Toward a Eurasian Bestiary:
(I)   OTTER in Tibeto-Burman and Mon-Khmer
*(II)   JACKAL in Sino-Tibetan and Indo-European**

James A. Matisoff
University of California
Berkeley

PART I: *OTTER in Tibeto-Burman and Mon-Khmer*

I.1   **Introduction**

One of the most interesting zoonyms to be found in *Sino-Tibetan: a Conspectus* [henceforth STC] is the word for 'otter' (#438, p. 107). On the basis of the following 11 forms, Benedict originally reconstructed this etymon as Proto-Tibeto-Burman [PTB] ***s-ram**, with the initial sibilant analyzed as an instance of the ***s-** "animal prefix":[1]

> Written Tibetan [WT] **sram**, Lepcha **săryom**, Miri **si-ram**, Nung **səram**, Jingpho [Jg.] **shăram**, Maru **xrɛn**, Phunoi **sam**, Garo **matram**, Dimasa **matham**, Lushai (Mizo) **sa-hram**, Mikir **serim**.

Later, however, he changed his mind, and removed the hyphen from the proto-form (***sram**), treating the sibilant-plus-liquid combination as a true consonant cluster, with no morpheme boundary intervening. This move was prompted by a reconsideration of Mizo **sa-hram** and Lepcha **săryom**, both of which look as if they derive from ***sa-sram**.[2]

Yet as Benedict himself realized, this reanalysis does nothing to explain the very strange Written Burmese [WB] word for 'otter': **phyam**. In a footnote (STC, n. 302) he offers a non-explanation without much enthusiasm:

> "…Burmese has **phyam** 'otter', which can be analyzed as a derivative of ***phram** < ***p-sram**, with the **p-** element of undetermined origin."

This Burmese form certainly "otter" undergo further scrutiny – a theory to account for its mysterious labial onset will be the main focus of the first section of this paper.

To begin with, it seems to me that Benedict's hypothetical ***p-sram** prototype is overly complex, and is merely an artifact of his concern to justify the "cluster" rather than "prefixal"

---

[1] This morphological element, which appears in words for parts of the body as well as in animal names, shows up sporadically throughout Tibeto-Burman [TB], and is obviously derived from the root-morpheme for 'animal; flesh' reconstructed as PTB ***sya** (STC #181).

[2] Both the Mizo voiceless sonorant **-hr-** and the Lepcha palatalized **-ry-** may be referred to an original **s** + **r** sequence. For the peculiar Lepcha development, see Benedict 1943.

interpretation of the **s** + **r** sequence. If anything, the WB form seems rather to favor Benedict's first idea that the ***s-** was prefixal, and therefore "optional" – the immediate ancestor of **phyam** would then be ***p-yam** < ****p-ram**.

In any case, the whole issue of "prefix" vs. "cluster" seems rather artificial in this instance. The sequence **s** + **r** is relatively hard to pronounce,[3] and it would not be surprising if it could not be unambiguously reconstructed morpheme-internally in the proto-language. Even the Mizo and Lepcha forms are not conclusive evidence for the proto-cluster theory. We can imagine a cyclical scenario: once the troublesome sequence of prefixal ***s-** plus root-initial ***r-** was eliminated (in Mizo via devoicing of the liquid, and in Lepcha by palatalizing it), the way was left open for *reprefixation* by the animal prefix, which has remained quite productive in these two languages:

> 1st prefixation: PTB ***sya** 'animal' + ***ram** 'otter' > pre-Mizo ***sa-ram** > **hram**;
> 2nd prefixation: Mizo **sa-** 'animal prefix' + **hram** 'otter' > **sa-hram**.[4]

The process whereby a full root-morpheme in a compound word gets reduced to a prefix has been called "prefixization" or the "compounding-prefixation cycle" (Matisoff 1978a:58–72, 2003:153–156). The closeness of bonding between the elements is not susceptible of precise quantification – the process represents an unstable continuum of phono-semantic interdependency.

The other forms presented in STC #438 are even less capable of proving that the word for 'otter' began with an intrinsic cluster. At first glance the Barish forms (Garo **matram**, Dimasa **matham**) look as if they might descend from ***m-sram**, but this is an illusion. According to Weidert (1987:122), the syllable **mat-** is a separate morpheme meaning 'squirrel; animal'.[5]

Most of the other forms in STC #438 are transcribed in the dictionaries from which they were taken with a vowel between the **s-** and the following **-r-**, which underlines the "non-clusterial" nature of the consonant sequence. This vowel is usually schwa (Nung **səram**, Jg. **shăram**, Lepcha **săryom**[6]), but sometimes it has a non-central quality even though it is still presumably not a fully-stressed segment (Miri **si-ram**, Mikir **serim**).

