

The background of the slide is an aerial photograph of a vast mountain range. The peaks are covered in snow and are partially obscured by a thick layer of white clouds that fills the lower half of the image. The sky above is a clear, pale blue. The text is overlaid on this background.

# **What favors the development of rara?**

## ***A Himalayan case study***

**Balthasar Bickel**

University of Leipzig

# Data and methods

- AUTOTYP project with Johanna Nichols (UC Berkeley) on the historical development of typological distributions:  
[www.uni-leipzig.de/~autotyp](http://www.uni-leipzig.de/~autotyp)
- CPDP fieldwork project on Kiranti languages in Nepal:  
[www.uni-leipzig.de/~ff/cpdp](http://www.uni-leipzig.de/~ff/cpdp)

## Relative vs absolute rara

- Relative rara = rare wrt surrounding area, i.e. enclave effects (Bickel & Nichols 2003)
- Absolute rara = rare wrt to the (currently attested) world

## Relative rara in the Eurasian Enclaves

- Bickel & Nichols's (2003) Eurasian Enclave Theory: the Caucasus and the Himalayas were only marginally affected by the postneolithic language spreads in Eurasia (Northern Steppe, Southeast Asia, Mesopotamia and South Asia)
- Prediction: a substantial number of typological variables evidence a difference between the enclaves and the rest of Eurasia.





# Relative rara in the Eurasian Enclaves

- Draw genealogically-balanced samples from AUTOTYP and WALS (major branch/genus-based)
- 112 Variables (several overlapping) with sufficient or near-sufficient datasets
- Accept effects with  $p < .05$  in a distribution-free permutation test (exact in 2x2 tables, randomization-based elsewhere, following Janssen, Bickel & Zúñiga 2005)
- Reject effects which are secondarily induced by some local areal skewing in the rest of Eurasia:
  - case (absence in SEA)
  - position of DEM (postposed in SEA)
  - order of S,V, and O (non-final arealities in SEA and EUROPE)
- Reject effects which are universally correlated (e.g. accept at most one of OV/VO or GenN-NGen effects)

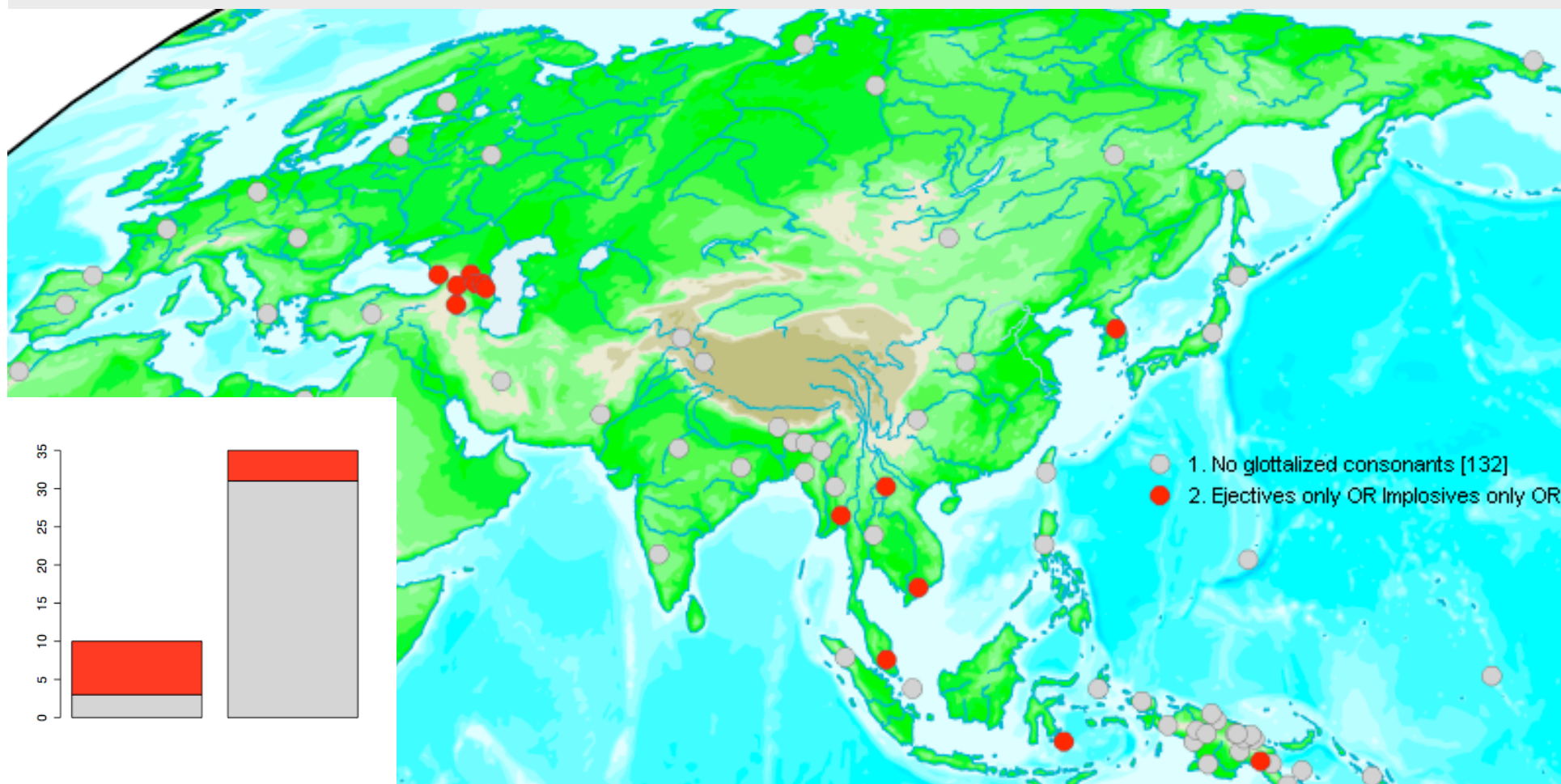
# Relative rara in the Eurasian Enclaves

MADGLO0	Glottalized C	WALSG	Maddieson 2005
MADUVU2	Uvular C (reduced to binary)	WALSG	Maddieson 2005
MADUVU0	Uvular C	WALSG	Maddieson 2005
MADVOI0	Voicing series ('none' removed)	WALSG	Maddieson 2005
SYN	Verbal Inflectional Synthesis	GEN	Bickel & Nichols 2005
POLYAGR	Obligatory agreement with A and P	GEN	Bickel & Nichols, nyp
PREROLE	Some agreement prefixed	GEN	Bickel & Nichols, nyp
DOBOPT	Inflectional Optatives	WALSG	Dobrushina et al. 2005
COMNUM5	Counting systems (subtypes collapsed)	WALSG	Comrie 2005
POSSCL	Inflectional possessive classes	GEN	Nichols & Bickel 2005
DRYPOS0	POSS Pf vs. Sf vs. both ('none' removed)	WALSG	Dryer 2005
CORNUM	N (genders)	WALSG	Corbett 2005
COMALN5	NP alignment (ACC subtypes collapsed)	WALSG	Comrie 2005
COMALP0	PRO alignment ('no PRO' rm, ACC collapsed)	WALSG	Comrie 2005

N = 14

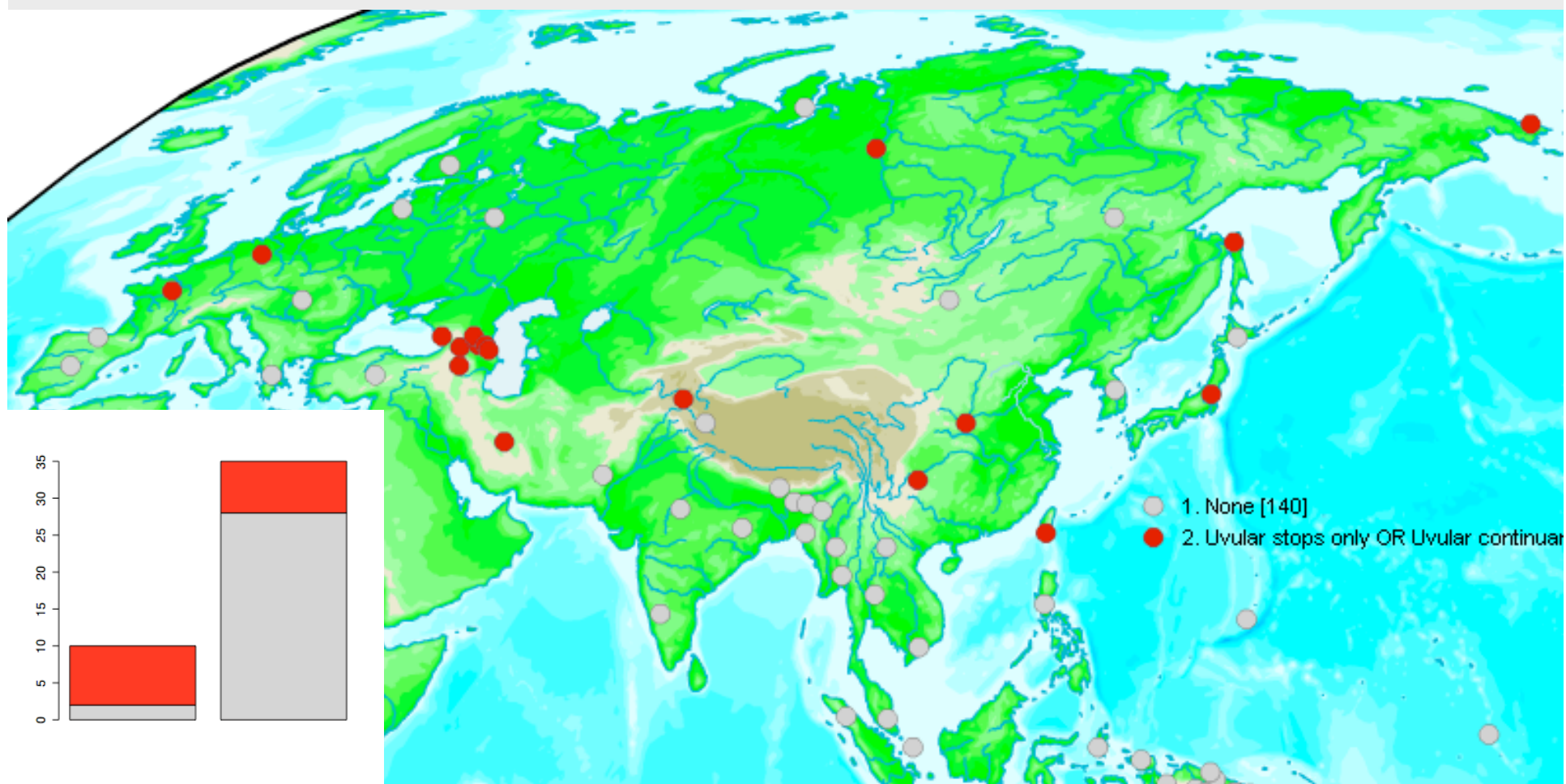
# Relative rara in the Eurasian Enclaves

