The clause linking verb suffix *-la* in Bulu Puroik and its diachronic origin

Ismael Lieberherr

September 29, 2017

PhD defence

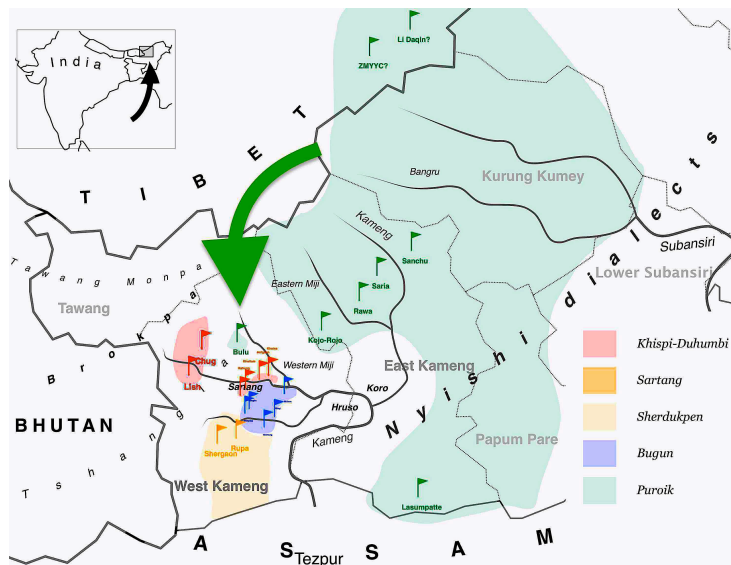## 1 The language and the PhD project

Figure 1: Map of western Arunachal Pradesh (Lieberherr and Bodt in press 2017)

"Three Ph.D. fellows will complete grammatical descriptions of **strategically chosen languages** of the Indian subcontinent"… "The choice of languages will be made on the basis of (1) the **urgency of documentation** in view of the degree of language endangerment, and (2) the **potential significance** of the chosen language in terms of what light the language may shed on the phylogeny of its language family and on Asian population prehistory."
`http://p3.snf.ch/Project-138331`

Potential significance of Puroik

- Original population of western Arunachal Pradesh (among the last sago cultivators in the Himalayas).

- Diversity of Puroik undocumented. Previous research on eastern varieties[1]

- Bulu Puroik is the last Puroik dialect spoken between Kho-Bwa languages and other Puroik

Endangerment: <20 speakers

- 6 cousin brothers living in Bulu

- their sisters living in other villages

- their non-Puroik wives (only some)

- their sons/daughters (only some)

Outcomes

- A Grammar of Bulu Puroik:
    - Phonology
    - Morphosyntax
    - Registers (Hunting language, Ritual language)

- Lexicon (around 4900 entries, 400 pictures, 4000 recordings)

- Annotated recordings archived on paradisec:
  `http://catalog.paradisec.org.au/collections/BULU1`

- Software (minimalistic plain text lexicography and corpus) forthcoming on Github:
  `https://github.com/metroxylon`

- Papers (e.g. Lieberherr 2015 about affiliations of Puroik)

---

[1] Tayeng 1990, Remsangpuia 2008, Soja 2009, Li 2004

2 Syntax and function of verbs marked with *-la*

Verb suffix on verb stem: *(ba-)*ROOT(-DERIV)(-DERIV)-*la*

Non-finite: Morphological

- never combined with tense-aspect-marker: *-na* FUT, *-tʃa* PRF, *-ba* PRS (Impossible: *-tʃa-la*, *-na-la*, *-ba-la*)

- however: past is unmarked, i.e. on the surface it is not possible to distinguish a non-finite form from a finite past

Non-finite: Syntactic

- rarely on predicates of "independent" sentences (but possible)

A then B then C (one after another)

(1) [*guu=na atʃoi vuu -la*]₁   [*rɨɨ -la*]₂ [*guu ham   uN -la*]₃   [*guu mɘruu*
    1SG=TOP far go.from.base-SEQ stay-SEQ 1SG house go.to.base-SEQ 1SG woman
    hɛɛ   ʃiN   ba-ʒeN]
    what wood NEG-carry
    'I went far, stayed, came home and my wife doesn't carry any wood.' [WOOD39:22]

A and B and C (at the same time)

(2) typical farewell to a person leaving

    [*amjɛɛ niŋ -la*]₁   [(*ham*) *uN-bo*]₂
    good   watch-SEQ (home) go.to.base-IMP

    'Go back carefully' (lit. 'watching well').