That leaves three more forms in set #438 that require some comment: WT **sram**, Maru **xrɛn**, and Phunoi **sam**.

Tibetan orthography never bothers to indicate the schwa-vocalism that must have occurred automatically between its numerous prefixes and their following root-initials (e.g. **bźi** 'four'

---

[3] In Indo-European languages this unstable combination is often broken up by an epenthetic stop (compare Sanskrit *sravati* 'flow' with its Russian and English cognates, *ostrov* 'island' and *stream*, respectively).

[4] This is analogous to the new English compound *house-husband*. The *hus-* of *husband* is itself a reduced form of *house*, though its morphemic identity has been lost for the average speaker, leaving the way open for a re-compounding with the same element.

[5] In other Barish languages cited by Weidert (*ibid.*), this same morpheme recurs in such a way as to conceal the original root-initial ***r-**: Meche **moṭam** (with the retroflex **ṭ** reflecting the **-t** + r- sequence), and Boro **motamʔ** (where the **r-** has disappeared altogether).

[6] As hypothesized above, the **să-** in the Lepcha form is actually a reprefixation, since the earlier ***s-** prefix got swallowed up after palatalizing the root-initial. In Mizo, judging from the orthography, the animal morpheme retained its full low central vowel /-a/, so that these animal names are more like *compounds* than like prefix + root combinations (e.g. **sa-khi** 'barking deer', **sa-kor** 'horse', **sa-phu** 'pangolin', **sa-zu** 'rat').

[bəźi], **brgyad** 'eight' [brəgyat], etc.). In this respect **sram** is no different from all the other WT animal names that begin with **s-** (e.g. **sbal-pa** 'frog', **sdig-pa** 'scorpion', **spre** 'monkey').

The Maru (Burmish) form is rather different. Like the other Lolo-Burmese [LB] languages, Maru preserves older TB prefixes only sporadically, when the phonological environment is particularly favorable, typically when the following root-initial is a 'weak' or 'preemptible' resonant (**r- l- y- w-**) or nasal.[7] In this case, the backing of the fricative ***s** to **x** evidently allowed the */s + r/ sequence to become pronounceable as /xr/, presumably in close juncture with no schwa intervening, so that both the prefix and the root-initial survived as separate segments. In other words, even though synchronically Maru /xr/ may be a cluster, it descends historically from a */prefix + root-initial/ sequence.

The Phunoi (S. Loloish) form **sam** seems to represent a clear case of what I have called *prefix preemption*[8] – the old prefix found itself before a 'weak' root-initial consonant and drove it out, so that only the prefix has left a distinct segmental reflex (i.e. the former prefix has evolved into the synchronic root-initial): ***s-ram** > **s-[]am** > **sam**. Other Loloish languages have treated the otter similarly, as in the second syllables of Lahu **ɣ̀ɪ-šo-lo** and Akha **uiˇshmˇ**.[9] (An unprefixed PLB form ***ram** would have developed into Lh. **ɣo**, not **šo**.)

Already it is clear that the TB languages show extreme variation and unpredictability in the behavior of prefixes before non-stop initials. Before zeroing in on our main concern – an explanation for the labial onset of WB **phyam** – let us quickly list the words for 'otter' in about 20 other TB languages that are not cited in STC #438. As we shall see, the treatment accorded to the */s + r/ combination in a given language – i.e. whether the prefix was simply dropped, or retained (with or without a cushioning vowel), or fused into a single segment with the root-initial, or fused with subsequent reprefixation, or substituted for by another prefix, or whether it "preempted" the initial itself – is quite independent of the subgroup of TB to which the particular language belongs.

I.2   **Karenic**

The Sgaw Karen form **shɣɔ** 'otter',[10] with an aspirated sibilant followed by ɣ (< ***r**), is practically the mirror-image of Maru **xrɛn** (above). Both languages succeed in breaking up the undesirable combination */s + r/ by means of backing and fricativization – Maru did it by backing and spirantizing the *prefix* (***s-** > **x-**), while Sgaw did it by backing and spirantizing the *root-initial* (***-r** > **-ɣ**). In both cases new heterorganic (velar-apical or apicovelar) clusters were created – more easily pronounceable in close juncture than a sequence of apical sibilant plus apical liquid.

