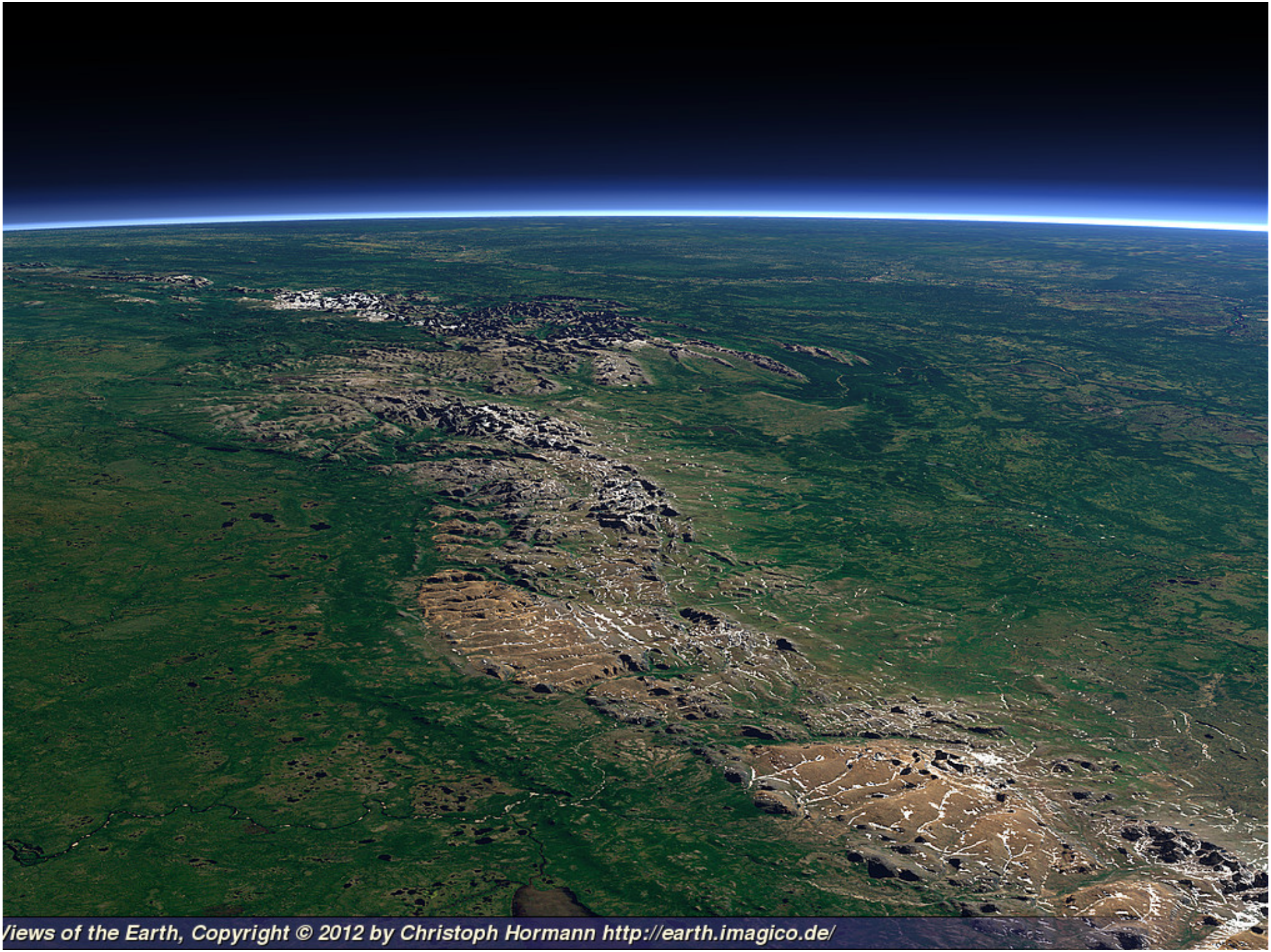
Johanna Nichols johanna@berkeley.edu
University of California, Berkeley
Linguistic Convergence Laboratory, Higher School of Economics, Moscow
University of Helsinki

Linguistics, Archaeology, & Genetics 2:

The origin and expansions of Uralic-speaking populations

MPI-SHH, Jena Nov. 29-Dec. 1, 2017





iews of the Earth, Copyright © 2012 by Christoph Hormann http://earth.imagico.de/

Previous slide: The Ural mountains.

Questions:

Did speakers of Proto-Uralic ever see these mountains? hear about them?

If so, from east or west?

And from home or on distant travels?

Likely answer to all: Yes.

Issues

Date for the Uralic dispersal; for the Finno-Ugric dispersal
Location of Uralic homeland; Finno-Ugric homeland
Cause(s) and mechanism(s) of the Uralic spread

Outline of argument:

Language movement trajectories on and around the Eurasian steppe
The Fur Road
Uralic typology, Indo-European contacts
Proto-Uralic and Proto-Finno-Ugric in space and time
Contacts

General trends of language spread and distribution in Eurasia

The Eurasian steppe is a spread zone
Accretion zones at the steppe periphery
Northward spreads from the steppe-forest interface

The Eurasian steppe spread zone

Dominant spread direction east to west -- from the Bronze Age on.

Earlier, more complex; domestication frontier spread W > E from W steppe.

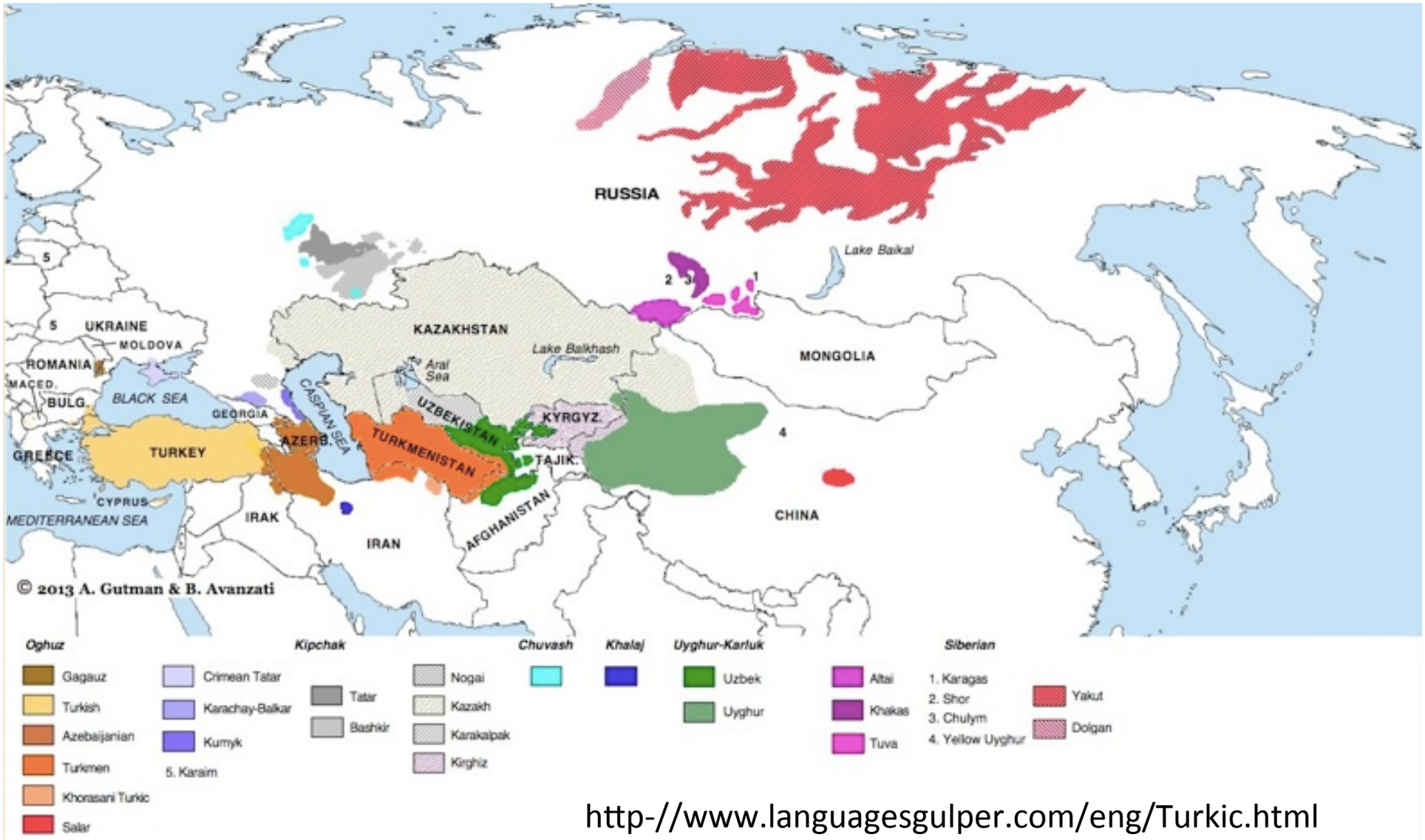
East to west: Iranian, Turkic (several languages each), Mongolic