Maddieson 2005: Glottalized C, WALSG = 45



# Relative rara in the Eurasian Enclaves

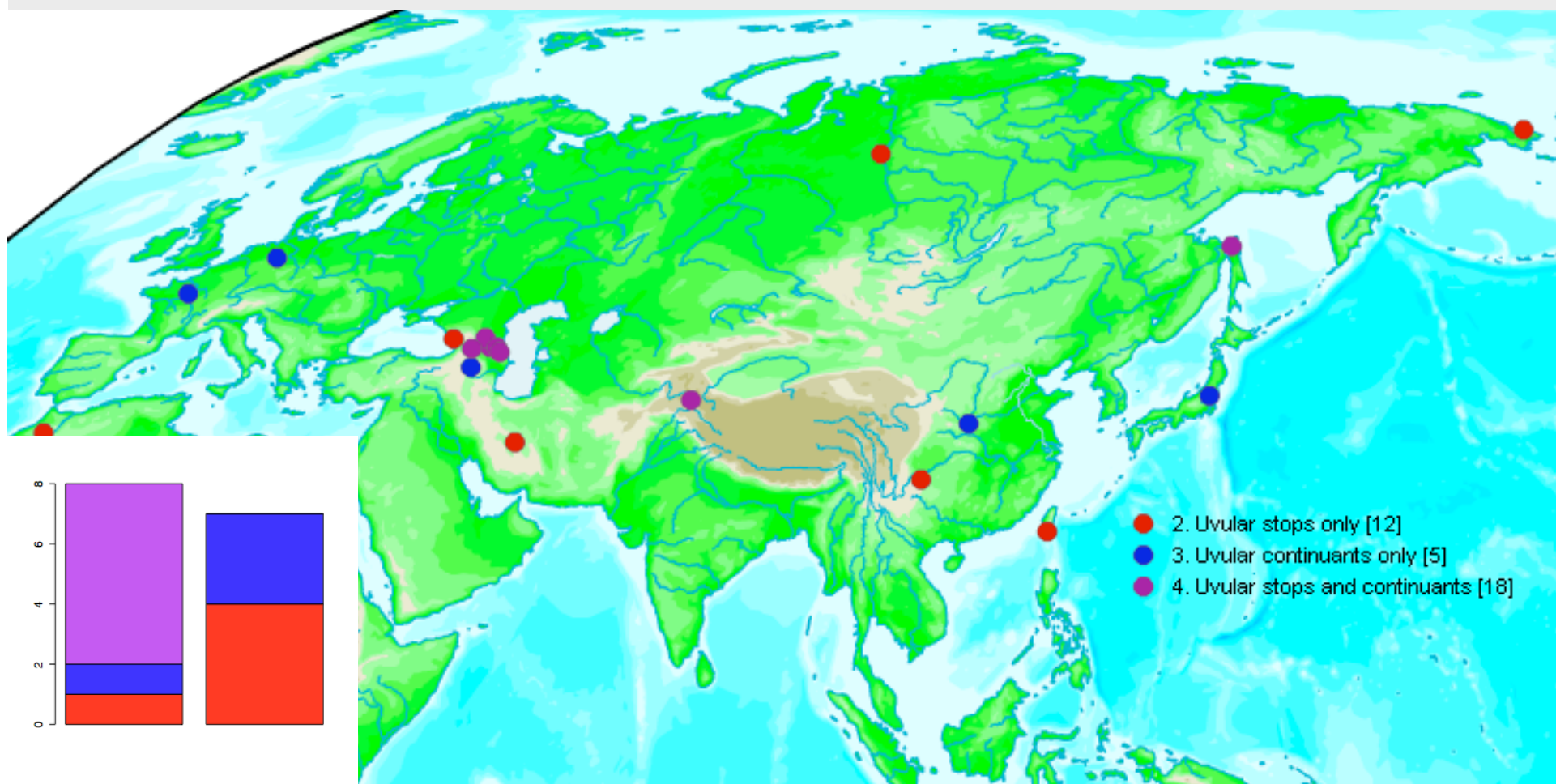
Maddieson 2005: Uvular C, WALSG = 45





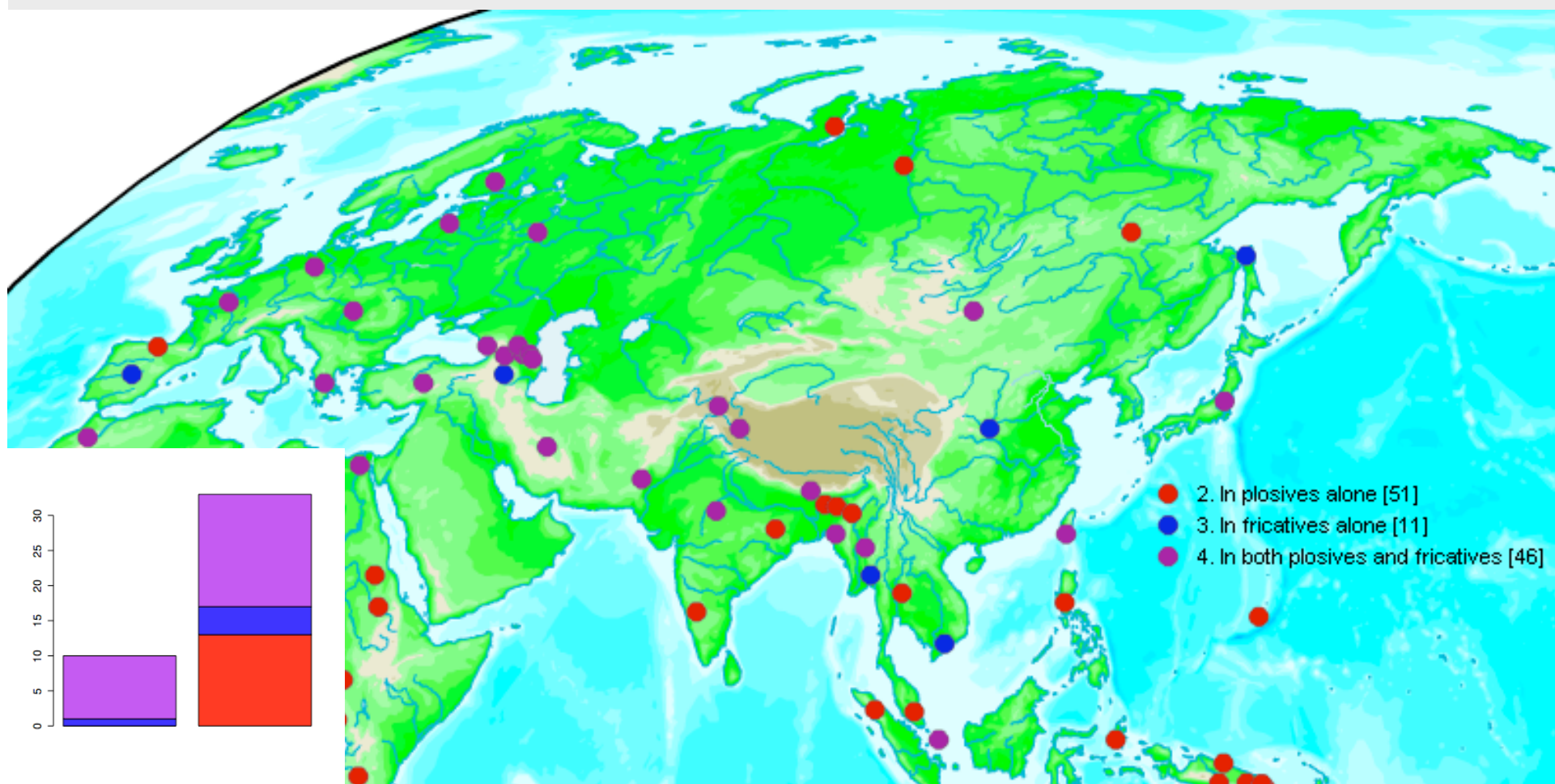
# Relative rara in the Eurasian Enclaves

Maddieson 2005: Uvular Series, WALSG = 15



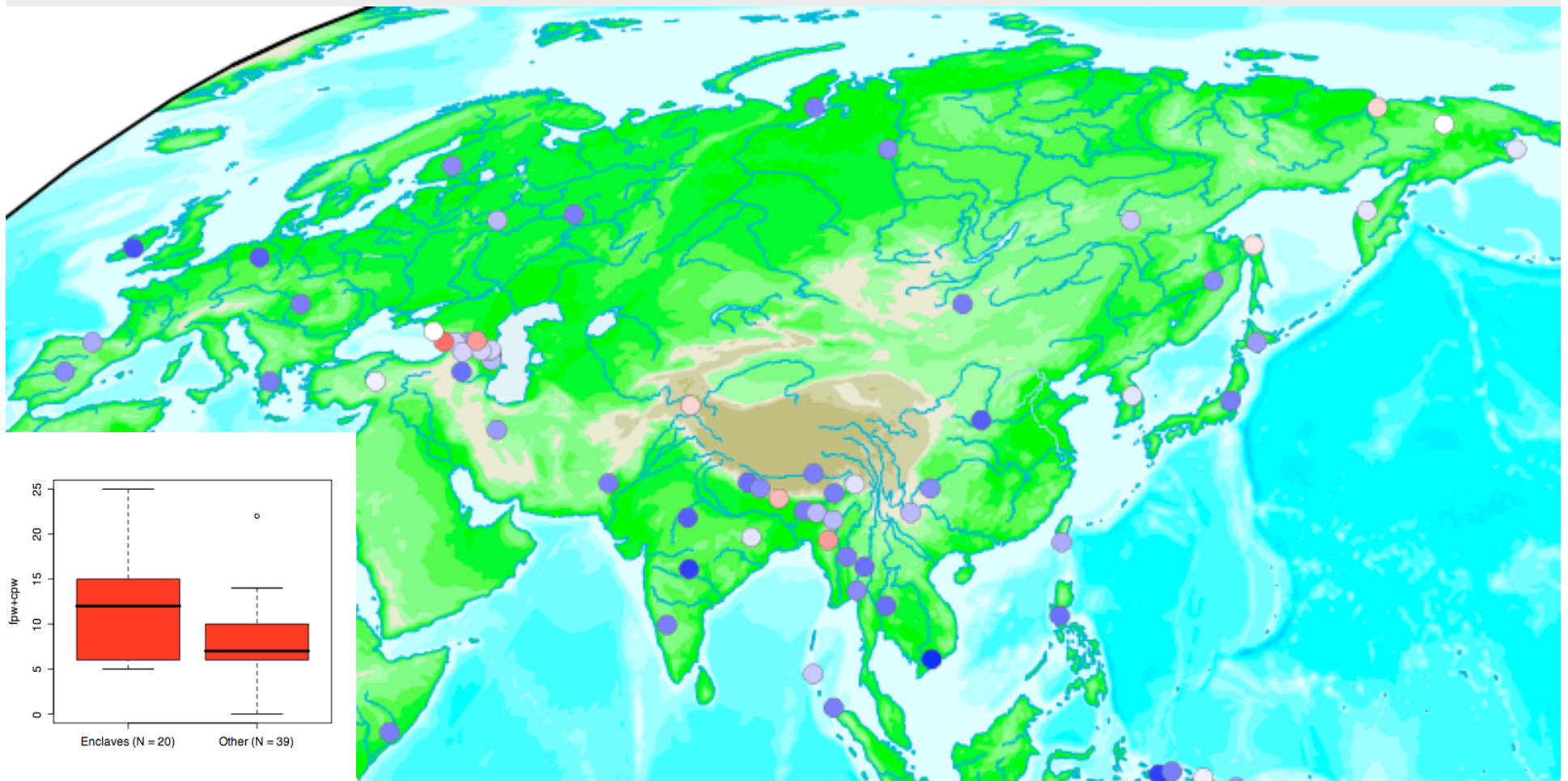
# Relative rara in the Eurasian Enclaves