---

[7] Cf. e.g. Maru **bìt** 'four' < PTB ***b-ləy**.

[8] See Matisoff 1972:275, 1975:166, 1978a:53, 1979:24 & 2003:95, 153. Other well-known examples of this phenomenon include Akha **myɔ̀q** 'lick' < ***m-lyak**, Lahu **sɪ̄** 'seven' < ***s-nit**, Lh. **nɪ̄** 'penis' < ***m-ləy**, etc.

[9] The first syllables of both these forms (Lh. **ɣ̀ɪ**, Ak. **uiˇ**) mean 'water' (cognate to WB **re** < PLB ***rəy¹**). Contra my note 304 in STC (p. 108), the third syllable of the Lahu form does *not* mean 'big'; the Lahu morpheme 'big' is under the high-rising tone /ló/, and is a loan < Tai (cf. Si. *lŭaŋ*). (Both Lh. **-o** and the Ak. syllabic nasal **-m** are the regular reflexes of PLB ***-am**.)

[10] This form is cited in the Karenic section of STC, p. 140. According to the data in Jones (1961:132–133), this root occurs only in the Sgaw dialects and Palaychi, but not in Pho or Pa-O.

To WT **sram** and Lepcha **sǎryom** (< *sǎ-s-ram) discussed above, we may add 3 forms from TB languages of Nepal: Tamang **chu-rɑm** (Mazaudon 1993–1994), Limbu **haːm-ba** (van Driem 1987), and Chepang **karyamʔ-jaʔ** (Weidert 1987). The Tamang form is a compound, with the first syllable **chu** meaning 'water',[11] and the second syllable descending from the unprefixed root *-ram.

The initial **h-** of Limbu **haːm-ba** (the **-ba** comes from a widespread "bulk-providing" noun-suffix in TB – cf. WT **sbal-pa** 'frog', **sdig-pa** 'scorpion', Section I.1) is apparently a "fusional reflex" of *sr. This is a typical development in TB, occurring also e.g. in Loloish:

$$
\begin{array}{ccccc}
\text{PTB} & & \text{PLoloish} & & \\
\text{*s-l, *s-y, *s-r} & & \text{*hl, *hy, *hr} & & \\
\text{*ʔ-l, *ʔ-y, *ʔ-r} & > & \text{*ʔl, *ʔy, *ʔr} & > & \text{Lh. } \mathbf{h} \ ^{12}
\end{array}
$$

Chepang **karyamʔ-jaʔ** has a different prefix altogether than the *s- we have encountered up to now. It seems to be the same "velar animal prefix" that has been noted long ago by Tibeto-Burmanists, and surmised to be of Mon-Khmer origin,[13] occurring in such animal names as *k-la 'tiger', *k-laŋ 'eagle, vulture, bird of prey', *k-roŋ 'wildcat', *k-rak 'chicken', *k-r-wat 'leech', *k-r-wak 'rat', *k-m-raŋ 'horse'. Note that these examples also all have liquids after the *prefix!

At least three interpretations of this variability in prefix are possible. (a) Either Chepang innovated by substituting a new *kǎ- prefix in place of the older *sǎ- prefix; or (b) even at the PTB stage there was prefixal variation with this root, so we should set up the etymon as *s-ram ⪤ *k-ram (equivalently $\left.{}^{*s-}_{*k-}\right\}$ ram); or else (c) Chepang (like Lepcha and Mizo) has a reprefixed form, deriving from *k-s-ram. Perhaps Chepang, like Lepcha, developed secondary palatalization from prefixed **s-**, thus accounting for the **-y-** in **karyamʔ**.[14]) The first interpretation might be called "diachronically paradigmatic" – i.e. it supposes a substitution of one thing by another through time. The second one is "synchronically paradigmatic" – i.e. it assumes the possibility of substitution of one thing by another at the same time. The third interpretation is (diachronically) syntagmatic – i.e. it supposes that both prefixes co-occurred before the same root in a given language, even though the older prefix had been reduced or modified before the younger one was tacked on. Chepang **karyamʔ** < *kǎ-s-ram would then be analogous to Mizo or Lepcha *sǎ-s-ram.[15]

---

[11] Lexosemantically this is the same formation (WATER + OTTER) as Lahu **ɣɪ̀-šo(-lo)** and Akha **uiˇshmˇ** (above, Section I.1), though the Tamang morpheme for 'water' reflects a different TB etymon *tsyu (cf. WT **čʼu**) than PLB *rəy¹. For *tsyu, see TBRS #161:229–230.