Northward spreads

Across the entire forest-steppe ecotone, languages spread north:

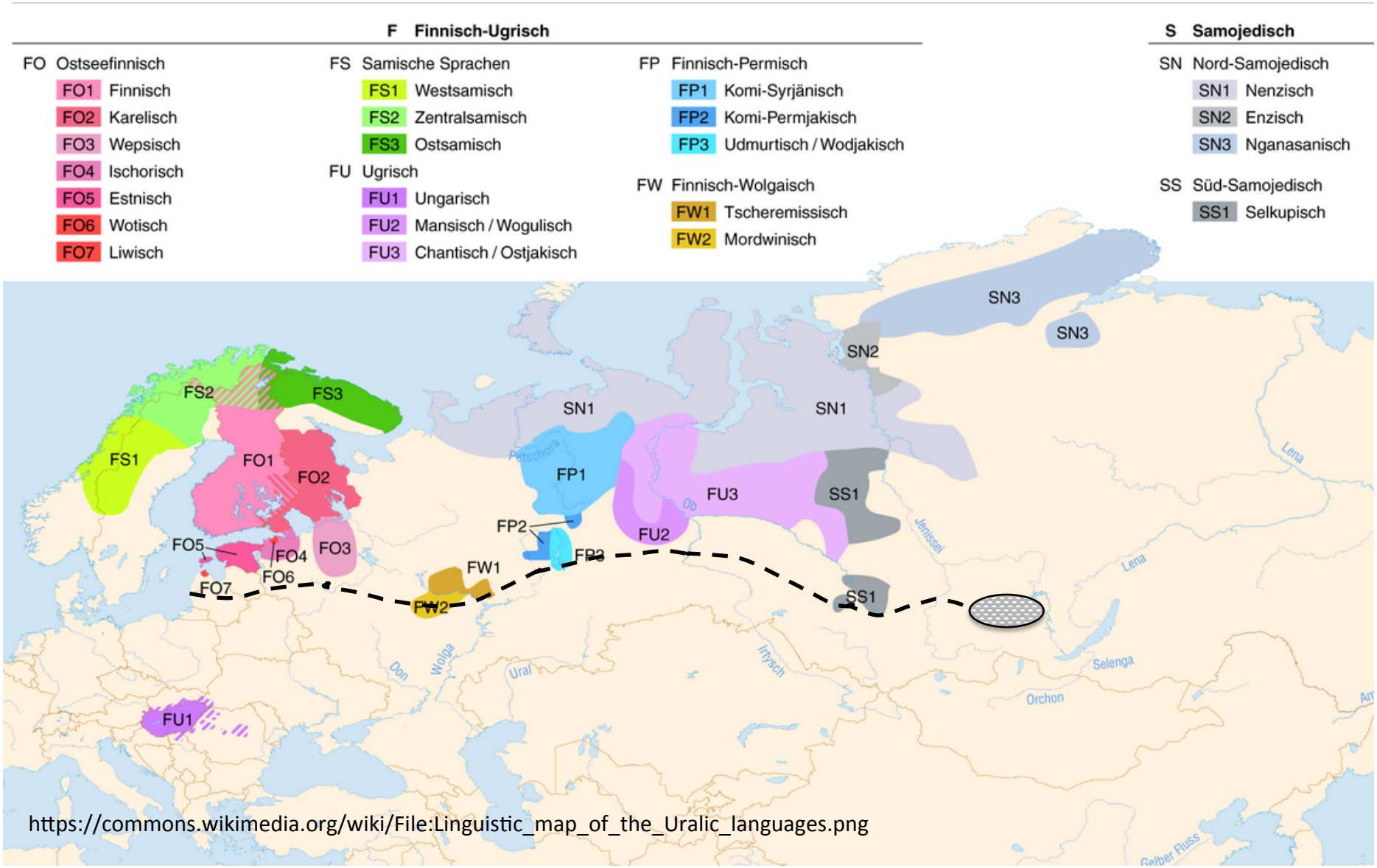
Germanic, Saami, Finnic, Permian, Khanty, Samoyedic, Yeniseian, Turkic, Tungusic

Turkic languages. Sakha/Yakut (red) represents a northward spread.



http-//www.languagesgulper.com/eng/Turkic.html

Uralic languages Dashed line: Former southern limit of Uralic; compactly Uralic to the north. Stippled oval: approximate former range of extinct Southern Samoyedic.
Northward spreads: Saami; Finnic; Permic; Khanty; Samoyedic



https://commons.wikimedia.org/wiki/File:Linguistic_map_of_the_Uralic_languages.png

Accretion zones (a.k.a. residual zones)

Attractor locations: languages enter more often than they leave; more often survive than often go extinct; undergo contact but not much shift. Minimal spreading inside the area. Result: diverse and old language populations.

Accretion zones, cont.

Languages move in; rarely move out.

Caucasus: No spreads out.
Balkan: No spreads out (since the original spread of agriculture).
E Circum-Baltic: Only Saami-Finnic (northward spread)
Middle Volga: Maybe Permic northward spread (if from nearby)
Upper Yenisei: Samoyedic northward spread; Yeniseian northward spread

Center-of-gravity illusion:

Distant sisters from distant parts of the catchment zone end up in the accretion zone (a pile-up). Examples:

All four branches of Iranian represented in Central Asian Mountains (Edelman 1968)
Both Turkic branches are found in the Middle Volga area (Chuvash, Tatar)
Several IE branches in Balkan peninsula
All Baltic languages in the E Circum-Baltic area

Summary on accretion zones

They may **look** like centers of phylogenetic gravity, but these are secondary accumulations.

Placing a proto-homeland in an accretion zone is risky and requires very strong argumentation

(much stronger argument than mere phylogenetic diversity in the area)

Especially risky when that accretion zone is at the edge of a major spread zone

The Fur Road (Barfield 2009)

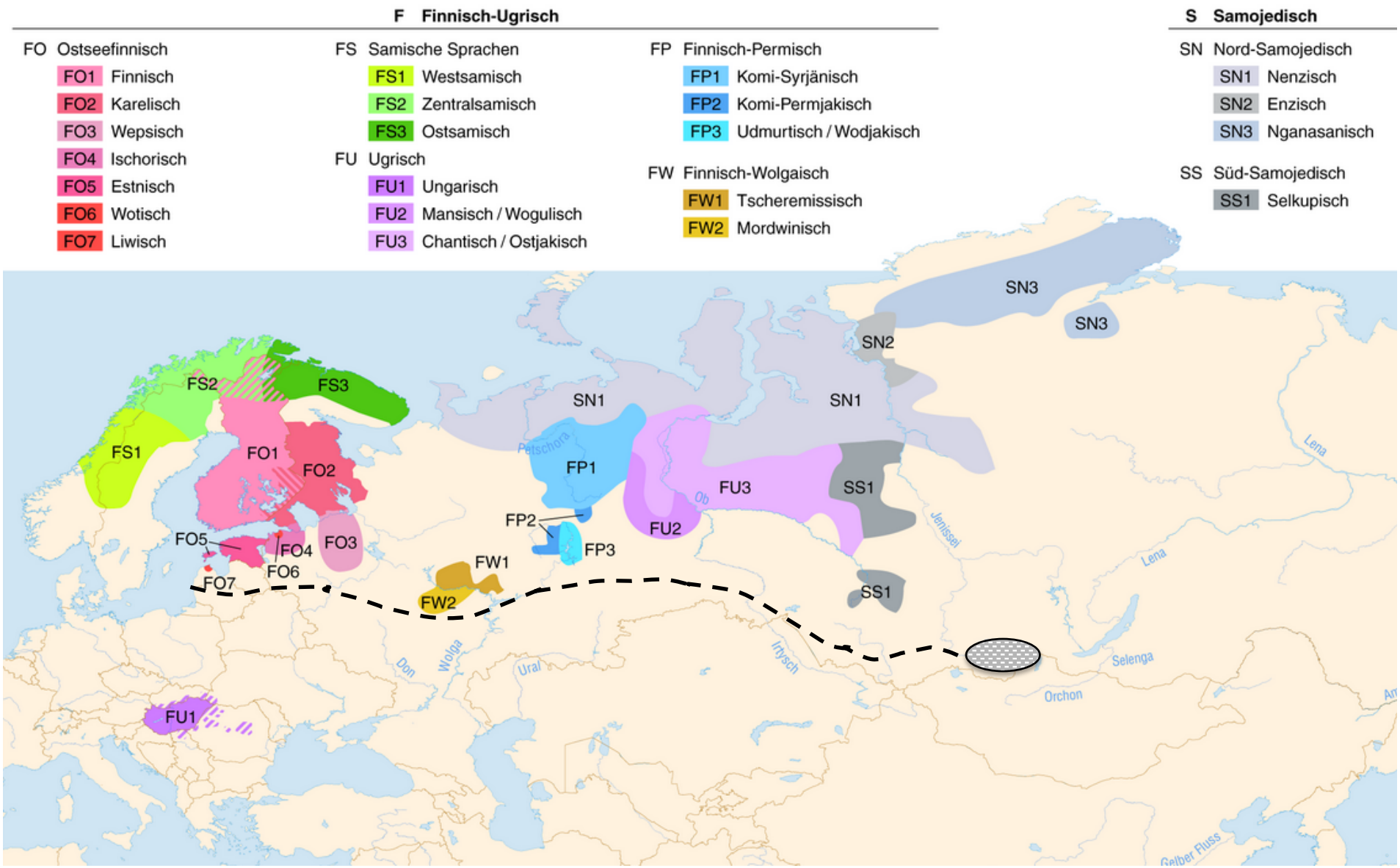
In the northern U.S.-southern Canada, extensive trade networks along rivers
- archaeologically (near-) invisible
- probably accounts for at least some of the Algonquian spread
- back-and-forth dominance of Cree and Ojibwe in protohistorical and historical times
- a major trade item: Plains buffalo hides

In the Eurasian forest-steppe zone north of the western and central steppe, Fur Road ran along the major east-west rivers.