Maddieson 2005: Voicing Series, WALSG = 43



# Relative rara in the Eurasian Enclaves

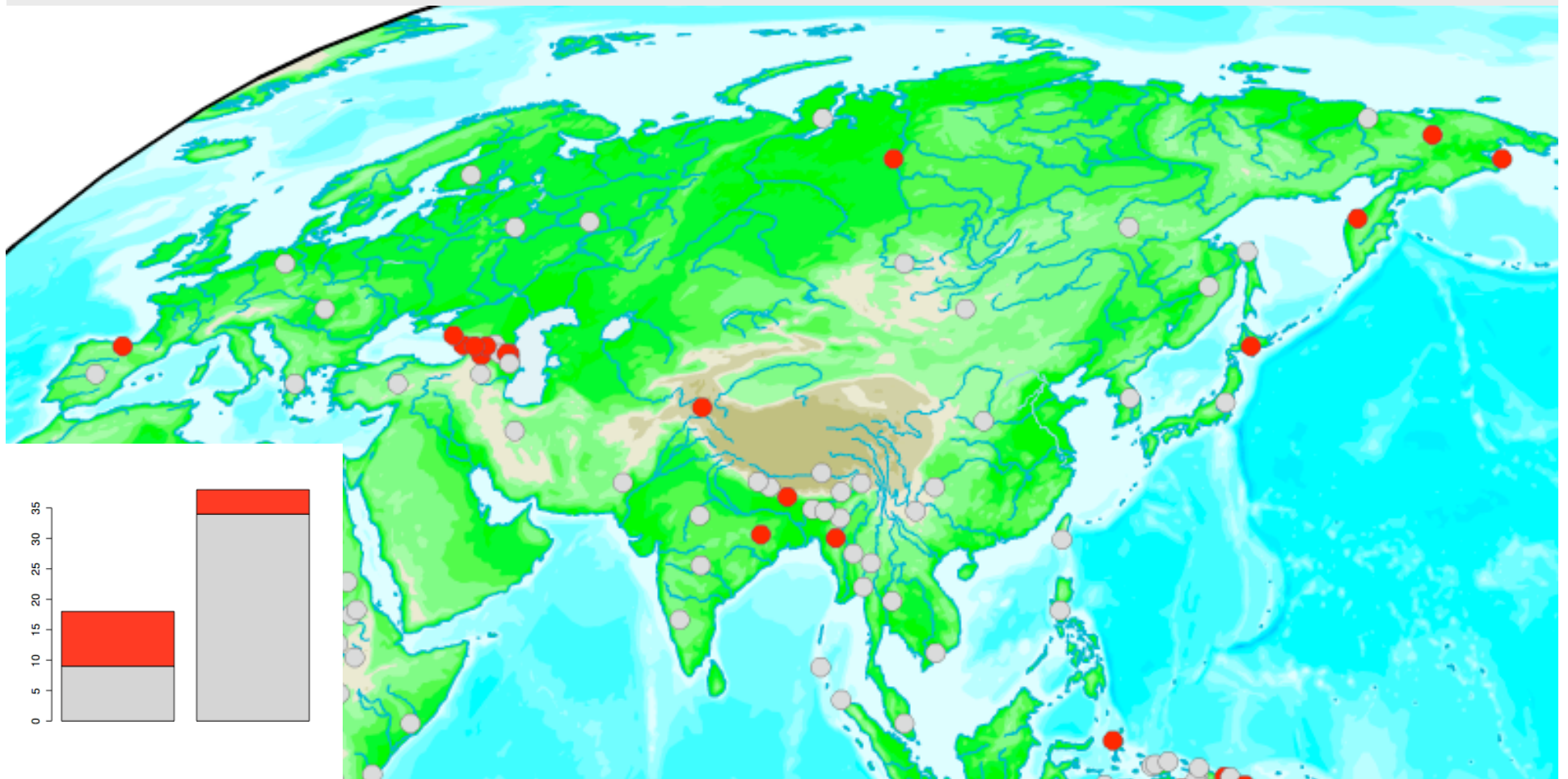
Bickel & Nichols 2005: Verbal Inflectional Synthesis, GEN = 55





# Relative rara in the Eurasian Enclaves

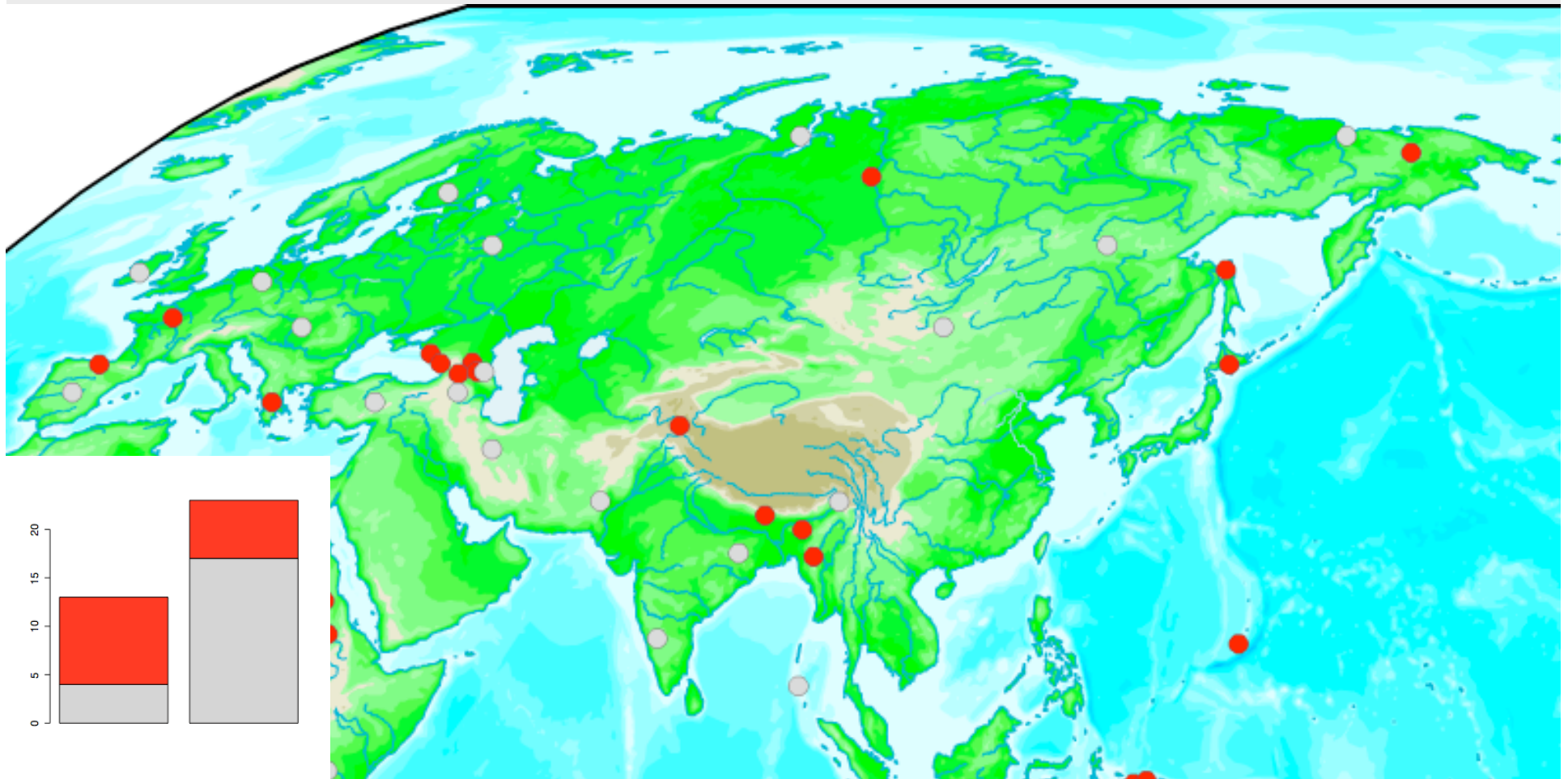
Bickel & Nichols NYP: Obligatory Polyagreement, GEN = 56





