Genetic stability in the contact zone of Sino-Tibetan and Indo-European

Balthasar Bickel
University of California, Berkeley and University of Zürich
[mailto:bickel@socrates.berkeley.edu; http://socrates.berkeley.edu/~bickel or /~autotyp]

I. Areal spread features in the Himalayan ST/IE contact zone

Sino-Tibetan (ST) and Indo-European (IE) are prime examples of how strongly a language family can typologically diversify under the pressure of areal spread features:

**Phonology:** mildly to heavily fusional phonological word structure (ω)

(1) a. Nepali (IE) (ω gari-raheko-cha-s)
   do-PROGRESSIVE-NPT-2SG
   ‘You are doing it right now.’

b. Belhare (ST) (ω cokg-hett-u-ga)
   do-TEMPORARY-3[SG].O-2[SG].A
   ‘You are doing it right now.’

c. Belhare (ST) (ω mi-ηη-u-ukg-att-u-n-chi-nn-hak=cha)
   3NSG.A-NEG-roast-bring.down-PT-3O-NEG-NSG.O-NEG-N=ADD
   ‘They didn’t even roast it for them down here.’

(2) a. IE in Europe: smallish ω’s, often including clitics, e.g. English
   (ω You’re) (ω doing it).

b. ST in SEA: small ω’s, often splitting grammatical words, e.g., Lai Chin (ST)
   (ω na-tuk) (ω nhaa) (ω lāay).
   [GrWd 2SG.A-hit.with.stick:Σ2 3PL.O FUT]
   ‘You will hit them.’

**Morphology:** much V-compounding (synthesis), less N-compounding

(3) a. Nepali (IE) āi-pug- ‘come-reach’, i.e., ‘arrive’

b. Thulung (ST) rom-phā ‘come-reach’, i.e., ‘arrive’ (Ebert 1994)

(4) a. IE in Europe: much N-compounding, hardly any V-compounding

b. ST in SEA: much N-compounding, much V-compounding, e.g., Lahu
   phi̯-qhè làʔ-nɔ ‘dog-dung finger’, i.e. ‘index finger’
   phe-chiʔ ‘restrain-bind’, i.e., ‘tie up’ (Matisoff 1973)
**Syntax:** convivial or conjunctional cosubordination (chaining with underspecified operator scope; ± finite)

today evening DEM-FOC-DAT curry make-CONV eat-INF must-3SG.NPT
‘Tonight [we] should cook this as a curry and eat it.’

b. Limbu (ST) — biha-n balla ke-ma:nd-u-ա nj ke-dhaŋ-e-i?:
marriage-ART.NOM finally 2[SG.A]-finish-3O-CONV 2[SG.A]-come.up-PT-Q
‘Did you come up after you had finally finished [attending] the wedding?’
— ā ma:nd-u-ա nj than-ա.
yes finish-3O-1SG.A-CONV come.up-1SG.[S].PT
‘Yes, I finished it and came up.’ (van Driem 1987:284)

c. Belhare (ST) khar-e kī jutta nj-in-ghutt-he-ga i?
‘Did she go and buy you shoes?’
or ‘Did she buy you shoes when she went?’

(6) a. IE in Europe: subordination (disjunct. operator scope) vs. coordination (conj. scope)

b. ST in SEA: Verb serialization (= mostly unmarked cosubordination)

all paddy carry.on.shoulder put into PERSUASIVE
‘Please carry all the paddy [home] and put it into [the storeroom].’

**Semantics:** PATH as a MOTION concept in verbs¹ (Talmy 1985, Slobin & Hoiting 1994, etc.)

jungle-LOC run-CONV enter-3SG.PT

house-LOC run-CONV [3SG.S-]enter-PT

‘He ran into the jungle.’ (‘Il est entré dans la forêt (en courant).’)

(8) IE in Europe: PATH as MOTION (Rom.) or PATH as GROUND PROPERTY (Germ., Slav.)

ST in SEA: PATH as MOTION but DEPENDENT on (serialized with) co-occurring manner verbs in head function; e.g., lo? ‘into, enter’ in (6b) kɔ lɔ? ‘put into’ only with motion verbs (Matisoff 1973:222)

---

**Discourse:** INFORMATIONAL DIPTYCH as a pan-SA announcement strategy

(9) Belhare *emu cok-yu / cek-yu bhanе* ... follows a narrative/report
   how [3SG-do-NPT [3SG-say-NPT TOPIC/COMP

Nepali *ke gar-cha / bhan-cha bhanе*... follows a narrative/report
   what do-3SG.NPT say-3SG.NPT TOPIC/COMP

Maithili *kathi kar-ait / kah-ait chai je* ... follows a report
   what do-PART say-PART AUX.3 COMP

‘And then s/he did/said…’ *(literally: ‘What s/he does/says [is] that…’)*

II. Less diffusion of interface principles

**Phonology/morphology interface:** ST languages differ from IE languages in showing a strong bias against metrical constituents containing concatenative morpheme boundaries. This is the Sino-Tibetan TAUTOMORPHEMICITY Principle.²