Reason

(3) [*adɘɘ=la.ɛNku ba-ʒeN-muɛN -la*]₁ [*guu ʃiN   ʒeN-tʃiʔ-na   baʔ*]
    child=SOC   NEG-carry-can -SEQ 1SG   wood carry-OBLG-NPST EXIST
    'Because [my wife] cannot carry [wood] with the baby, I have to carry wood.'
    [WOOD39:28]

Posture verbs

(4) a. POSTURE + *rɨɨ*
       reʔ -la rɨɨ
       lie-SEQ sit
       'to be lying'
    b. POSTURE + *lii*
       reʔ -la lii
       lie-SEQ put

'to lie someone down'

Progressive

(5) ACTIVITY + *rɨɨ*

    ruʔ -la   rɨɨ
    bark-SEQ sit

    'be barking'

"Altaic sense" inappropriate:
"a non-finite verb form whose main function is to mark **adverbial subordination**" (Haspelmath 1995, p. 3)

Ok, in a South Asian sense:
"a converb is a **non-finite, clause-linking inflectional verb form** that is dependent upon a matrix verb and cannot function as a sentential argument" (Coupe 2017, p. 8)

- sequence of clauses (Eng. 'and, then')

- adverbial linking (Eng. 'if, when, because')

- complex predicates (Other languages: lexical, inflectional, auxiliary construction)

rank	morph	gloss	#
1	*-la*	SEQ	495
2	*dʒi*	ANA	287
3	*grɨɨ*	1PL	216
4	*=ku*	LOC	199
5	*uN*	'go.to.base'	185
6	*la*	CONJ	165
7	*-na*	NPST	164
8	*-ka*	REL	136
9	*ba-*	NEG	135
10	*vɛɛ*	3SG	134

Figure 2: 10 most frequent morphemes in texts

3 Diachronic origin of -*la*

Diachronic origin of *la*

Synchronically connected: conjunction *la*

(6) [… *məhjɛN atsoN ɬuʔ-∅*]₁ [*la* *vɛrii ba-dɛN-dyiN*]₂
earth   center fall-PST CONJ 3PL   NEG-know-again
'The sulphur springs fell down in the center of the earth and they didn't realise.'
[SULPH23:24]

(7) [*səmjɛʔ   tʃakui vjɛɛ -la*]₁ [*la* *ii-kəljuu-la   rɨi*]₂
amaranth rice   throw-SEQ CONJ die-pretend-SEQ stay
'He threw cooked amaranth grains on [on his anus], and pretended to be dead.'
[MASANG02:50]

Conjunction and verbal morpheme

Clause linking

1. [CLAUSE-*la*]₁ [CLAUSE]₂ (verb suffix)

2. [CLAUSE]₁ [*la* CLAUSE]₂ (conjunction)

3. [CLAUSE-*la*]₁ [*la* CLAUSE]₂ (suffix and conjunction)

Reanalysis
ambiguity: unmarked past = non-finite stem

(8) a. [*nam-∅*]₁ [*la* *fintʃuɛʔ*]₂
smell-PST CONJ dispose

b. [*nam -la*]₁ [*fintʃuɛʔ*]₂
smell-SEQ   dispose
'He smelled and threw it away.'

Another source for *la*: case markers

Case/postposition > converb suffix
Typologically most common source for converbs (Haspelmath 1995, p. 17). Very common in TB (Genetti 1986, Thurgood 1986)

Locative
Proto-Puroik **-la* 'locative'

Locative to co-temporal subordinator ('when')

- Bulu Puroik =*ku* (modern locative)

- Duhumbi (Chugpa) -*kʰo* (but on nominalisation -*ba*) Bodt 2017, p. 794

- Tshangla -*ka/ga* Bodt 2017, p. 799

- Galo CLAUSE=*lo* (Post 2007, p. 793)

Locatives in Puroik and beyond

Proto-Puroik **-la*
-*la* all other Puroik dialects except Bulu =*ku*

Traces of *la*-locative in Bulu Puroik
nuu la 'here', *tɛɛ la* 'there'
dʒi la local 'there', temporal 'then'
= *la pu* 'ABL'

Other TB languages:
-*lo* Galo and "most Tani languages" (Post 2007, p. 737)
-*la* Koro (unclassified TB) (Anderson and Murmu 2010, p. 8)
*la, *na Tibetan (LaPolla 2004, p. 44)
*ra, *la, *ri other Bodic (Noonan 2008, p. 17)
*lə Karen (LaPolla 2004, p. 44)
lo Lahu (Loloish) (Matisoff 2006, p. 147)
(-*loʔ* Duhumbi (ablative and clause linker) (Bodt 2017, p. 426))

Locative also source of conjunction?

apheresis *la* < *dʒila*?