Economic mechanism: Furs from northern forests traded to European and steppe wealthy. Traders, trading posts, trade colonies along the rivers.
- Major trade hubs at confluences with north-flowing rivers.
- Trade language spread along rivers, and was adopted by hunters closest to trading posts. Gradual northward spread.
- Likely mechanism: local chiefs' daughters marry rich traders (as with French fur franchises in North America)

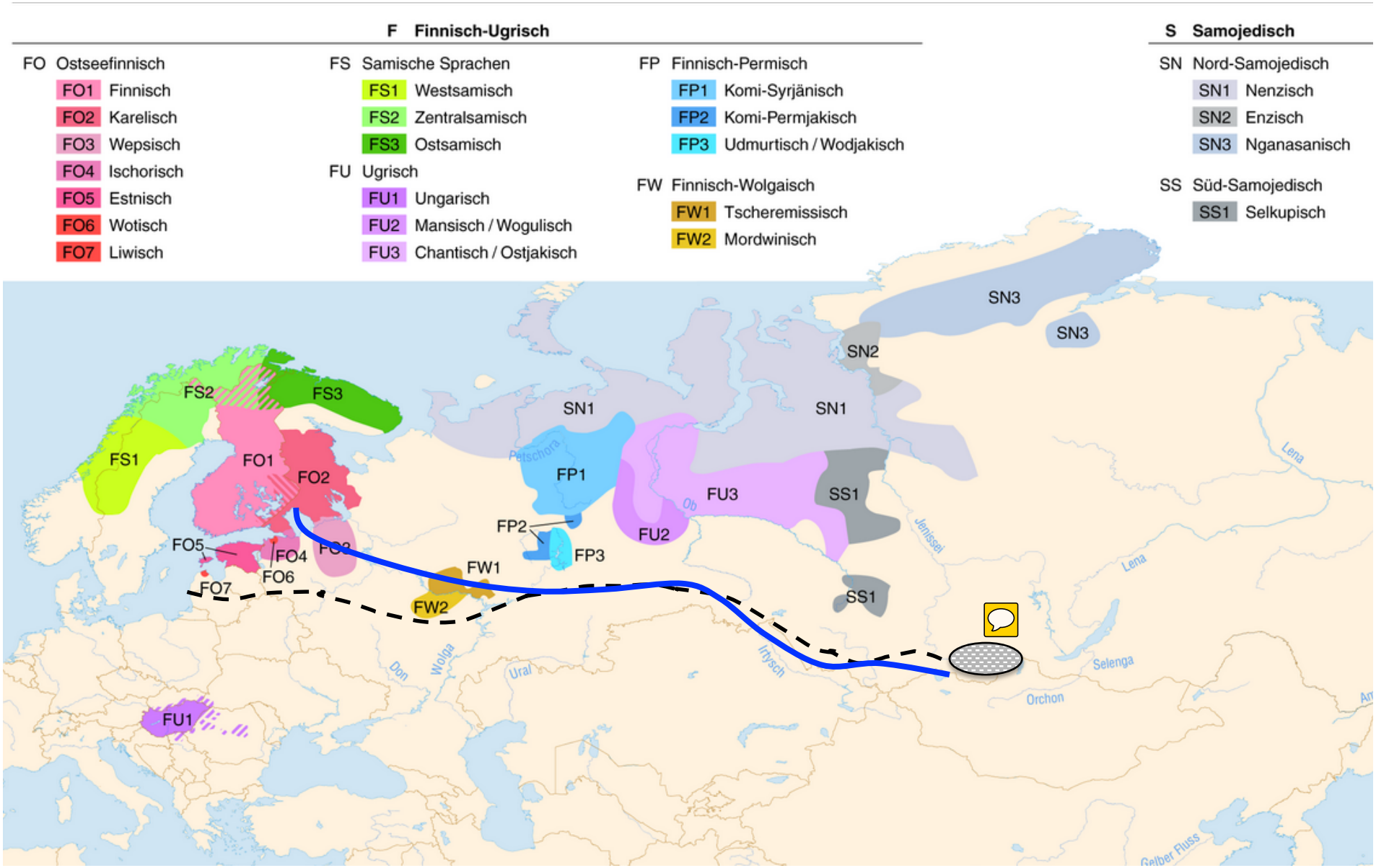
Barfield 2009; Ives 2003, 2006, others; Nichols & Rhodes in press

Uralic languages Dashed line: Former southern limit of Uralic. Compactly Uralic to the north. Stippled oval: approximate former range of extinct Southern Samoyedic.



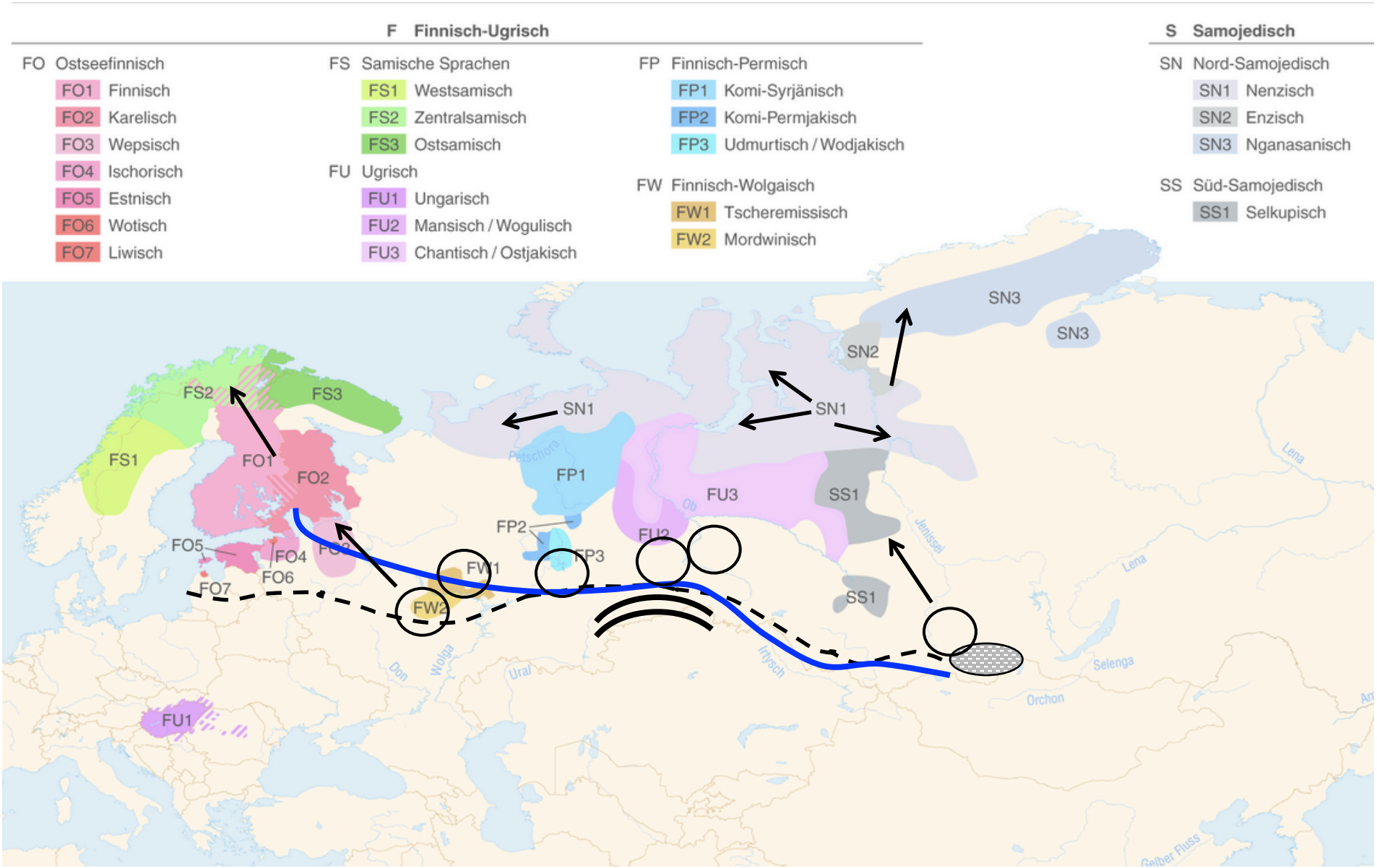
https://commons.wikimedia.org/wiki/File:Linguistic_map_of_the_Uralic_languages.png

Uralic languages Blue solid line: Fur Road.



https://commons.wikimedia.org/wiki/File:Linguistic_map_of_the_Uralic_languages.png

Uralic languages Circles: approximate branch homelands



https://commons.wikimedia.org/wiki/File:Linguistic_map_of_the_Uralic_languages.png

The Fur Road

The only Eurasian language family whose spread could have come as a Fur Road trade language is Uralic.