[12] Exs: PTB *s-la 'god; spirit; beautiful' > PLB **hla³** > Lh. **ha**
            PLB *hya¹ 'swidden; mountain field' > Lh. **hɛ**
            PTB *b-g-ryat 'eight' > PLB *ʔrit > Lh. **hí**

[13] See STC:107 and Matisoff 1969:190–199, 1973:n.37, 2003:138. Smith (1975) discussed this prefix with respect to several MK languages of Vietnam. This prefix is also relevant to the PTB etymon *s-kywal discussed in Part II, below.

[14] Not enough is yet known about Chepang historical phonology to be sure. The final glottal stop represents a phonation type (creaky voice) in Weidert's transcription, and is perhaps itself an evolutionary by-product of the **s-** prefix.

[15] The earnest reader by now also sees analogies with Benedict's hypothetical pre-WB form *p-sram, not so different after all from *p-s-ram! But for the source of this **p-** we must wait for the discussion below.

This third interpretation is somewhat reinforced by two forms from Kaduic languages cited in Luce 1986: Gănan **kɔ¹hɑm⁴** and Kădu (=Kantū) **ku²hɑm²**. Although little is known about the phonological history of the obscure Kaduic (or "Luish" group),[16] it seems likely that the **h** in these two forms (as in Limbu **haːm-ba**, above) also derives from ***s-r-**, so that here too we might guess at a prototype like ***kV-s-ram**.[17]

Luce also gives two interesting forms from dialects of Sak, another Kaduic language spoken in the Chittagong Hills of W. Burma and Bangladesh: Bawtala **p'æ̃⁴** and Dodem **pɛ̃¹**. These look very much like WB **phyam** (now pronounced **phyã** in the standard Rangoon dialect), and might well turn out to be loans from a Burmese dialect (perhaps Arakanese).

I.4   **Kamarupan**[18]

Weidert [W] and Luce [L] provide words for 'otter' in a number of TB languages of Northeast India.

Abor-Miri-Dafla (= "North Assam group" or "Tani" [Sun 1993]): Apatani **¹sɯ ²riŋ** (W), Gallong **ˆhoram** (W) (cf. also Miri **si-ram** and Mikir **serim** [STC #438]).

Barish (= Bodo-Garo): Chang **lám**,[19] Khiamngan **¹yam**, Meche **moṭam**, Garo **mat-ram**, Boro **motam?** (all W). (STC #438 cites this Garo form, and also Dimasa **matham** [see above, Section I.1].)

Kuki-Chin-Naga: Zemei **¹he ⁴ram**, Rongmei **a-ram**, Angami **²khuo ⁵rha**, Lotha **¹de ³ran**, Ao **³ši ¹m** (all W); Khumi (Ahraing dial.) **săham¹**, (Awa dial.) **tsă xã⁴**, Haka (= Lai) **s'ai⁵ hrɛm²**, Womatu **mˋʃram¹** (all L).[20]

Many of these forms reflect the ***sV-** prefix: Apatani **¹sɯ ²riŋ**, Ao **³ši ¹m** (here the root is reduced to a simple syllabic nasal), Khumi **săham¹** and **tsă xã⁴**; and maybe also Gallong **ˆhoram** and Zemei **¹he ⁴ram**. Haka **s'ai⁵ hrɛm²** (like Mizo **sa-hram** and perhaps the Khumi forms) looks like it has undergone reprefixation, deriving from ***sV-s-ram**.

Angami **²khuo ⁵rha** looks very much like the Kaduish forms cited above, Gănan **kɔ¹ ham⁴** and Kădu **ku² ham²**. If these words are not actually compounds, they could also represent reprefixed formations, deriving from ***kV-s-ram**.

The syllabic nasal that begins the Womatu form **mˋʃram¹** might be a reduction of the same morpheme **mat** 'squirrel; animal', that Weidert claims for the Meche, Garo, Boro, and Dimasa forms (see above, Section I.1); yet in the case of Womatu there is a double prefix (< ***m[at]-s-ram**).

---

[16] See STC, p. 5.

[17] It is of course also possible that these forms are compounds, and that the initial syllables **kɔ¹-** and **ku²-** have a separate lexical meaning (like 'water')! See also the Angami form **²khuo ⁵rha** cited in below in Section I.4.

[18] I have adopted this neutral geographical term (from *Kāmarūpa*, the old Sanskrit word for the Assam region), as a cover term for the TB languages of NE India, whose precise subgrouping remains uncertain. See GSTC, n. 36. For controversy surrounding this term, see Burling 1999 and my reply (Matisoff 1999).