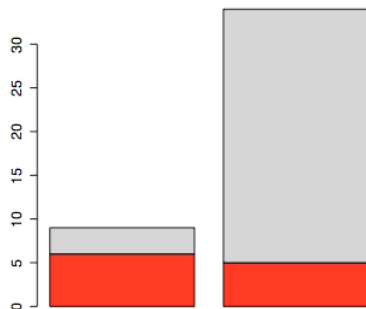
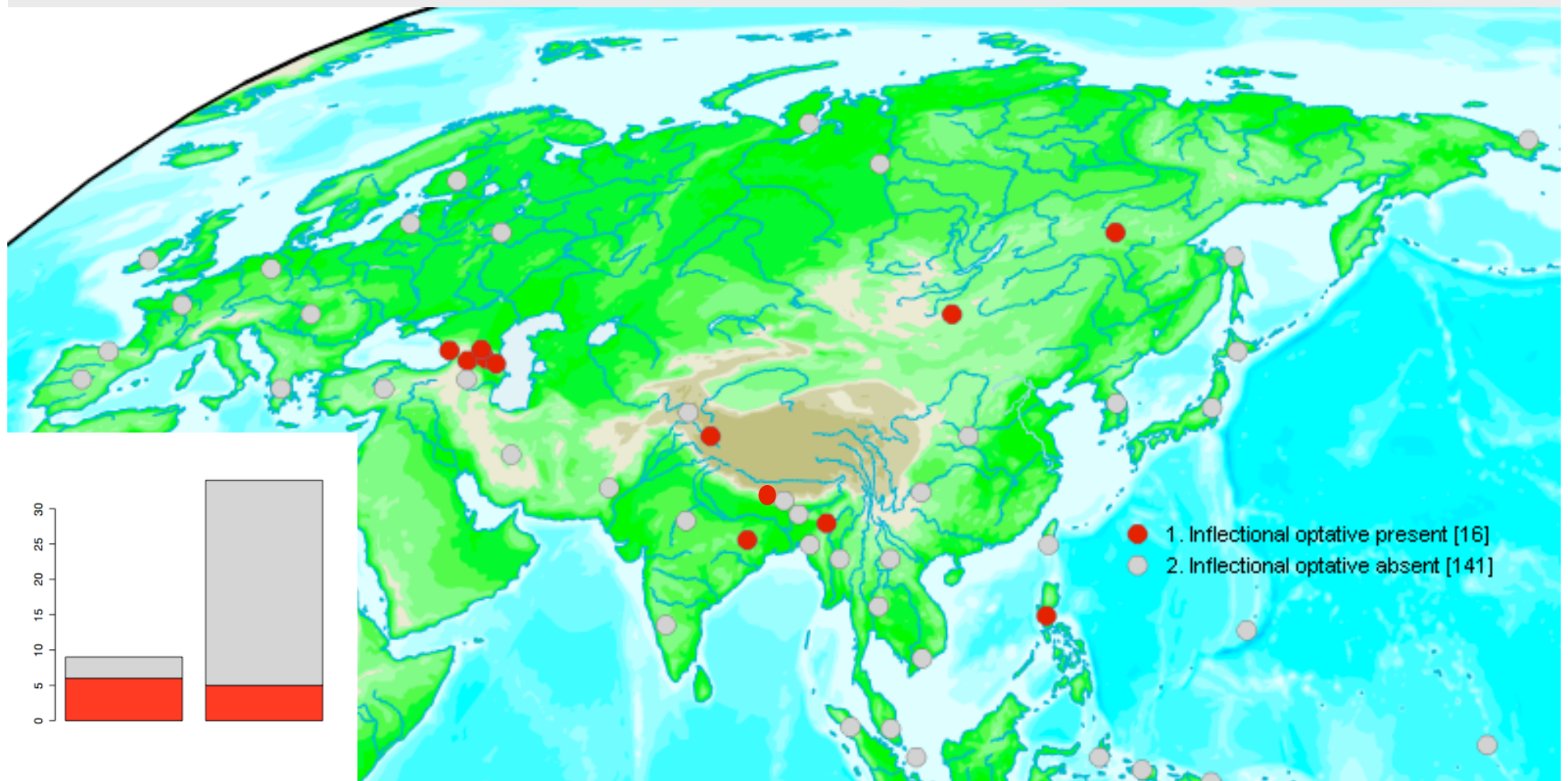
# Relative rara in the Eurasian Enclaves

Bickel & Nichols NYP: Prefixal AGR, GEN = 36



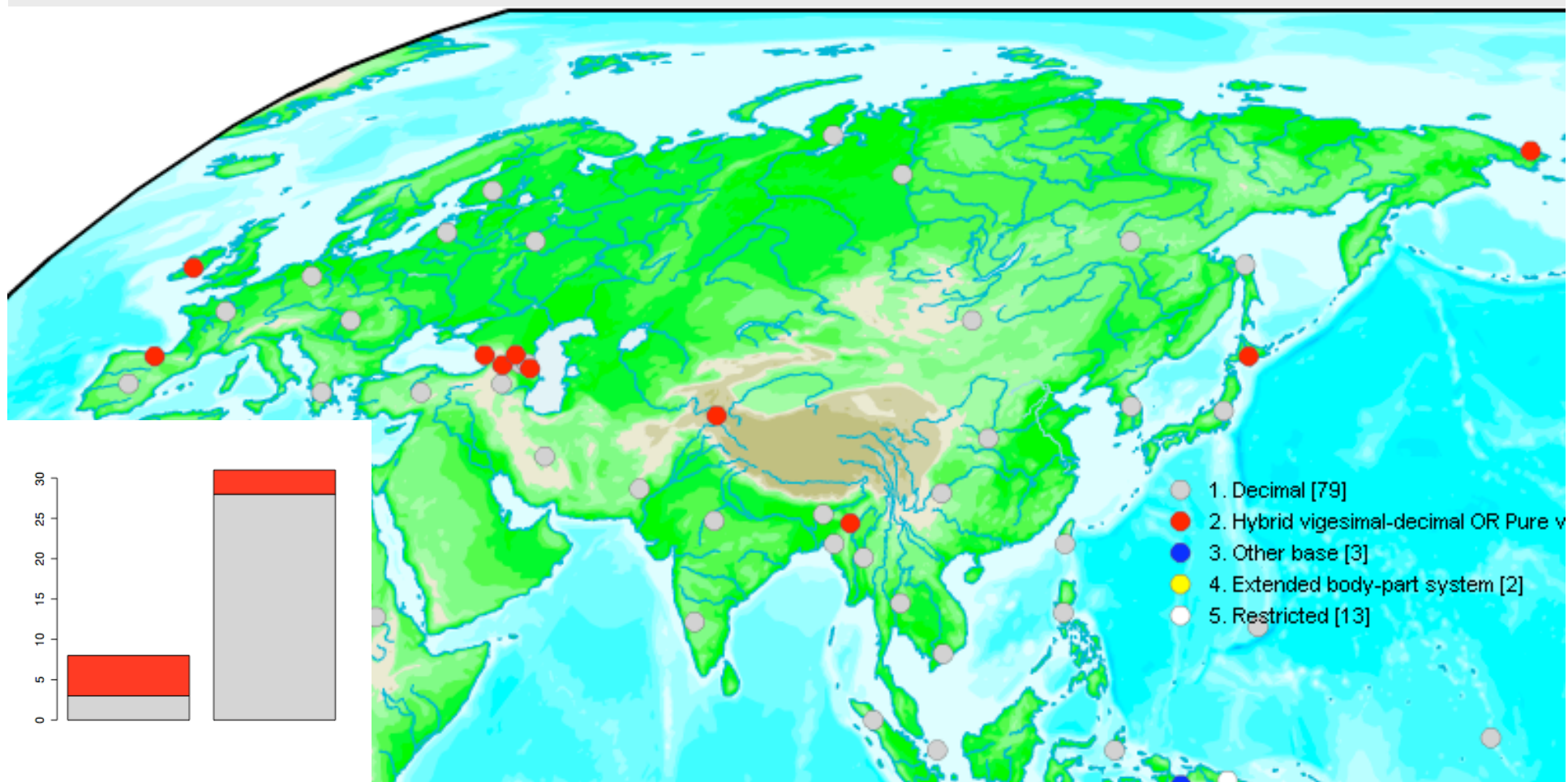
# Relative rara in the Eurasian Enclaves

Dobrushina et al. 2005: Inflectional Optative, WALSG = 43



# Relative rara in the Eurasian Enclaves

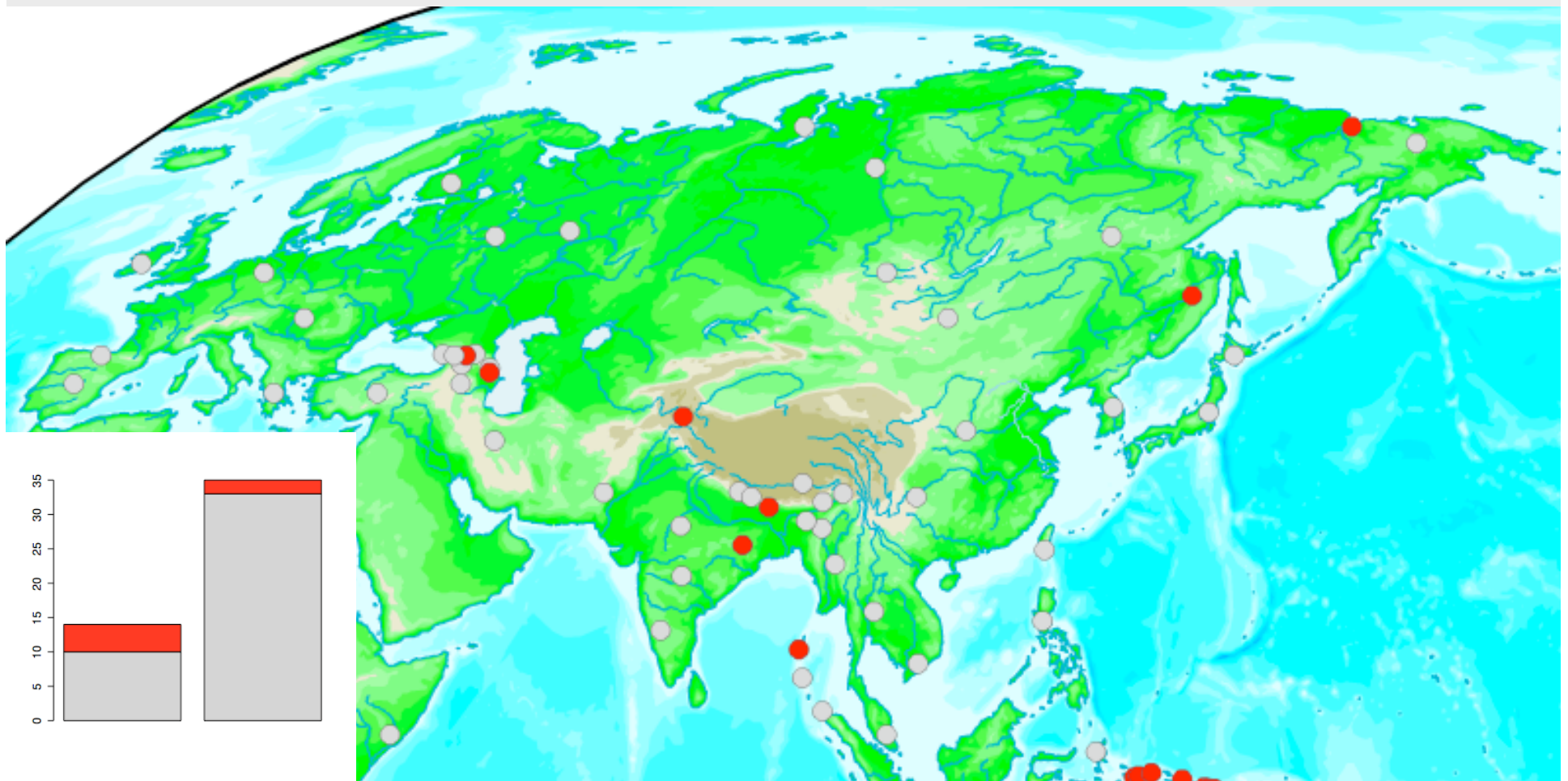
Comrie 2005: Counting Systems, WALSG = 39