Are northward spreads inevitable?

(Are spreads to higher latitudes inevitable?)

Not the rule in North America.

Athabaskan	Southward
Algonquian	Predominantly eastward
Siouan	west to Plains
Salish	Radial (north, south, inland)
Chinookan	Upriver (east)
"Penutian" families	West and south

Not the rule in Africa, South America *
Partly in Australia *

* For what it's worth; cone-shaped geographies must affect spread trajectories

Why northward only in Eurasia?

Fur trade, from northern forests to steppe and Europe
(also amber, metals, finished metal items east-west)

vs.

Buffalo hide trade, Plains to elsewhere

Barfield 2009; Ives 2003, 2007, others

Proto-Uralic typological geography

Typologically, Proto-Uralic clusters with the (greater) Pacific Rim population of language

High causativation, inflectional person, fairly high POS flexibility, head-final, …

and Proto-Indo-European clusters with (what can be extrapolated to) the western Eurasian population

(and both PU and PIE belong to the northern Eurasian linguistic population)

Nichols 2007, 2017

An example (one of several typological variables with this distribution)

Realization of the causative alternation

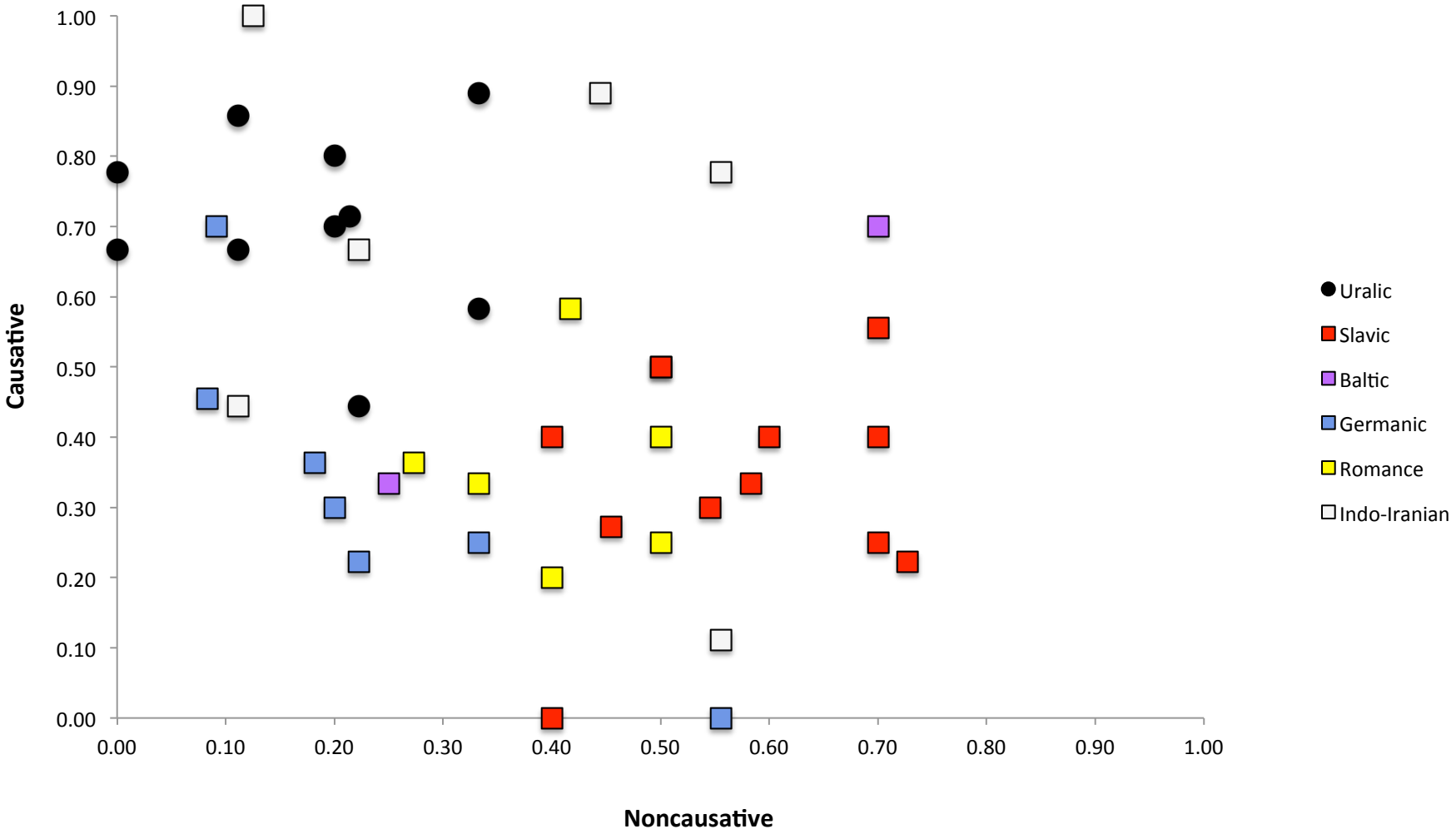
	'sit down'	'seat, have/let/make sit'	
Spanish	sentar=**se** seat=REFL	sentar seat	(decausativizing)
Ingush	wa-xou down-sit	wa-xoa-**d.u** down-sit-CAUS	(causativizing)

(blue = relevant derivational morphology)

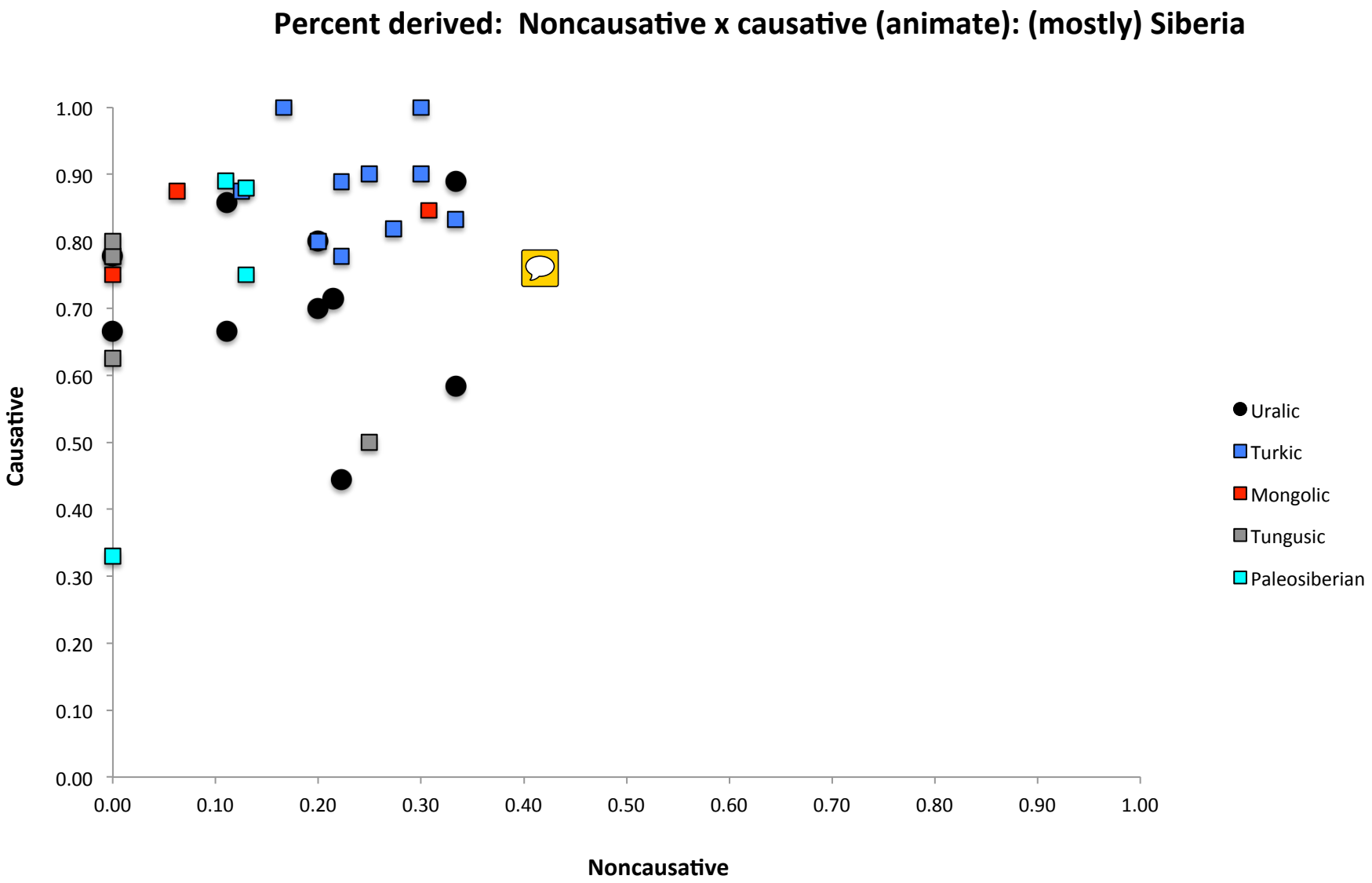
These are two of several possible patterns of derivation.