(9) [*VERB-∅*]₁ [*la* *vɛɛ=na klaa   rin ba-rin-mulɔ*]₂
VERB-PST CONJ 3SG=TOP where run NEG-run-DESP
'He didn't know where to run.' [WAR01:37]

(10) [*VERB-∅*]₁ [*dʒila vɛɛ=na hɛɛ ba-tsaʔ-mulɔ-la*]₂
VERB-PST CONJ 3SG=TOP what NEG-do-DESP-SEQ
'He didn't know what to do.' [WAR02:09]

Etymology of conjunction *dʒila*

- *dʒi* anaphoric pronoun + **-la* locative

- Semantic change: THERE > THEN (space to time)

- like *dʒilapu* 'after this' = *dʒi* + =*lapu* ablative

4 Summary

Summary

Main points

- Verb suffix -*la* is the most frequent morpheme in Bulu Puroik with multiple functions:
  - clause chaining
  - adverbial linking
  - first part of complex predicates

- Bound morpheme -*la* is synchronically connected to conjunction *la* (…-*la la* … construction), maybe diachronically (reanalysis)

- Hypothesis: ultimate origin of both is locative **-la*

References

Anderson, Gregory D. S. and Ganesh Murmu (2010). "Preliminary notes on Koro, a 'hidden' language of Arunachal Pradesh". In: *Indian Linguistics* 71, pp. 1–32.

Bodt, Timotheus Adrianus (2017). "A Grammar of Duhumbi". PhD thesis. Bern: Institut für Sprachwissenschaft, Universität Bern.

Coupe, Alexander R. (2017). "On the diachronic origins of converbs in Tibeto-Burman and beyond". In: *Sociohistorical Linguistics in Southeast Asia: New horizons for Tibeto-Burman studies in honor of David Bradley*. Ed. by Picus Ding and Jamin Pelkey, pp. 210–237.

Dondrup, Rinchin (1990). *Bugun Language Guide*. Itanagar: Government of Arunachal Pradesh.

Genetti, Carol (1986). "The development of subordinators from postpositions in Bodic languages". In: *Annual Meeting of the Berkeley Linguistics Society*. Vol. 12, pp. 387–400.

Haspelmath, Martin (1995). "The converb as a cross-linguistically valid category". In: *Converbs in cross-linguistic perspective*. Ed. by Martin Haspelmath and Ekkehard König, pp. 1–55.

LaPolla, Randy J. (2004). "On Nominal Relational Morphology in Tibeto-Burman". In: *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-cherng Gong on his Seventieth Birthday*. Ed. by Ying-jin Lin et al. Taipei: Institute of Linguistics, Academia Sinica, pp. 43–74.

Lǐ, Dàqín 李大勤 (2004). *Sūlóngyǔ yánjiū 苏龙语研究 [Research on Puroik]*. Běijīng: Mínzú chūbǎn shè 民族出版社 [National Minorities Publisher].

Lieberherr, Ismael (2015). "A progress report on the historical phonology and affiliation of Puroik". In: *North East Indian Linguistics*. Ed. by Linda Konnerth et al. Vol. 7. Canberra: Australian National University: Asia-Pacific Linguistics, pp. 235–286.

Lieberherr, Ismael and Timotheus Adrianus Bodt (in press 2017). "Sub-grouping of Kho-Bwa based on cognate core vocabulary". In: *Himalayan Linguistics* 16.2.

Matisoff, James A. (2006). *English-Lahu Lexicon*. Vol. 139. University of California Press.

Noonan, Michael (2008). "Case compounding in the Bodic languages". In: ed. by Greville G. Corbett and Michael Noonan, pp. 127–148.

Post, Mark W. (2007). "A Grammar of Galo". PhD thesis. Bundoora: La Trobe University.

Remsangpuia (2008). *Puroik Phonology*. Shillong: Don Bosco Center for Indigenous Cultures.

Soja, Rai (2009). *English-Puroik dictionary*. Shillong: Living Word Communicators.

Tayeng, Aduk (1990). *Sulung Language Guide*. Itanagar: Directorate of Research, Government of Arunachal Pradesh.

Thurgood, Graham (1986). "Lolo-Burmese subordinators from case postpositions: several partial etymologies". In: *Proceedings of the Second Annual Conference of the Pacific Linguistics Conference*. Ed. by Scott DeLancey and Russell S. Tomlin, pp. 449–453.