[19] Weidert classifies Chang as Barish, though other writers have included it in the "Northern Naga" group. See Marrison 1967, French 1983.

[20] Mru **tăklui²** looks quite unrelated to our ***-ram** etymon. It is conceivably cognate to the Jingpho word **məjoi** (Hanson 1906/1954:649) that only occurs coupled with the usual word for 'otter' in poetic style (**shăram-măjoi**).

The first element of Lotha **¹de ³ran** remains a mystery.

Finally, Chang **lám**, Khiamngan **¹yam**, and Rongmei **a-ram** are of particular interest, since they show no trace of the ***sV-** prefix at all. Either they never had it, or they lost it without trace.

I.5 **Chinese**

Before coming around at last to an explanation for WB **phyam**, let us see what we can find on the Chinese side.

The Chinese root-word for 'otter' is 獺 (Mand. *tǎ* or *tà*). This is now a bound morpheme in Mandarin, where the name of the fresh-water species of the animal is expressed by a disyllabic compound with the first syllable meaning 'water', *shuǐ-tǎ* 水獺.[21]

The initial of the 'otter' morpheme in Old Chinese is especially interesting, as can be inferred from the phonetic series to which the character belongs. All the other words in the series are reconstructed with the lateral initial ***l-**, and are still pronounced with *l-* in Mandarin:

GSR #272		OC		MC	gloss	Mand.	Mand. tone
a-d	剌	*lât	/	lât	'wicked; slash, cut in two'	là	4
e	賴	*lâd	/	lâi	'gain, get hold of; lean on, depend on, rely on; gain, advantage'	lài	4
f	瀨	*lâd	/	lâi	'shallow water flowing rapidly over sand'	lài	4
g	籟	*lâd	/	lâi	'tube; musical pipe consisting of three reeds; rustling of the wind'	lài	4
h	藾	*lâd	/ and	lâi	'artemisia'	là~lài	4
		*lât	/	lât			

(To these we may add 癩 'leprosy' [Mand. *lài*] [not in #272].)

The only word in this series to have a dental stop in its initial is 'otter' itself. In order to account for the same phonetic being used here as for the words in *l-*, Karlgren reconstructs an initial cluster consisting of an aspirated dental stop with lateral release (a sort of lateral affricate):

---

[21] Lexically this is exactly parallel to TB compounds like Lahu **ɣɪ̀-šo(-lo)** and Akha **uiˇshmˇ** (above).

GSR #272		OC		MC	gloss	Mand.	Mand. tone
i	獺	*t'lât	/	t'ât	'otter'	tă ~ tà	3~4
			and			chă	
		*t'lat	/	t'at			

This reconstruction agrees so beautifully with the ***s-r** onset of PTB ***s-ram** that one feels prepared to go to any lengths to demonstrate the cognacy of 獺 to our TB etymon! Schuessler (2009:235) reconstructs ***rhât** – also not a bad comparison with PTB ***s-r**.

The nuclear vowel (OC ***-â-** / PTB ***-a-**) is also a perfect fit. However, there is a big problem with the syllable-final consonant: the Chinese word ended in a dental stop **-t** (cf. Cantonese *shuí-ch'aåt*[22] and Sino-Japanese *datsu*[23]), while the TB root ends in **-m**. The difficulty is twofold: the stop vs. nasal discrepancy, and the dental vs. labial point of articulation:

```
PTB   *s-r   a   m
OC    *t'-l   â   t
```

Yet neither of these problems seems insuperable. There must have been something variable about the OC final in any case – this is a series where Karlgren reconstructs ***-d ~ *-t** variation, implying that the pre-OC final was more sonorant-like than stop-like. Furthermore, another character in the same phonetic series (that does not appear in GSR) definitely has a reading with a nasal final: 懶 (sometimes written 嬾) 'lazy' (Cant. *lăan*, Mand. *lăn*).[24]

We should also note that all the words in this series reconstructed with OC ***-d** are under the 去声 (Mand. Tone 4). According to the well-known theory of Haudricourt (1954), subsequently adopted by many other scholars, the 去声 had its origin in an ***-s** suffix. Perhaps then we should reconstruct 'otter' at the PST level as

***s-ram-s.**

We could then explain the Chinese shift from final labial to dental in terms of assimilation to the suffix: PST ***s-ram-s** > pre-Old Chinese ***t'lans** or ***hrans**.[25, 26] Killing two birds with the same "**-s** tone", we could go on to assume that the pre-Archaic final cluster ***-ns** was 'hardened' into a **-t**, probably via epenthesis of the homorganic stop with subsequent loss of the sibilant:

$$\text{*t'lans} > \text{*t'lan}^{\text{t}}\text{s} > \text{*t'lãts} > \text{OC *t'lât}$$

---

[22] See Cowles 1965, p. 941.