Grünthal & Nichols 2017, Grünthal et al. in progress, Nichols et al. 2004, …, Nedjalkov 1969

Percent derived: Noncausative x causative (animate): (mostly) W Eurasia

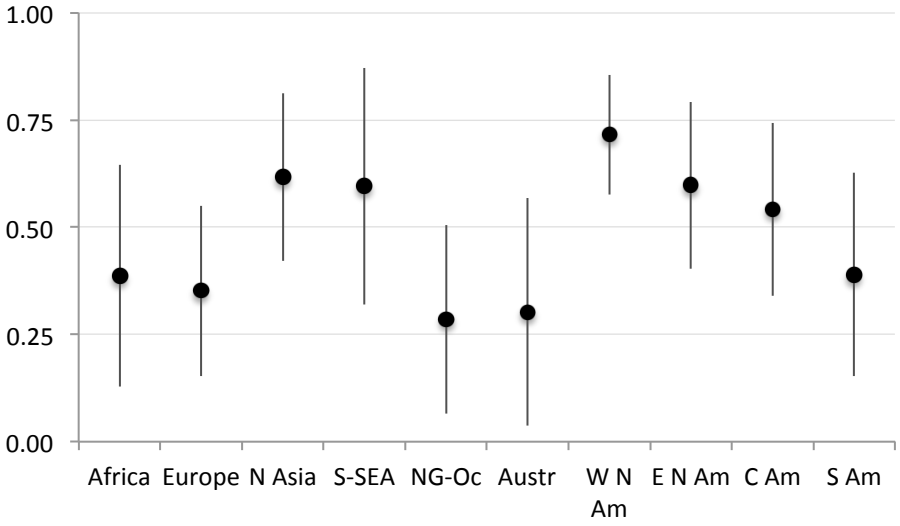


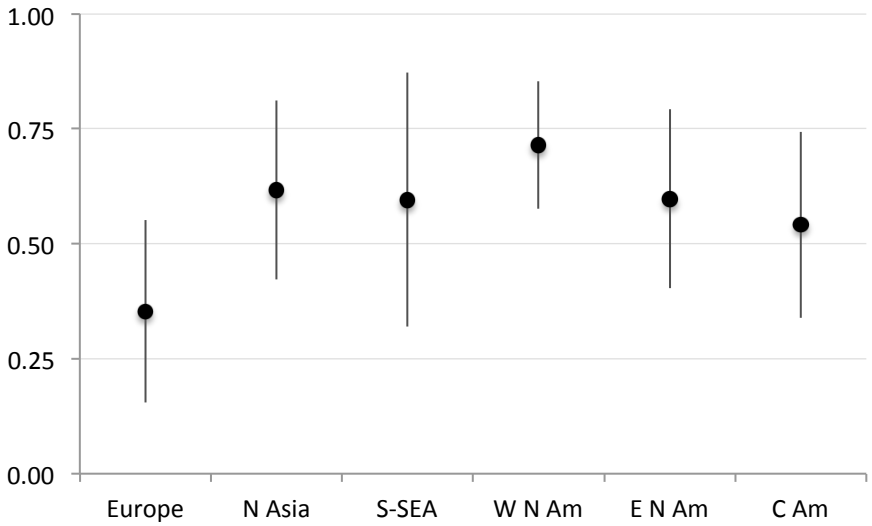
NB: Branch sisters cluster together. Uralic very different from European branches, similar only to Indo-Iranian (100% language at top left is Hindi).



NB: Uralic fits well into Siberian languages.

Causative alternation: Continental mean percent causativized.
Left: all continents; right: N. hemisphere. Mean ± 1sd.



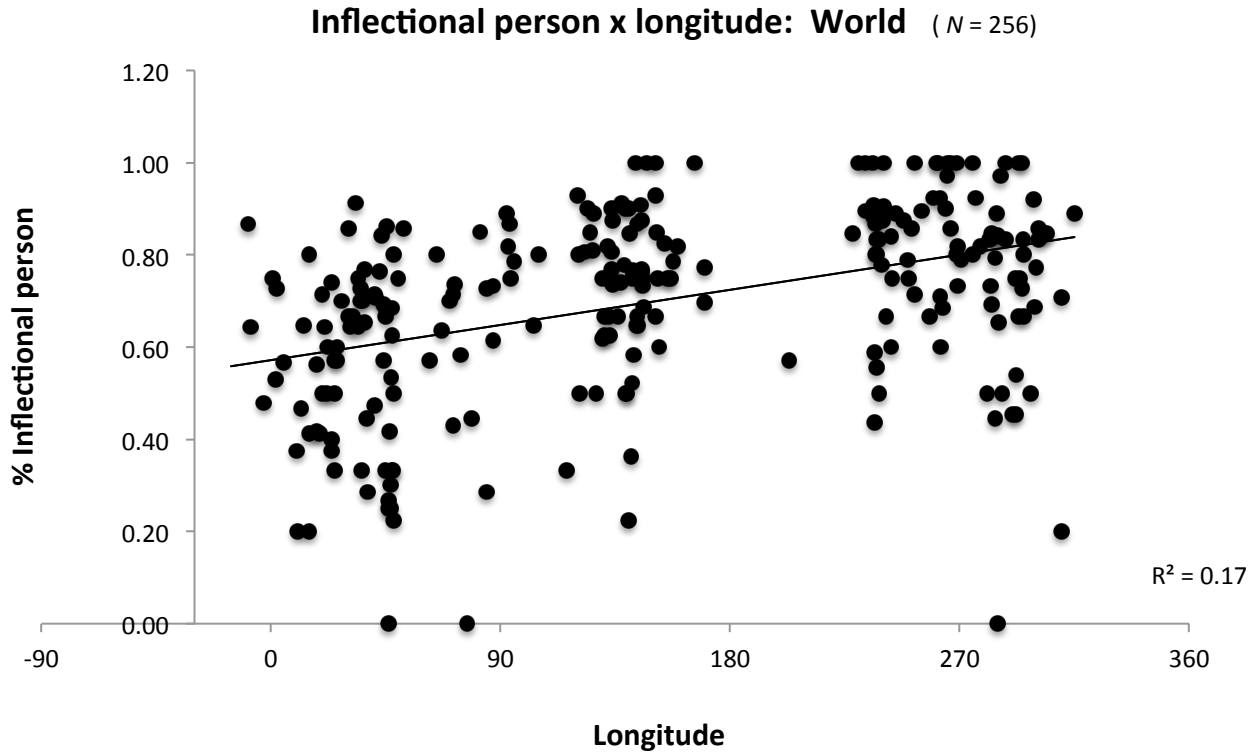


Distribution: Peak at North Pacific Rim.