# Relative rara in the Eurasian Enclaves

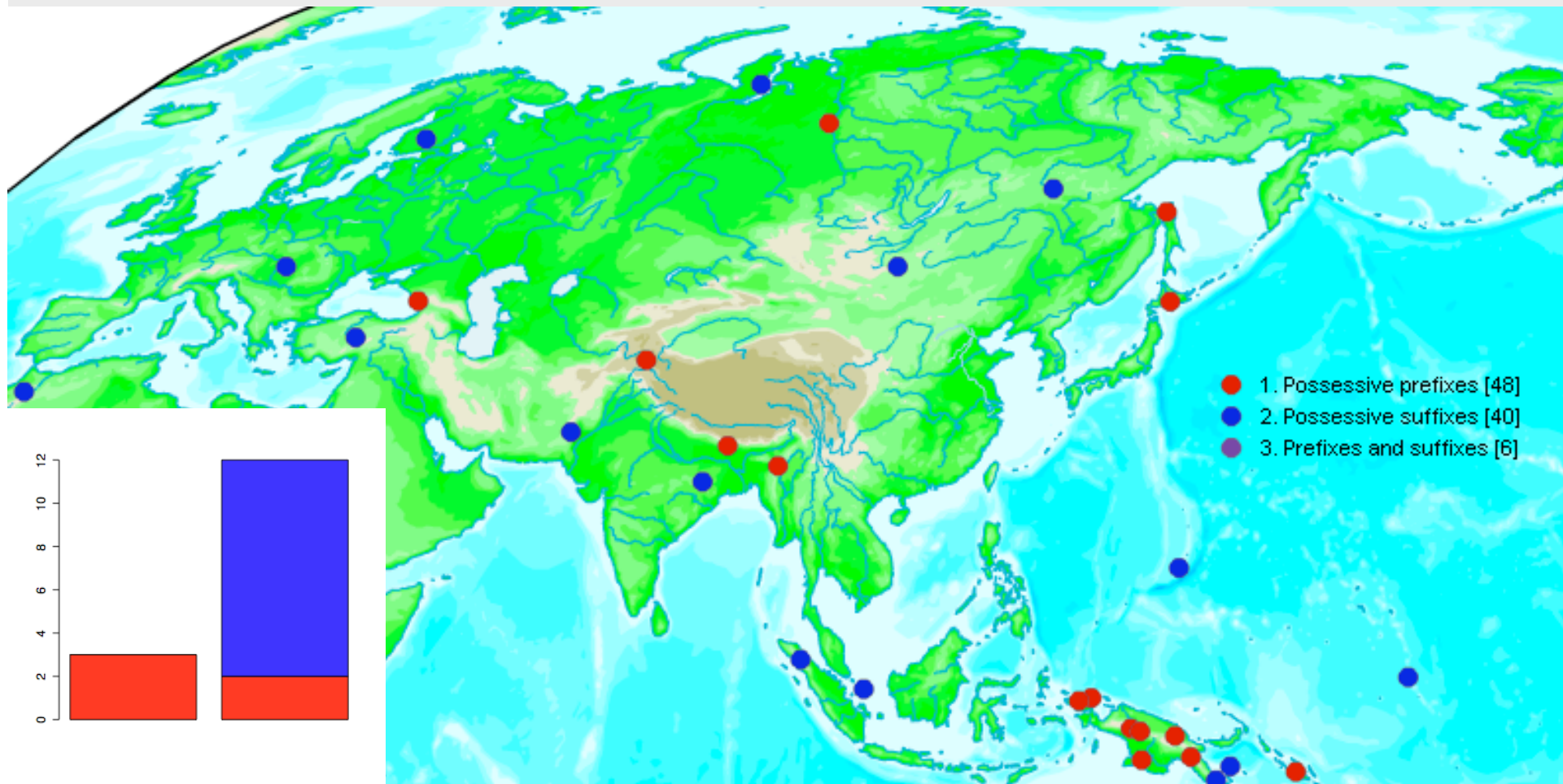
Nichols & Bickel 2005: POSS classes, GEN = 49





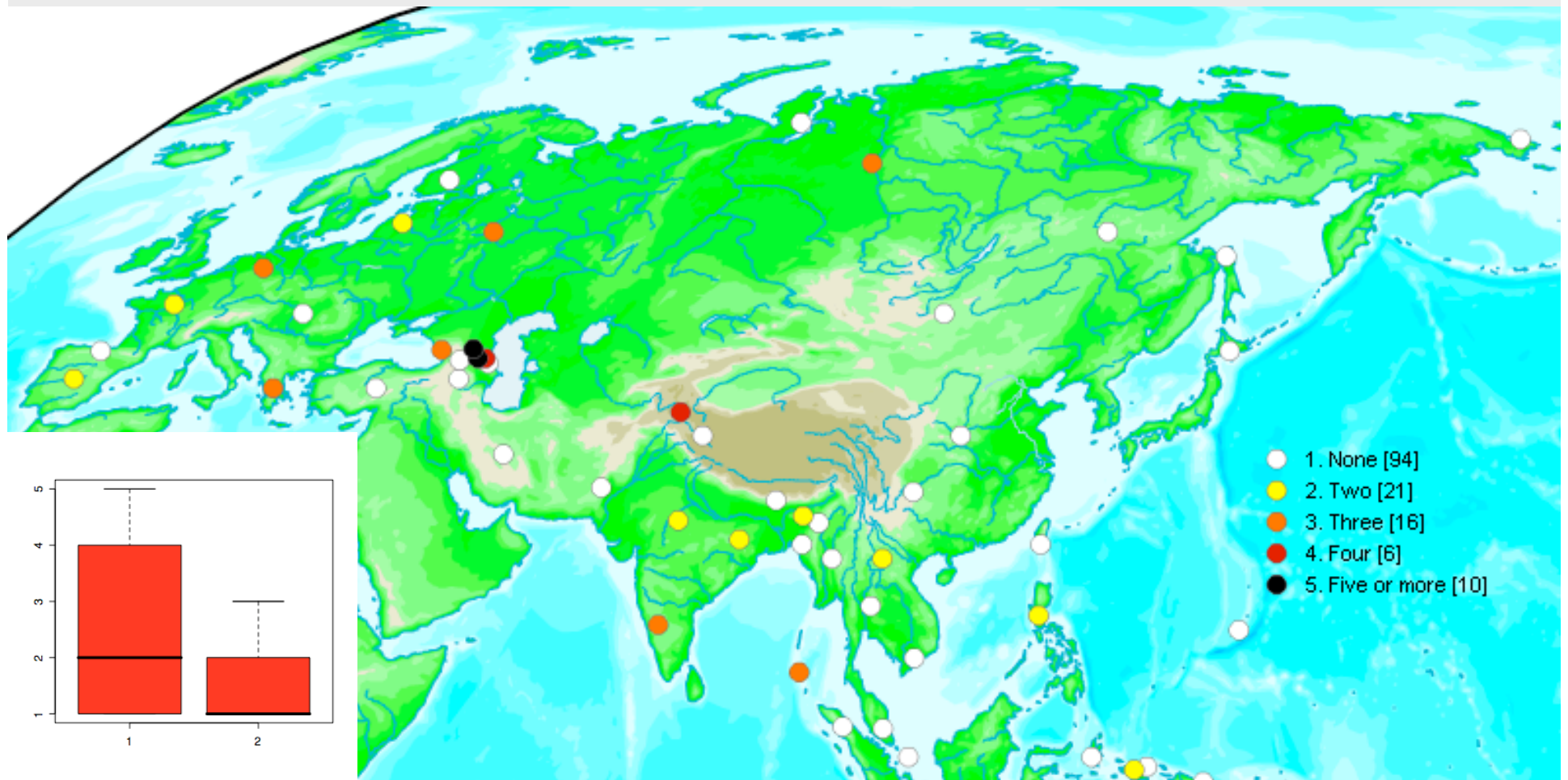
# Relative rara in the Eurasian Enclaves