(10) Tautomorphemic σ in SEA-ST
   a. Syllabic morphemes, cf. Lai Chin ex. (2b)
      Garo: all but one morpheme are syllabic (Burling 1961:6);
      Exceptional morphemes tend to become nonconcatenative, e.g. *-t and *s- causatives often develop into phonation and aspiration, respectively:
      Lai Chin *fiag* ‘be clear’ ~ *fi后排* ‘make clear’ (Peterson 1998)
      Lai Chin *pit* ‘be blocked’ ~ *phet* ‘block’
   b. Onset-free syllabification:
      Garo (*ca?-(a), not *(ca)(a?)-a* ‘eat-HAB’, as shown by *(a)? (Burling 1961:5)
      (*kat-(a), not *(ka)(t-a) ‘go-HAB’, as shown by [kʰat’a], not [kʰat’a]
   c. ₫-Prothesis (morpheme-juncture gemination):
      Meithei *thəm-u* ‘keep-IMP’ → (*thəm-)((a) m-u) (Burling 1961:67)
      təw-e ‘do-ASS’ → (*təw-)((a) w-e), not *(tə)((a) w-e) (Burling 1961:23)

(11) Tautomorphemic σ in SA-ST
   a. Belhare: only 7 out of 80 allomorphs are subsyllabic
      (-ŋ ‘1SG.A’, ‘1SG.S’ in neg. form; -m ‘1/2PL.A’; -n ‘NEG’ after V; -t ~ -? ‘NPT’)
      Fate of of *-t and *s- as in SEA, e.g.
      Belh. *pok- ‘rise’ ~ phok- < *s-pok ‘raise’; on the fate of *-t, see below.

---

cf. Tibetan ergative WT -s > [+front], [H?], e.g., WT kho-s ‘he-ERG’ = /khø^h/.

b. Onset-free syllabification
Dolakha Newar (\(\alpha\)ye-)(\(\alpha\)e) ‘come-N’, not *(\(\alpha\)ye:) (Genetti 1994:30)
Belh. (\(\alpha\)yu-)(\(\alpha\)a) ‘go.down-IMP’, not *(\(\alpha\)yu)

(12) Tautomorphemic \(\phi\) in SEA-ST
a. Mostly 1 morpheme = 1\(\sigma\) or 2\(\mu\) = 1\(\phi\) = 1\(\omega\)
b. Underparsing of \(\sigma\), as in the Lai Chin ex. (2b), p-clitics in sesquisyllabic\(^3\) \(\omega\)’s

\(\omega\) na (\(\phi\) tuk) \(\omega\) (\(\phi\) nhaa) \(\omega\) (\(\phi\) âay)
2SG.A hit.with.stick:Σ2 3PL.O FUT

(13) Tautomorphemic \(\phi\) in SA-ST:
  a. Mostly 1 morpheme = 2\(\mu\) (CVC or CVCV) = 1\(\phi\)
  b. \(\cap\)-Prothesis (morpheme-juncture gemination):
Maivä-Mevä Limbu huk-en ‘hand-ART.NOM’ → (\(\phi\) huk-)\(\phi\) en ‘the hand’,
not *(\(\phi\) hu)\(\phi\) k-en or *(\(\phi\) hu)k-en (Michailovsky 1986)
Belh. lap-uk-ma ‘catch-bring.down-INF’ → (\(\phi\) l)ap\(-\)\(\phi\) uk\(-\)ma, not *(\(\phi\) l)\(\phi\) b-uk\(-\)ma
  c. Underparsing of \(\sigma\) (even at the cost of degenerate feet)
Belh. lap-u-k=cha ‘catch-3O-2A=ADD’ → (\(\phi\) l)ap\(-\)\(\phi\) b-u-k=cha, not
*(\(\phi\) l)ap\(-\)\(\phi\) b-u-k=cha
Belh. lap-u ‘catch-3O’ → (\(\phi\) l)ap\(-\)\(\phi\) b-u, *(\(\phi\) l)ap\(-\)\(\phi\) u
  d. Underparsing of segments (deletion): the fate of *-t (and *-s) in Belhare
n-lu-t-att-u-n ‘NEG-tell-T-PT-3O-NEG’ → n(\(\phi\) l)u\(\phi\) ,at\(-\)u-n, not *n(\(\phi\) l)\(\phi\) r-at\(-\)t-u-n
hi-t-ma ‘be able-T-INF’ → (\(\phi\) hi)ma, not *(\(\phi\) hi-t)\(\phi\) ma or *(\(\phi\) hi)\(\phi\) t\(\phi\) ma
hir-e ‘be able-T-PT’ → (\(\phi\) hi)r-e , not *(\(\phi\) hi)\(\phi\) e or *(\(\phi\) hi)

(14) Contrast to IE in SA:
Maithili restriction of inflectional desinences to 2\(\sigma\), with two effects:\(^4\)
  a. Suffix lot 3 Allomorphy -ainh ~ -nh

---

\(^3\) Matisoff, J.A. 1999. Genetic vs. areal linguistics in Southeast Asia: prosodic diffusibility in Southeast Asian languages. Ms. UC Berkeley.

b. Only one triple-agreement form:
\[ \text{dekhau-l-i-au-nh} \]
show-PT-1NOM-2NONHON.NONNOM-3HON.NONNOM
‘I showed him/her to you.’ or ‘I showed you to him/her.’ or ‘I showed his/her X to you.’ (Y.P. Yadhav, p.c.)
but not *-ahikunh ‘2MIDHON.NOM-3NONHON.NONNOM-3HON.NONNOM’,
*-ahinhunh ‘2MIDHON.NOM-3HON.NOM-3HON.NONNOM’, etc.