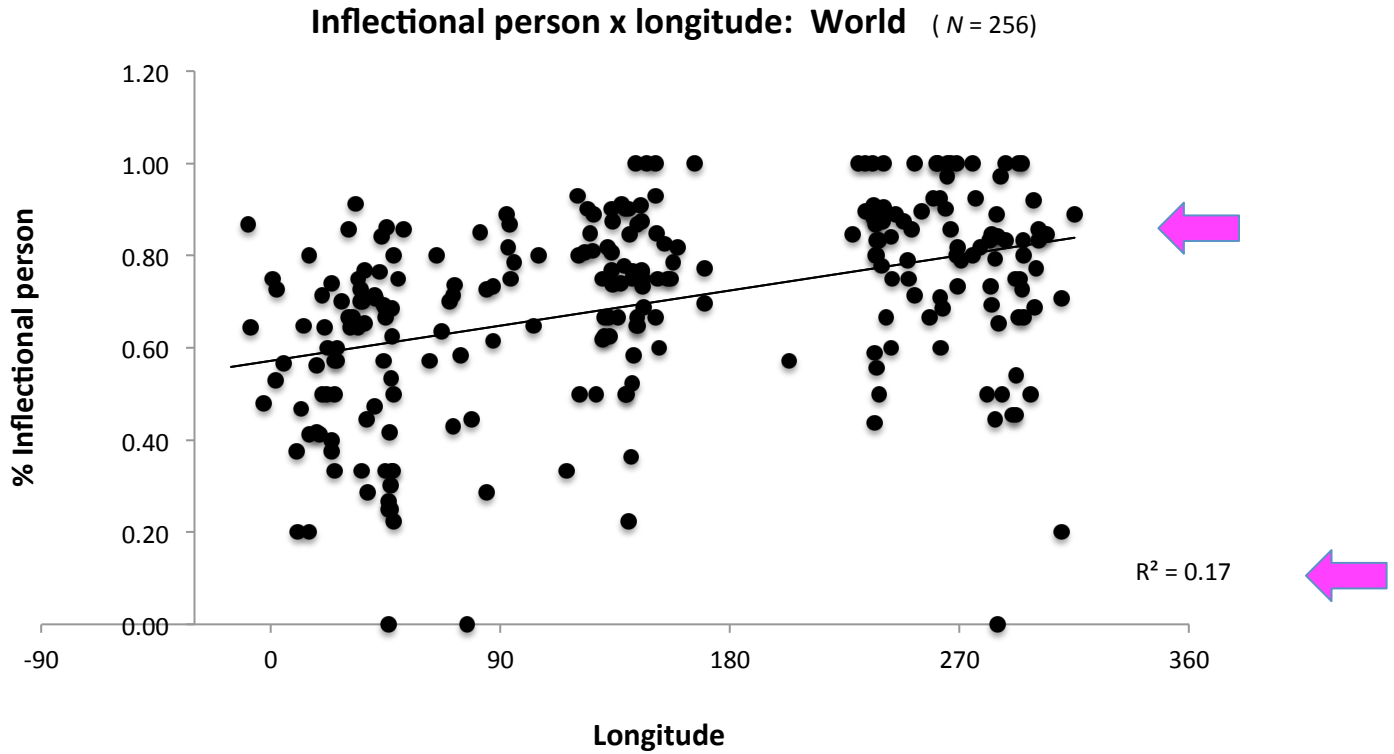
Distribution: Fortescue 1998. AUTOTYP continent breakdown (Nichols et al. 2013)

Another example: Inflectional person

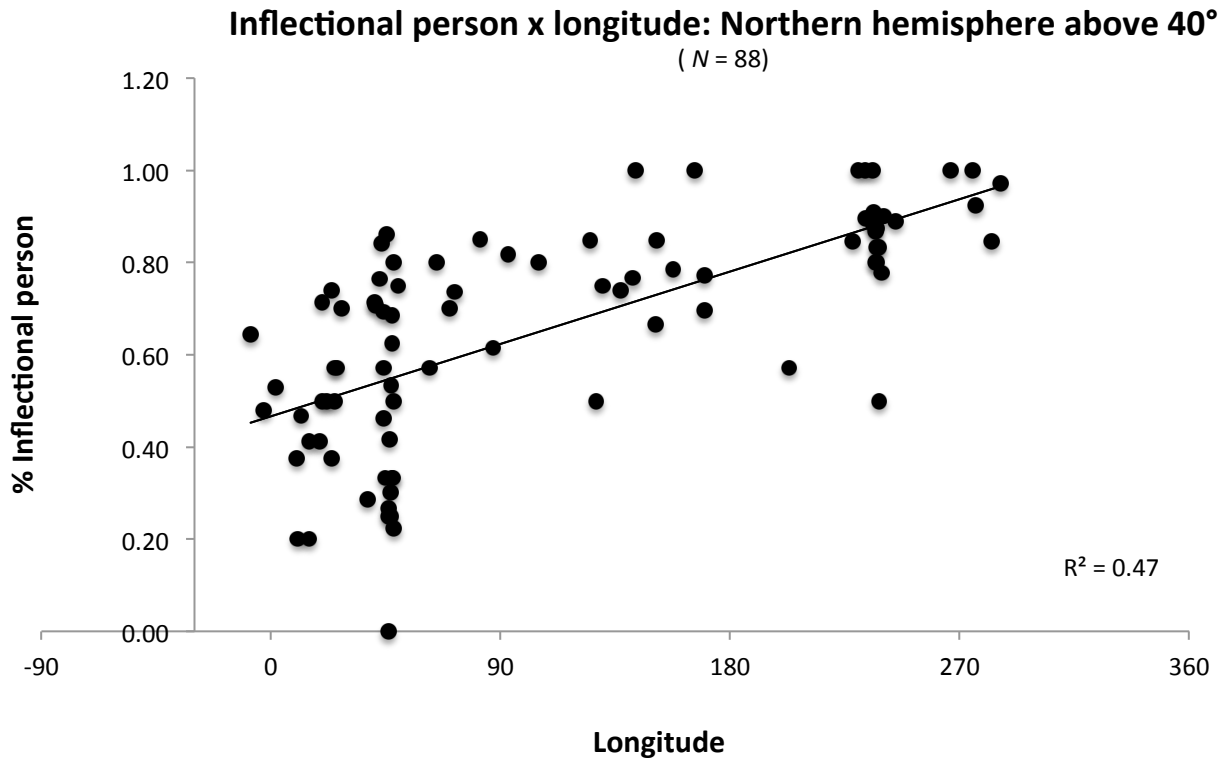
(the extent to which the category of person behaves like inflectional morphology, vs. like lexical items)



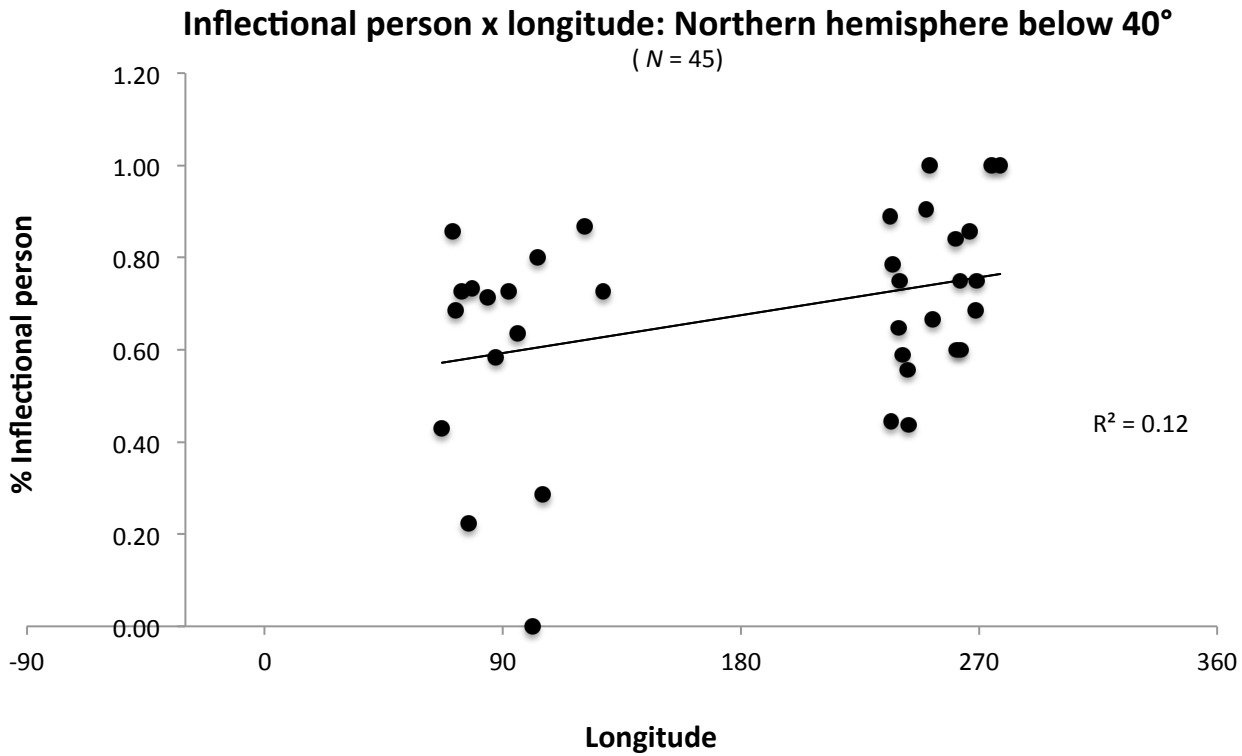
Europe, Africa Asia, New Guinea-Oceania, Australia Americas