Dryer 2005: Position of POSS affixes, WALSG = 15



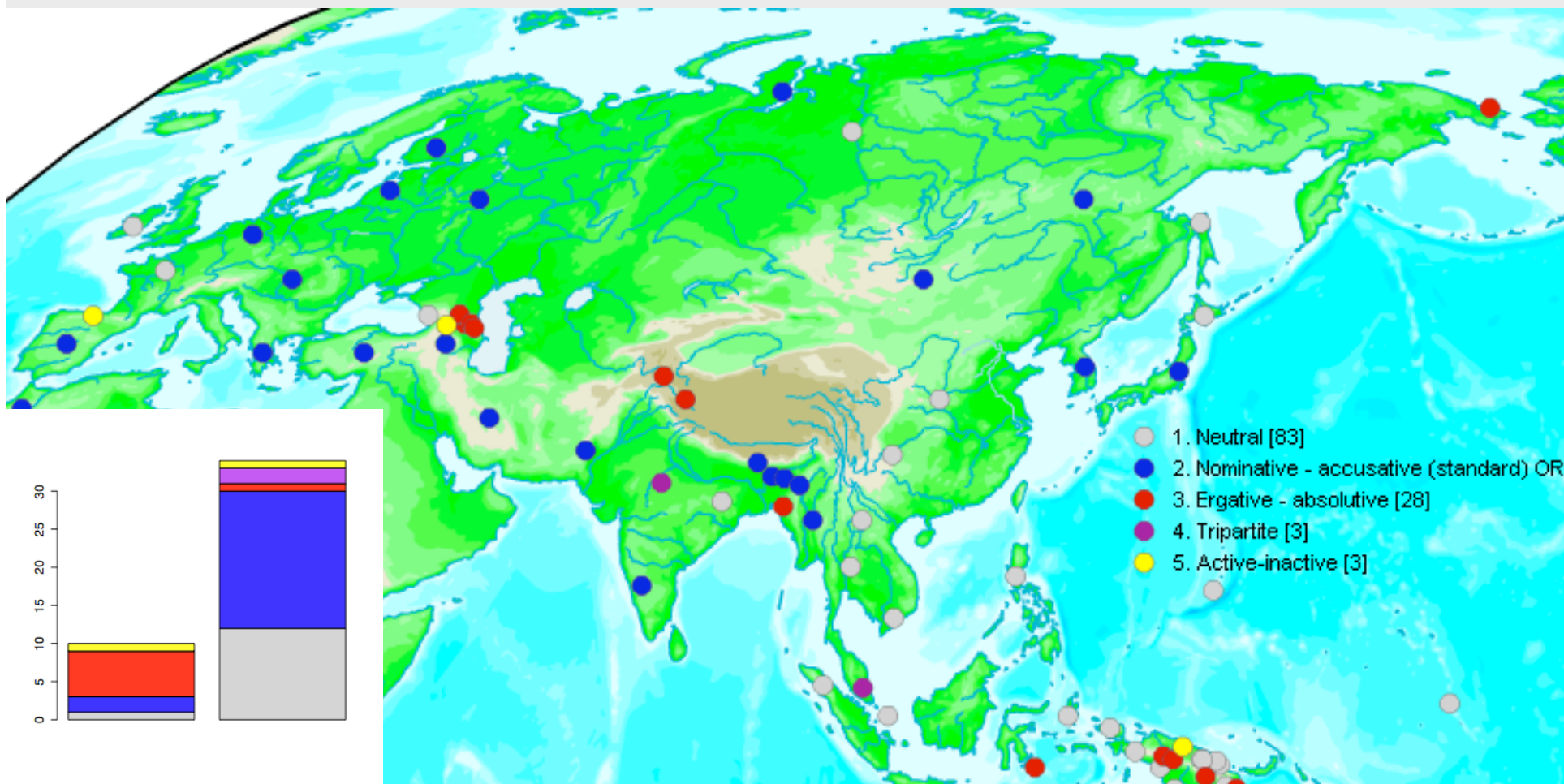
# Relative rara in the Eurasian Enclaves

Corbett 2005:  $N$  (genders), WALSG = 44



# Relative rara in the Eurasian Enclaves

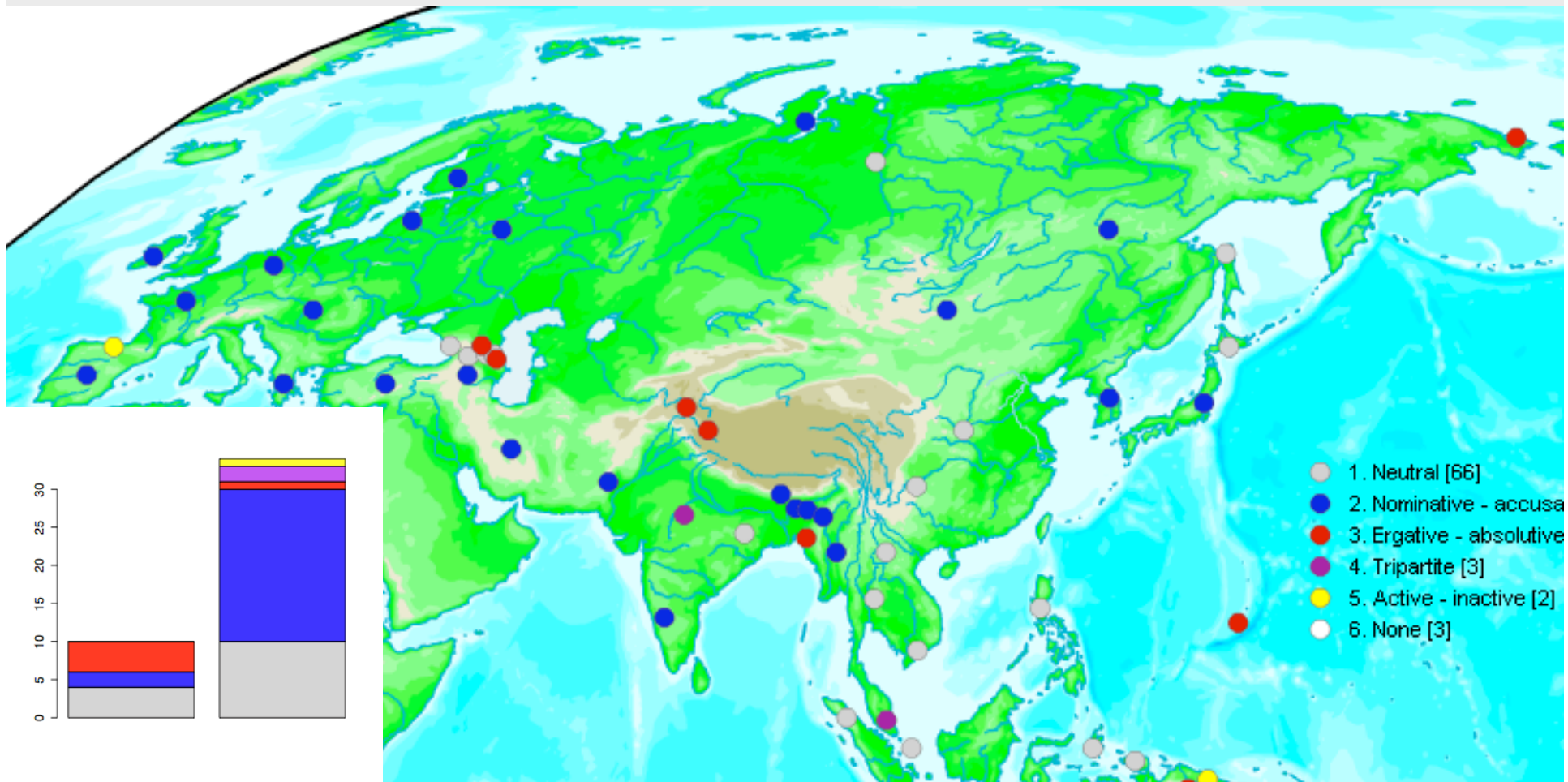
Comrie 2005: NP alignment, WALSG = 46





# Relative rara in the Eurasian Enclaves

Comrie 2005: PRO alignment, WALSG = 46

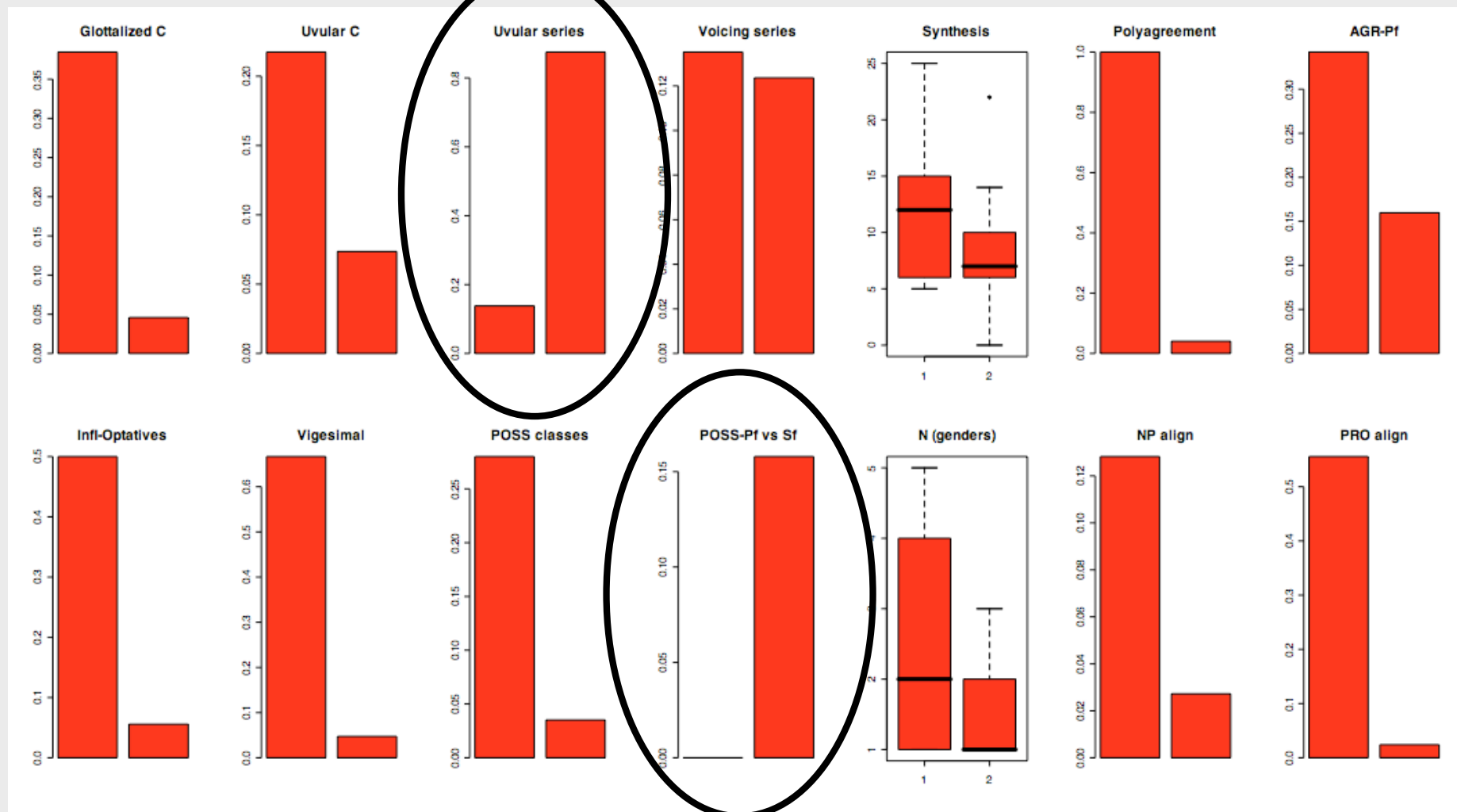




# Relative rara in the Eurasian Enclaves

- Bickel & Nichols 2003: Enclaves are not areas! They are the 'left-over' of areas.
- Therefore, expect greater variance within than outside enclaves!
- Test this by comparing variances
- Measure variance for each region by
  - standard deviations for scalar variables
  - chisq-deviations from equal distributions (50:50, 33:33:33, 25:25:25:25, etc.) for categorical variables (plotted as the inverse of this)

# Relative rara in the Eurasian Enclaves



Larger variance of multinomial variables = smaller chisq-deviations from 50:50, 33:33:33, 25:25:25:25 etc. expectations.

