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

• CPDP fieldwork project on Kiranti languages in Nepal: 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)

Bickel & Nichols 2003 and in progress
Relative rara in the Eurasian Enclaves

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

\(N = 14\)
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

Relative rara in the Eurasian Enclaves

Other hypothesized effects (not tested yet):

• Complex or at least bipartite stems (Bickel & Nichols 2003), leading to endoclisis (Harris 2002 on Udi in the Caucasus; Bickel et al. 2005 on Chintang in the Himalayas), exuberant agreement (Harris 2006) and the like.


• 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
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
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
11. Spatial interjections: Chintang, Belhare
13. Aspirated stops alternating with breathy voiced stops (\(\text{ph} \rightarrow \text{bf} V \_V\)) : Limbu, Belhare
   ? Voiceless-voiced clusters and pseudo-geminates: Puma, Belhare
Upside-down ergativity splits

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

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>S</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1s</td>
<td>-η (&gt;3)</td>
<td>-na (&gt;2)</td>
<td>-ηa (~-οη /PST)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1d</td>
<td></td>
<td></td>
<td>-ci-∅</td>
</tr>
<tr>
<td>1p</td>
<td>-m</td>
<td></td>
<td>-i</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>tʌ-</td>
</tr>
<tr>
<td>3s</td>
<td>ø-</td>
<td></td>
<td>-i</td>
</tr>
<tr>
<td></td>
<td>pʌ- (&gt;1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3d</td>
<td>pʌ- -ci</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ni- (&gt;2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3p</td>
<td>mʌ- (&gt;3n s)</td>
<td></td>
<td>-ci</td>
</tr>
<tr>
<td></td>
<td>ni- pʌ- (&gt;1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ni- -nin (&gt;2, 3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*ben-ηa* ‘I come over’

*𝑝ʌ-en-ηa* ‘S/he heard me’

*enn-u-η* ‘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. ø khoŋ-ma nui-ka.
   [S] play-INF may-2sNPST
   ‘You may play.’

b. ø ø 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 endoclitic 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:

\[
\begin{align*}
[S^\cdot (\_\_ [S met]-na)-bi]-na-?ã-ni \\
do-1\rightarrow 2-V2: \text{BENEFACTIVE}-1\rightarrow 2-\text{eNPST}-p
\end{align*}
\]

‘I’ll do it for you (p).’

\[
\begin{align*}
[S^\cdot (\_\_ ko]-na)-gon]=lok... \\
walk-NA-V2: \text{AMBULATIVE}=\text{SIM}
\end{align*}
\]

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

\[
a-mas-u)-and-u)-bid-u-ku-m=ní
\]

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

‘You’ll lose it on him!’
Reflexive verb stems

Puma

\[ [\Sigma \cdot (\phi [\Sigma dher]-o\gamma)]-cen]-o\gamma \]
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’: \textit{teï},-yu  
tenchinma ‘to hit oneself’: \textit{ten-chiï},-yu

also: Yamphu (Rutgers 1998)
Middle EAT

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

\textit{khoŋ-ca-he}

play-MIDDLE- PST[3sS]

‘She played by herself.’
Triplication not via reduplication

Chintang (Rai et al. 2005)

*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=Ø (‘here’) series

<table>
<thead>
<tr>
<th></th>
<th>PROXIMAL</th>
<th>DISTAL</th>
<th>F=S=Ø</th>
<th>Ø≠S</th>
<th>R≠S</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UP</strong></td>
<td>toba</td>
<td>atu(ba)</td>
<td>bandu</td>
<td>tobandu</td>
<td>utu(ba)</td>
</tr>
<tr>
<td><strong>DOWN</strong></td>
<td>moba</td>
<td>amu(ba)</td>
<td>bamu</td>
<td>mobamu</td>
<td>umu(ba)</td>
</tr>
<tr>
<td><strong>ACROSS</strong></td>
<td>yoba</td>
<td>ayu(ba)</td>
<td>bayu</td>
<td>yobayu</td>
<td>uyu(ba)</td>
</tr>
</tbody>
</table>
Spatial interjections

Belhare:

\textit{tu!} vs. \textit{mu!} vs. \textit{yu!}

Chintang:

\begin{itemize}
\item \textit{to, toto, toi, togoi}
\item \textit{mo, momo, moi, mogoi}
\item \textit{yo, yoyo, yoi, yogoi}
\end{itemize}
Color-sensitive articles

Belhare (E. Kiranti, Bickel 2003):

\[
\begin{align*}
phabele\ = & \ m\ & khim \\
red=COLOR.ART.SG & & house \\
ei\ = & \ n\ & khim \\
big=ART.SG & & house
\end{align*}
\]

Distinction neutralized in the plural (and dual):

\[
\begin{align*}
phabele\ = & \ h\ & khim \\
red=ART.PL & & house \\
ei\ = & \ k\ & khim \\
big=ART.PL & & house
\end{align*}
\]
Unexpected voicing patterns

Common alternation in Eastern Kiranti: \( ph \sim bh \) etc. intervocalically

e.g. Belhare \( khi-thaŋs-e \) ‘quarrel-upwards-PST’: /khid\( ^h \)aŋse/
‘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.

Bickel & Nichols 2003
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 ≠ 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

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