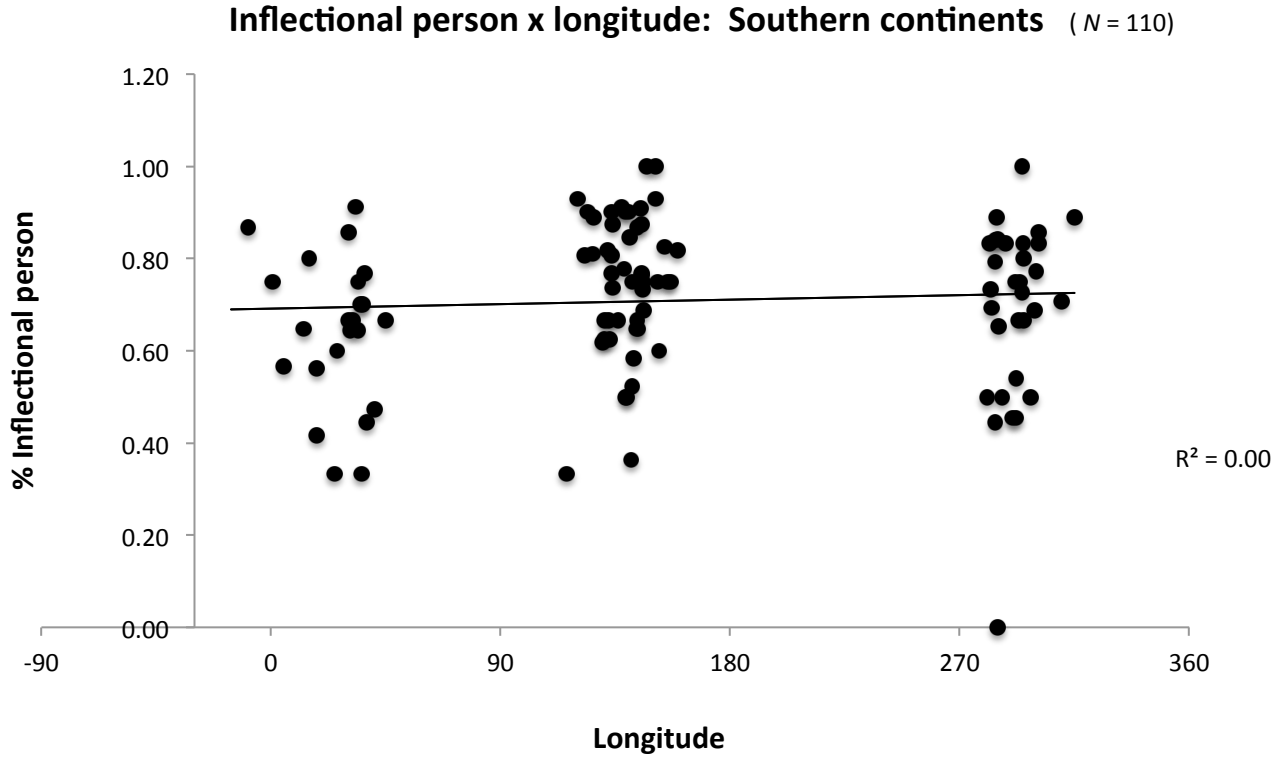
What to look for in these graphs:
Trendline slope. Is there a cline?
R^2 number: Strength of correlation. 0-0.5 none 0.10 weak 0.15 moderate 0.20 strong
(Linear trendline. Correlation coefficient is 0.412.)



Strong global cline across entire higher-latitude Northern Hemisphere.

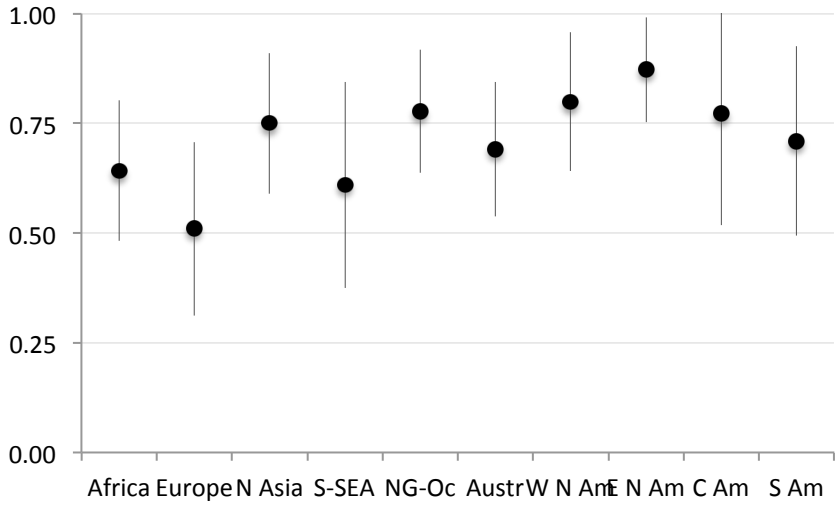


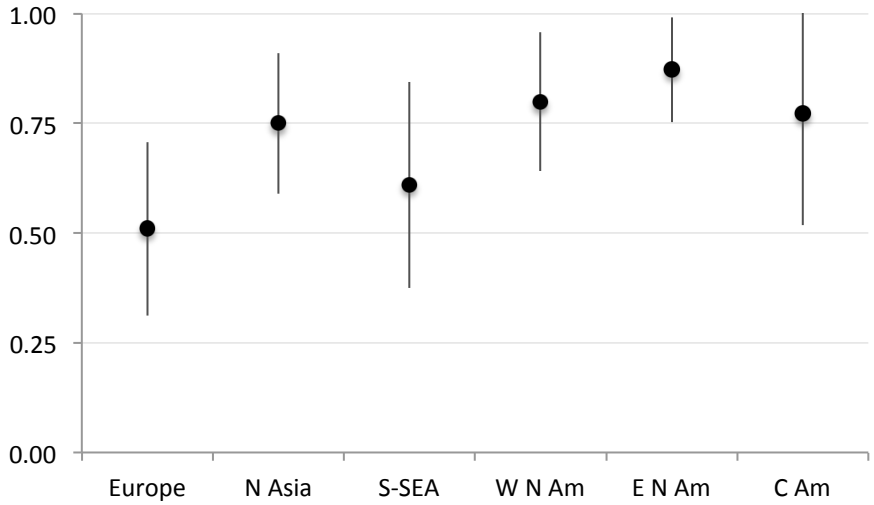
Moderate cline across northern lower latitudes.



No cline across the southern continents (Africa, NG-Australia, S. America).
(No surprise; no connections among them.)

Inflectional person: Continental means (±1 sd).
Left, all continents (*N* = 255); right, northern hemisphere only.

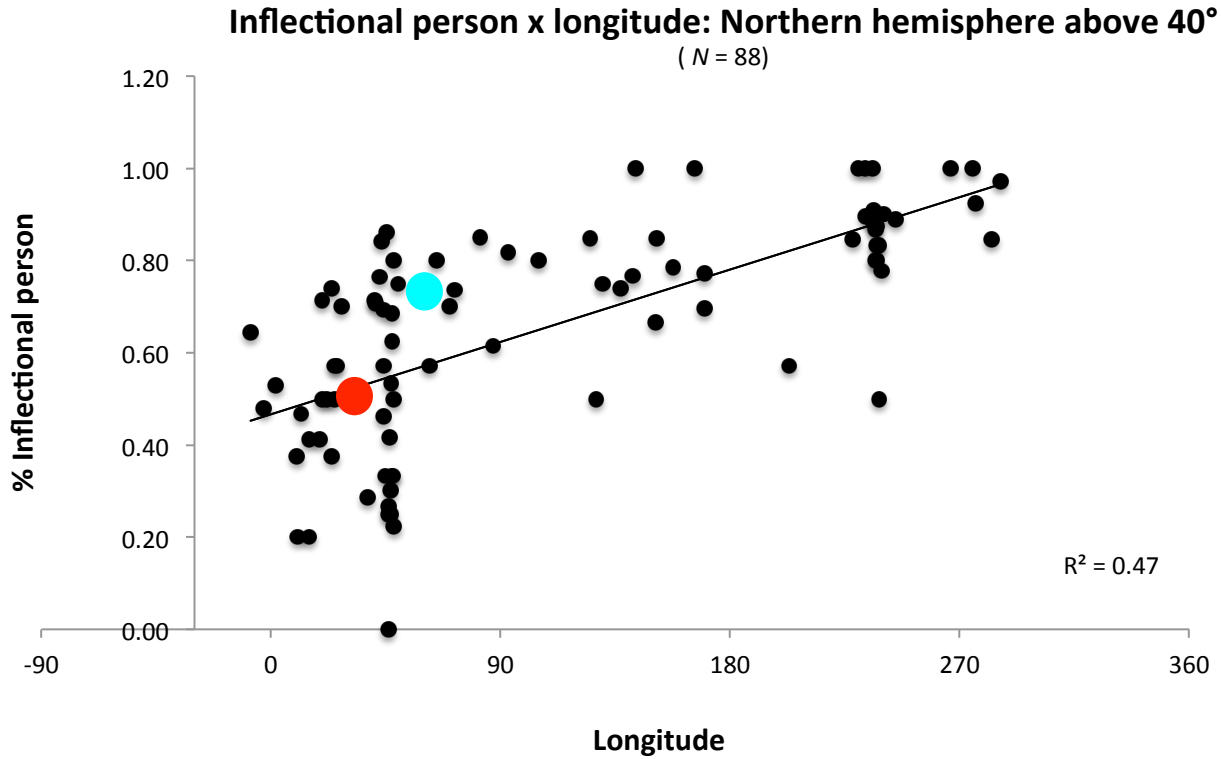




Distribution: E-W cline over entire Northern Hemisphere.