# Relative rara in the Eurasian Enclaves

Other hypothesized effects (not tested yet):

- Complex or at least bipartite stems (Bickel & Nichols 2003), leading to endocclisis (Harris 2002 on Udi in the Caucasus; Bickel et al. 2005 on Chintang in the Himalayas), exuberant agreement (Harris 2006) and the like.
- Conjunct/Disjunct agreement (Hale 1980, Hargreaves 2003, Bickel 2000, Curnow 2002)
- etc.



# Absolute or near-absolute rara in the Himalayas

1. Upside-down split ergativity: ergative alignment for first person, but not further down the referential hierarchy (Bickel 2000): Puma, Athpare, Limbu, Hayu, Khaling, Bahing, Bantawa
2. Syntactic ergativity in complementation (Bickel & Nichols 2001, Bickel 2004): Belhare
3. Antipassives (instead of passives) used for first person patient reference (Bickel & Gaenszle 2005): Puma
4. Free prefix ordering (Bickel et al. 2005): Chintang, Bantawa
5. Recursive inflection (Bickel et al. 2005): Dumi, Chintang, Athpare, Puma
6. Triplication that is independent of reduplication (Rai & Winter 1997, Rai et al. 2005): Bantawa, Chintang
7. Reflexives as verb stems (Rutgers 1998, Bickel 2003): Puma, Belhare, Yamphu
8. Middle voice < \*eat (Ebert 1994, Bickel 2003): Athpare, Bantawa, Belhare
9. Spatial cases ('up at tree', 'down at tree', etc. ) (Rai 1988, Bickel 1997): most Kiranti languages
10. Pronominal root distinctions for deictic transposition (Bickel 2001): Chintang, Belhare
11. Spatial interjections: Chintang, Belhare
12. Color-sensitive articles (Bickel 2003): Belhare
13. Aspirated stops alternating with breathy voiced stops ( $ph \rightarrow bhV\_V$ ) : Limbu, Belhare  
? Voiceless-voiced clusters and pseudo-geminates: Puma, Belhare

## Upside-down ergativity splits

Puma (S. Kiranti) and many other Kiranti languages (Bickel 2000)

	A	S	P
1s	-ŋ(>3) -na (>2)	-ŋa (~ -oŋ /PST)	
1d	-ci-∅		
1p	-m	-i	
2	tʌ-		
3s	∅- pʌ- (>1)	-i	
3d	pʌ- -ci ni- (>2)		
3p	mʌ-(>3n s) ni-pʌ- (>1) ni- -nin (>2, 3)	-ci	

*ben-ŋa* 'I come over'  
*pʌ-en-ŋa* 'S/he heard me'  
*enn-u-ŋ* 'I hear him/her'

*ben* 's/he come over'  
*enn-i* 's/he hears him/her'

# Syntactic ergativity in complementation

Belhare (E. Kiranti) (Bickel 2004)

a.  $\emptyset$  *khon-ma nui-ka.*

[S] play-INF may-2sNPST

‘You may play.’

b.  $\emptyset$   $\emptyset$  *lu-ma nui-ka.*

[A] [P] tell-INF may-2sNPST

‘You may be told.’ = ‘I/S/he may tell you.’

*Impossible:* ‘You may tell him/her.’

# ANTIPASSIVE and other generics for 1PO

Puma (S. Kiranti) (Bickel & Gaenszle 2005)

*kha-en-a.*

ANTIPASSIVE-hear-PST[3sS]

- a. 'S/he heard someone / people.' or 'S/he listened so as to find out whether or not there are people.' (does not entail existence of a specific undergoer referent)
- b. 'S/he heard us (incl).'



# Free prefix ordering

Chintang (E. Kiranti; Bickel et al. 2005)

*ma-* ~ *mai-* 'NEG'

*ma-* 'eP

*mai-* 'iP'

*kha-* '1nsP'

*a-* '2S/A'

*u-* '3A' (if P = 1s) or '3nsS/A' (elsewhere)

*na-* '3>2'

- selectionally restricted to verb stems
- resulting in syntactically integer words ( $X^0$ )
- no gapping under identity allowed (unlike clitics)
- enter dependencies with suffixes: *a-mai-kha-tup-t-a-ce*  
2-NEG-1nsP-meet-NEG-PST-d  
'You (s/d/p) didn't meet us (de).'

## Free prefix ordering

- Subcategorize prosodically for a p-word -- but ANY p-word in V
- P-word in Chintang:
  - optional ?-epenthesis at the left edge
  - only possible endoclititic host

*{u-kha-ma}-cop-yokt-e*

3nsA-1nsP-NEG-see-NEG-PST

‘They didn’t see us (pe).’

e.g. (kha)(?u)(ma)(cop)(yokte)

or (?u)(ma)(kha)(cop)(yokte)

or (kha)(ma)(cop)(?u)(yokte)

# Recursive inflection

Chintang (Bickel et al. 2005)

second stems (mostly grammaticalized) require a one-foot host:

$[\Sigma' (\phi [\Sigma \textit{met}]-na)-bi]-na-?ã-ni$

do-1>2-V2:BENEFACTIVE-1>2-eNPST-p

‘I’ll do it for you (p).’

$[\Sigma' (\phi [\Sigma \textit{ko}]-na)-gon]=lok...$

walk-NA-V2:AMBULATIVE=SIM

‘when s/he walks around...’ (no suffixes available in this mood)

$a-mas-u)-and-u)-bid-u-ku-m=ni$

2-loose-3P-V2:TELIC-3P-V2:BENEF-3P-NPST-2pA=FOC

‘You’ll lose it on him!’

## Reflexive verb stems

Puma

$[\Sigma' (\phi [\Sigma \text{dher}]-o\eta)-cen]-o\eta$   
hit-1sS/P.PST-REFL-1sS/P.PST

'I hit myself.'

Belhare: all verb stems CV ~ CVV / some desinences, e.g. *-yu* 'NPST':

*tenma* 'to hit': *teĩ-yu*  
*tenchinma* 'to hit oneself': *ten-chiĩ-yu*

also: Yamphu (Rutgers 1998)



## Middle EAT

Athpare, Bantawa, Belhare: verb 'eat' grammaticalized as Middle Voice (Ebert 1994, Bickel 2003):

*khon-ca-he*

play-MIDDLE- PST[3sS]

'She played by herself.'

# TriPLICATION not via reduplication

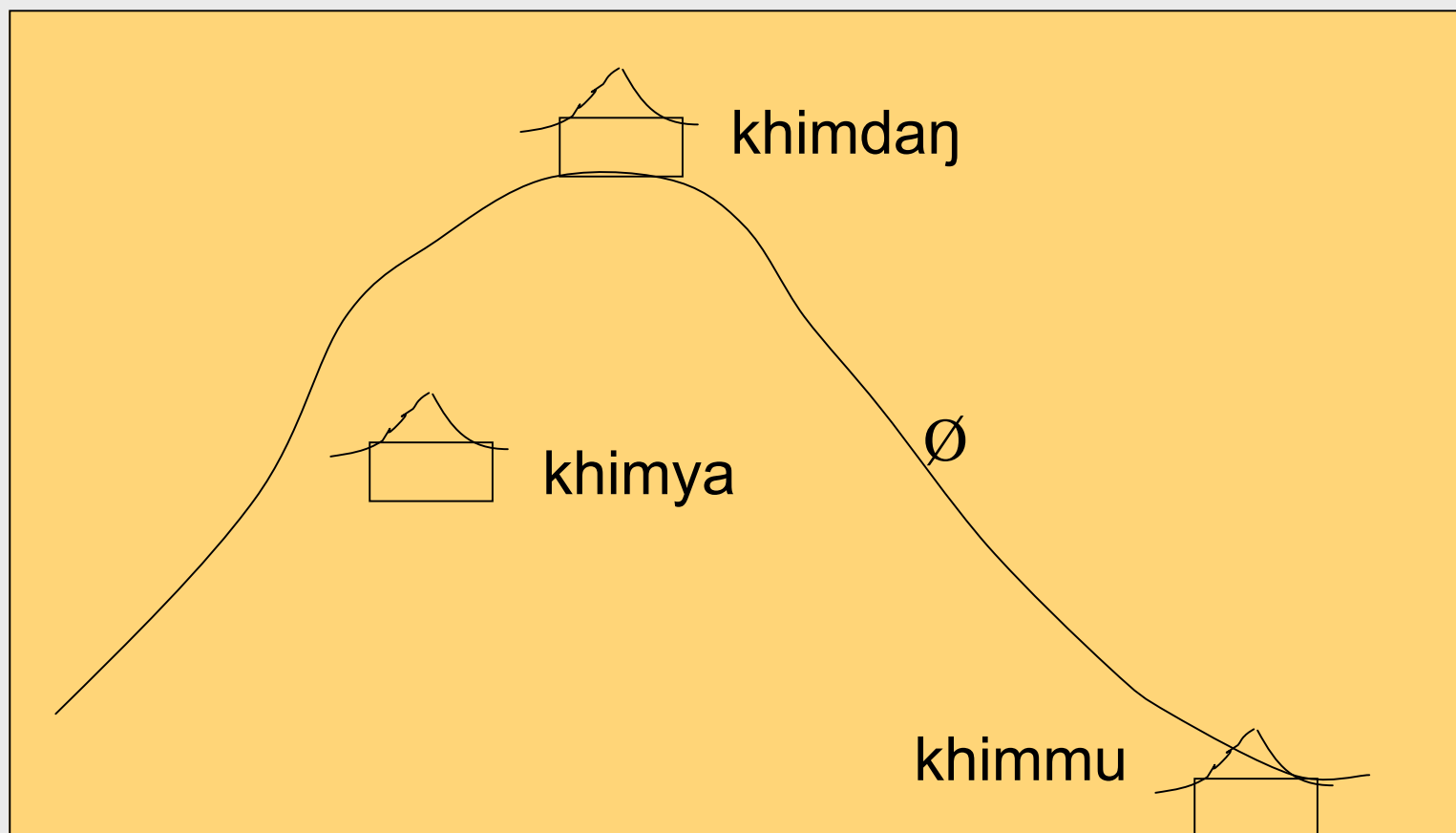
Chintang (Rai et al. 2005)

*rak-rak-rak-wa*            *mi om-no.*  
(burning red coal)-ADV fire burn-NPST  
'The fire burns very strongly.'