Syntax/semantics interface: ST languages differ from IE languages in mapping verb-defined semantic roles directly to grammatical relations (if there are any), without regard to information encoded by cases or phrase-structural positions. Grammatical relations in IE languages, by contrast, are systematically sensitive to lexical or constructional case frames or phrase structures. This is the Indo-European INTEGRATIVITY Principle.

(15) IE in both Europe and SA; Nepali:
   a.  \( \text{qar\text{\-}unu} \) ‘to fear’:\( \langle \text{exp., stim.}, \langle \text{NOM, sa\text{\-}nga} \rangle \rightarrow \text{exp.} \in \{S,A\} \)

   a’. \( \text{ma} \quad \text{bhut sa\text{\-}nga qar\text{\-}e} \).
1SG.NOM  ghost  with fear-1SG.PT
‘I was afraid of the ghost.’

   b.  \( \text{qar l\text{\-}gnu} \) ‘id.’:  \( \langle \text{exp., stim.}, \langle \text{DAT, sa\text{\-}nga} \rangle \rightarrow \text{exp.} \notin \{S,A\} \)

   b’. \( \text{ma-l\text{\-}i} \quad \text{bhut sa\text{\-}nga qar lag-yo} \) (*lag-\( \text{\-}e \)).
1SG-DAT  ghost  with fear feel-3SG.PT  feel-1SG.PT
‘I was afraid of the ghost.’

(16) Maithili (IE; Nepal)
   a.  \( \text{o qar\text{\-}l\text{\-}aith} \).
3HON.DIST:NOM  be.afraid-PT-3HON.NOM
‘S/he/they was/were afraid.’

   b.  \( \text{hunk\text{\-}a qar lag-l\text{\-}ainh} \).
3HON.DIST:DAT  fear[NOM]  feel-PT-3HON.NONNOM
‘S/he/they was/were afraid.

(17) ST in SA, e.g. Dolakha Newar (Genetti 1994)

   gy\text{\-}ye ‘to fear’: \( \langle \text{exp., stim.}, \begin{cases} \langle \text{NOM, NOM} \rangle \\ \langle \text{DAT, NOM} \rangle \end{cases} \rightarrow \text{exp.} \in \{S,A\} \)

   a.  \( \text{chi h\text{\-}atta gy\text{\-}at-an} ? \) (Genetti 1994:202)
2SG.NOM  why  fear-2SG.PT
‘Why were you afraid?’
b. *thau-ta* gibiŋg *ma-gyāt-ki* (Genetti 1994:53)

\[
\text{REFL-DAT nothing:NOM \ NEG-fear-1SG.PT (NB: REFLEXIVE in 1SG use)}
\]

‘I wasn’t afraid at all.’

(18) Belhare

a. *kitma* ‘to fear’:

\[
\text{<exp., stim.>, <ERG, NOM> \rightarrow exp.} \in \{S,A\}, \text{stim.} \in \{S,O\}
\]

a’. *han-na* tombhira *kii-t-u-ga* i?

\[
\text{2SG-ERG lynx[SG,NOM] \ fear-NPT-3[SG]O-2[SG.A] \ Q}
\]

‘Are you afraid of the lynx?’

b. *niũa* tima ‘to like’:

\[
\text{<exp., stim.>, <POSS, NOM> \rightarrow exp.} \in \{S,A\}, \text{stim.} \in \{S,O\}
\]

b’. *ŋka* *hale* *hani-ŋiũa* *ka-tiu-s-ik-kha*.

\[
\text{1SG[NOM] before \ 2PL.POSS-mind[NOM] \ 1SG.O-spend-TR-2[PL.A]-PERF}
\]

‘Before, you liked me.’

The same difference between ST and IE can be observed in control, raising, relative, converb constructions.5

III. Conclusions

- **Finding:** Despite extremely intense language contact (with systematic bilingualism, code-switching, and language shift), principles regulating the Phonology/Morphology and the Syntax/Semantics Interface have by and large resisted diffusion in the ST/IE contact zone.

- **Hypothesis:** Interface principles have generally a lower diffusion potential than most single-mode principles.

- **Explanation:** Interface principles are less cognitively transparent (less accessible for copying or substratal retention) and have less immediate communicative value than most single-mode principles. In this regard, they compare to inflectional classes, one of the genetically most robust pattern of grammar.

**Abbreviations**