PIE (red) and Proto-Uralic (turquoise)

Reconstructed typological values at approximate homelands (Dniepr, Urals)



PIE is at the typological mean for languages to its west. PU is near the mean for languages to its east. Likewise for many other variables.

What Afanasievo* tells us about Samoyedic and early IE

3 early eastward outposts of IE, inchronological order:

- Afanasievo (upper Yenisei) (from Yamnaya culture)
  - Language unknown, material culture early IE
- Andronovo (NE Urals, N Kazakhstan) (maybe from Abashevo)
  - Proto-Indo-Iranian to early Indo-Iranian
- Tocharian (Tarim Basin) (entry route and date unknown)
  - IAMC (Inner Asian Mountain Corridor; Frachetti) entry route not impossible
  - Isolate branch of IE

Given the mobility of the IE frontier, and of pastoral cultures, no reason to equate Tocharian with Afanasievo just because both are eastern outliers.

- 2 Pre-Tocharian words in Proto-Samoyed (Janhunen 1983), but no reason to assume they were borrowed directly from an adjacent neighbor; both may have spread widely.

Conclusion: Proto-Samoyedic was probably somewhere in or near the Minusinsk Basin. Pre-Tocharian was influential, or a source of trade items, in the vicinity.

*Afanasievo: very Yamnaya-like pre-Andronovo site in Minusinsk basin. Somehow this society leapfrogged across the entire central steppe while it was still occupied by hunter-gatherers.

What the Proto-/early Indo-Iranian loans into Proto-/early Finno-Ugric tell us

I-I words vary chronologically from Proto-I-I to early Indic and early Iranian.
Differently distributed over the different Finno-Ugric branches.
None in Samoyedic.

Conclusion: I-I influence continued over some time and affected a geographically dispersed set of early Finno-Ugric dialects (branch ancestors) unevenly.

Likely interpretation: Finno-Ugric dialects extended over a sizable stretch of the northern frontier of Iranian speech.

> Iranian expanded very rapidly from NE Kazakhstan to the entire central steppe and eastern part of the western steppe. Long northern frontier for Iranian.

Holopainen c. 2018

The middle Volga and later Finno-Ugric chronology

Political map, reflecting Uralic language ranges fairly well

The middle Volga and later Finno-Ugric chronology

Pre-mid 1st century CE: Various Finno-Ugric tribes (and languages or dialects) in the area. Bulgar settlements from lower Volga begin 7th century.

Mari under Bulgar infuence by 8th century.
Powerful Bulgar state, 9th century
Mari-Bulgar intense contact, some absorption of Mari groups by
Tatar-Bashkir influence from 13th century on
Movements of indigenous people under pressure from Bulgars, Tatars, Russia
- Mari moves to right bank of Volga; later Meadow Mari emigration eastward
- Proto-Mordvin moves to current location (from probably closer to the Volga)

Conclusions:
There was more than one Finno-Ugric tribe (and language?) in the area in the early indigenous population. Chuvash (Bulgar) absorbed more than one of them.

Possibly even whole additional branches of Finno-Ugric in this accretion zone.

Johanson 2000

The Fur Road is the key to the whole model

1. Early movements on and near the steppe:
    - No overall directionality
        - Domesticate frontiers spread from W steppe, Central Asia, NW China
    - Beeline leaps to distant attractive sites
        - Informed by information about geographically distant places of interest, and travel routes
    - Movement into accretion zones, language buildups there

2. Earliest Uralic: Samoyedic eastward movement is consistent with general pre-Bronze-Age movements: bidirectional, so eastward is possible.

3. Proto-Uralic:
    - Pre-Proto-Uralic originated somewhere in or near the forest-steppe ecotone
        - (likely including the Ural mineral deposits and metal lodes)
    - Pre-Proto-Uralic enters the Fur Road economy, becomes the trade language
        - -- at some point in time and space
    - Proto-Uralic has a linear range and an expanding bidirectional frontier
        - Ancestral Samoyedic at the eastern frontier

The Fur Road as key, cont.

4. Proto-Finno-Ugric
   - Wide E-W distribution from N. Kazakhstan to western forest-steppe
     - Not point-like in space, perhaps never point-like
     - Contact with Indo-Iranian: northern Kazakhstan to NE Russia,
       - Indo-Iranian dialect diversity, over some stretch of time
     - So this contact episode was not point-like in time or space

5. Uralic branch spreads
   - Trading posts, trade colonies (or similar; not optimal terminology)
     - brought pristine, conservative early Uralic far to the east and west
   - Local spreads from trading posts to indigenous populations
     - Likely mechanism: intermarriage; both indigenous and trader spouses from powerful families. Likely language outcome: bilingualism, good command of both languages. No appreciable decomplexification of Uralic.

6. Expected genetic profile
   - Uralic signature strongest where Pre-Proto-Uralic entered the Fur Road
   - Strong (but not unmixed) all along the Fur Road (maybe now pushed north)
   - Progressively weaker to the north, and locally more mixed.

The Fur Road as key, cont.

7. Expected archaeological signature
   - Mostly invisible. Trading posts and canoes maybe visible.
   - Raw materials move N to S (and S to N?); value-added trade items E-W
   - Technological innovations spread rapidly E-W along Fur Road; then N

8. Falsifying and supporting this picture
   - Linguistics:
     - Decomplexification (if any) to the north and along smaller rivers
       - – but really no reason to assume rapid shift and decomplexification
     - Local contact influence especially to north
     - Conservative Proto-Uralic profile far to E and W (and in general)
   - Archaeology:
     - Hypotheses about directionality of technological innovations, raw material flow, trade item flow, cultural contacts
     - Evidence of wealth (important driver of trade in luxuries, e.g. furs)
   - Genetics: Detectable Uralic signature with a long E-W trajectory, shorter and steeper S > N trajectory, and maybe a more concentrated center
     (Lactose tolerance part of the early signal??)

Thanks!

NSF 9222294
Kone Foundation
Russian Academic Excellence Project 5-100

johanna@berkeley.edu

References

Barfield, Thomas. 2009. Introduction: Frontiers and border dynamics. Bryan K. Hanks and Katheryn M. Linduff, ed., *Social Complexity in Prehistoric Eurasia: Monuments, Metals, and Mobility*, 235-240. Cambridge: Cambridge University Press.

Fortescue, Michael. 1998. *Language Relations across Bering Strait*. London and New York: Cassell.

Grünthal, Riho and Johanna Nichols. In press. Transitivizing/detransitivizing typology and language family history. *Lingua Posnaniensis* 58:2:

Ives, John W. 2003. Alberta, Athapaskans, and Apachean origins. Jack W. Brink and John F. Dormaar, ed., *Archaeology in Alberta: A view from the new millennium*, Medicine Hat, Alberta: Archaeological Society of Alberta.

Ives, John W. 2007. Review of Athapaskan migrations. R. G. Matson and Martin P. R. Magne, ed., *The Archaeology of Eagle Lake, British Columbia*, Tucson: University of Arizona Press.

Janhunen, Juha. 1983. On early Indo-European-Samoyed contacts. ed., *Symposium saeculare Societatis fenno-ugricae*, 115-128. Helsinki: Finno-Ugric Society.

Johanson, Lars. 2000. Linguistic convergence in the Volga area. Dicky G. Gilbers, John Nerbonne and Jos Schaeken, ed., *Languages in Contact*, 165-178. Amsterdam: Rodopi.

Nedjalkov, V. P. 1969. Nekotorye verojatnostnye universalii v glagol'nom slovoobrazovanii. F. Vardul', ed., *Jazykovye universalii i lingvisticheskaja tipologija*, 106-114. Moscow: Nauka.

Nichols, Johanna. 2007. A typological geography for Indo-European. Karlene Jones-Bley, Martin E. Huld, Angela Della Volpe and Miriam Robbins Dexter, ed., *Proceedings of the 18th Annual UCLA Indo-European Conference*, 191-211. Washington, DC: Institute for the Study of Man.

Nichols, Johanna. In press. Person as an inflectional category. *Linguistiic Typology* 21:3:

Nichols, Johanna, David A. Peterson and Jonathan Barnes. 2004. Transitivizing and detransitivizing languages. *Linguistic Typology* 8:2: 149-211.

Nichols, Johanna and Richard A. Rhodes. In press. Vectors of language spread at the central steppe periphery: Finno-Ugric as catalyst language. *British Archaeological Reports (International Series)*