\**rak-rak(-wa)*

# Spatial cases and interjections

Belhare (Bickel 1997)



# Root-coded deictic transposition

Chintang (E. Kiranti): 3x5 demonstratives,

each in 5 locative cases,  $N = 75$  forms,

plus distance-iconic lengthening for all but the  $F=S=\emptyset$  ('here') series

	PROXIMAL	DISTAL	$F=S=\emptyset$	$\emptyset \neq S$	$R \neq S$
UP	<i>toba</i>	<i>atu(ba)</i>	<i>bandu</i>	<i>tobandu</i>	<i>utu(ba)</i>
DOWN	<i>moba</i>	<i>amu(ba)</i>	<i>bamu</i>	<i>mobamu</i>	<i>umu(ba)</i>
ACROSS	<i>yoba</i>	<i>ayu(ba)</i>	<i>bayu</i>	<i>yobayu</i>	<i>uyu(ba)</i>



# Spatial interjections

Belhare:

*tu! vs. mu! vs. yu!*

Chintang:

- *to, toto, toi, togoi*
- *mo, momo, moi, mogoi*
- *yo, yoyo, yoi, yogoi*

# Color-sensitive articles

Belhare (E. Kiranti, Bickel 2003):

<i>phabeleŋ=ma</i>	<i>khim</i>
red=COLOR.ART.SG	house
<i>eiʔ=na</i>	<i>khim</i>
big=ART.SG	house

Distinction neutralized in the plural (and dual):

<i>phabeleŋ=ha</i>	<i>khim</i>
red=ART.PL	house
<i>eiʔ=kha</i>	<i>khim</i>
big=ART.PL	house

## Unexpected voicing patterns

Common alternation in Eastern Kiranti: *ph* ~ *bh* etc. intervocalically

e.g. Belhare *khi-thaŋs-e* ‘quarrel-upwards-PST’: /khi<sup>h</sup>θaŋse/  
‘S/he quarreled with someone standing further uphill.’  
(Bickel 1996:60)

## WHY?

Why are all these rara where they are?

Two issues:

- The *source* of rara
- The *survival and statistical visibility* of rara

## One source of rara: local “blends”

Puma (S. Kiranti) (Bickel & Gaenszle 2005)

*kha-en-a.*

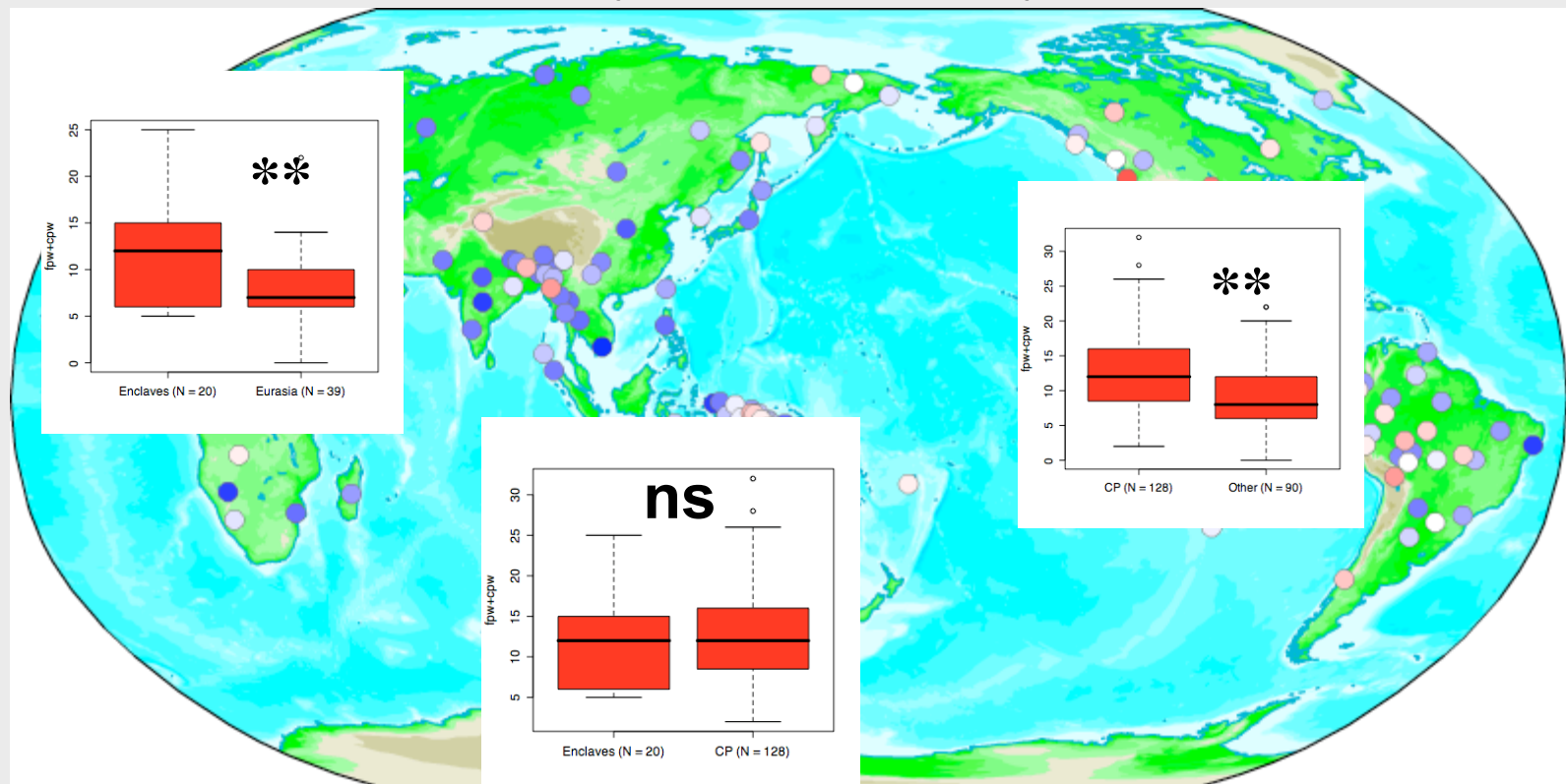
ANTIPASSIVE-hear-PST

- a. ‘S/he heard someone / people.’ or ‘S/he listened so as to find out whether or not there are people.’ (does not entail existence of a specific undergoer referent)
- b. ‘S/he heard us (incl.)’

- generic PO = 1PO found in several branches of the family
- only in the south: political alliances with Maithili-speaking kingdoms
- Maithili similar role in the Central Himalayas as French in Europe
- parallels patterns in Maithili: ‘zero-ing’ of 1st person for politeness reasons (Bickel et al. 1999)
- Kiranti peculiarity: zero-ing only of 1PO, not of 1A or 1S because Recipients are particularly sensitive socially.

# Another source of rara: enclave effects

- Some relative rara, e.g. a high degree of synthesis, appear to stem from the same population that characterize the Circumpacific area, so they are perhaps very old





# Visibility of rara

- What favors the visibility (and hence survival) of rara in the Himalayas and the Caucasus is the **absence of massive cross-family substrate interference (language shift) over at least 2Ky.**
- No evidence for mass shifts in the Himalayas before the Gorkha (Nepali-speaking) conquests starting in the 1770s.
- By contrast:
  - Large language spreads that affected northern Eurasia, South Asia, and Southeast Asia over several millenia resulted in a decrease of rarities.
  - Post-1770 mass shift (into Nepali) results in a decrease of rarities
  - (note the difference between Kusunda and Nahali (Watters 2005))
- **Absence of mass shifts and substrates  $\neq$  isolation! (cf. Maithili influence in the Southern Kirant)**
- **Local stabilization (Nettle 1999) is not necessarily a factor (cf. free prefix ordering)**

## If this is right...

- The statistical visibility of rara (the fact they are numerically rare but diachronically strong enough to be detectable) is not caused by anything structural (suboptimality etc.)
- Instead, we perceive 'rara' because other patterns have spread far, through massive substrate effects
- The frequency distributions we perceive statistically cannot *by themselves* evidence *structural* rara and universals, i.e. rara and universals as properties of the human language faculty.
- To claim structural rara and universals we also need to show that the current distributions are independent of any population history behind them (areas: Dryer 1989, earlier areas: Maslova 2000)
- no small task... and, anyway, not my task here...

# Acknowledgments: AUTOTYP

- **The AUTOTYP research team**
  - Johanna Nichols (Co-Director, Berkeley)
  - Balthasar Bickel (Co-Director, Leipzig)
  - René Schiering (Post-Doc, Leipzig)
  - Kristine Hildebrandt (Research Associate, Manchester)
  - Fernando Zúñiga (Research Associate, Santiago/Chile)
  - *RAs in Berkeley*: Gabriela Caballero, Nicole Marcus, Suzanne Wilhite
  - *RAs in Leipzig*: Alena Witzlack, Anja Gampe, Jenny Seeg, Thomas Goldammer, Sven Siegmund, Franziska Crell
  - *Past team members*: Sebastian Hellmann, Michael Riessler, Kathi Stutz, Josh Wilbur, Rebecca Voll, Sandra Biewald, Aimee Lahaussois-Bartosik, Dave Peterson, Keith Sanders
- **Funding**: Swiss NSF Grant Nos. 08210-053455 (1998-2001, Bickel) and 610-0627 (2001-2002, Bickel), German DFG Grant No. BI 799/2-1 (2003-2007, Bickel), US NSF Grant No. 96-16448 (1998-2001, Nichols), U. Leipzig and UC Berkeley funds

# Acknowledgments: CPDP

- **The CPDP research team**

- Balthasar Bickel (Leipzig)
- Martin Gaenzle (Leipzig)
- Novel K. Rai (Kirtipur)
- Vishnu S. Rai (Kirtipur)
- Elena Lieven (Leipzig)
- Sabine Stoll (Leipzig)
- Goma Banjade (Kirtipur)
- Netra Paudyal (Kirtipur)
- Manoj Rai (Kirtipur)
- Ichcha P. Rai (Kirtipur)
- Manoj Rai (Kirtipur)

- Rikhi Maya Rai (Kirtipur)
- Janaki Rai (Kirtipur)
- Narayan Sharma Gautam (Kirtipur)
- Arjun Rai (Kirtipur)
- Sri Kumar Rai (Biratnagar)
- Prem Dhoj Rai (Beltar)
- *Student assistants in Leipzig:*
  - o Sindy Poppitz
  - o Sven Siegmund
  - o Diana Schackow

- **Funding:** VW DOBES Grant No. II/79 092, 2004-2007 (Bickel); U. Leipzig funds