[23] Similar stopped finals for GSR #272a 'wicked' and #272f 'shallow water' are attested by the Sino-Japanese readings *ratsu*. The other characters in #272 are pronounced *rai* in Sino-Japanese.

[24] I believe this is a graphic elaboration of the same word as GSR #272a-d 剌 'wicked', since 懶 has an alternate Mand. reading *lài* with this very meaning. Laziness and wickedness are related notions. Cf. the old jingle about the proper number of hours one should sleep at night: "Nature requires five, Custom gives seven, Laziness takes nine, And Wickedness eleven." Note the double reading *rai ~ ran* 'lazy' in Sino-Japanese.

[25] A similar suggestion was made in VSTB (Matisoff 1978a:117) with respect to the Chinese word 肺 (Mand. *fèi*) 'lung', reconstructed by Karlgren as OC ***p'įwăd** (GSR #501g), so that it could be brought into relationship with my TB reconstruction ***p-wap � *s-wap**.

[26] The hesitation between tones 3 and 4 in the modern Mandarin form may be another indication of some complication in syllable-final position.

If this line of reasoning holds water, this etymon would be a striking instance of the preservation of the sibilant animal-prefix in Chinese.

I.6   Mon-Khmer

By now the suspense concerning WB **phyam** must be nearly unbearable.

The idea of a possible Mon-Khmer [MK] origin for this mysterious labial initial first occurred to me while Gérard Diffloth and I were idly browsing through his masterly *Proto-Waic*[27] one day in 1987. Diffloth's PWaic reconstruction for 'otter' is ***phiʔ**, but it turns out that this is only one tendril of a general MK root attested in at least 10 branches of the family. In April 1988, Diffloth kindly supplied me with all of the following forms:[28]

Khmer: **bhee**
Pearic: Chong **pḥe̱e**
Katuic: Bru **pḥɛɛ**; Pacoh **pihəy**, **pisay**
Bahnaric: Bahnar **pḥɛy**; Brao **phay**; Rongao **pḥɛɛi**; Hrê **bəhe**; Cua **phai**; Chrau **phii**
Monic: ***pḥɛɛʔ**; Nyah Kur ***pḥɛ̱ɛʔ**
Aslian: Semelai **bəheʔ**
Viet-Muong: Vietnamese **rái**; Thavɯng **psəːɜ**
Khasi: **kəsiʔ**
Khmu: **veʔ**
Palaungic: ***pheeʔ**; PWaic ***phiʔ**; AngKuic **psiʔ**

Shorto (2006:#260) reconstructs Proto-MK ***bheʔ**, and cites forms from several Chamic (Austronesian) languages that have borrowed this word from MK (e.g. Cham **phày**, Jarai **pəhaːi**, Röglai **bəhaːy**).

It therefore seems highly likely that WB **phyam** is a reduction of a two-syllable prototype like ***bheʔ-ram**, with its first syllable borrowed from MK and tacked onto the native TB second syllable. This phenomenon of "prefixization", already mentioned above (Section I.1) with respect to the reduction of ***sya** 'animal' to the "animal prefix **s-**", has demonstrably occurred in a number of other former compounds, including:

> 'ant' WB **pərwak** < ***bəw** 'insect' + ***-rwak** 'ant' (HPTB:154)
> 'lungs' Mizo **tsywap** < ***tsəy** 'lung' + ***-wap** 'spongy' (VSTB:115)
> 'son-in-law' WB **səmak** < ***za** 'son' + ***mak** 'son-in-law' (HPTB:154)

As a sort of postscript, it is interesting to observe that a similar reductional process has occurred in the Latin word *lutra* 'otter' (> e.g. French *loutre*, Italian *lontra*), where the initial *l-* does not descend from the IE root ***wed-** ⪦ ***wod-** ⪦ ***ud-** 'water', but perhaps represents a fusion of the demonstrative *ille* which developed into the Vulgar Latin definite article.[29]

<center>***</center>

---

[27] LTBA 5.2:1–182.
[28] He jotted them down from memory on a scrap of paper during dinner at a conference in Pittsburgh.
[29] See Matisoff 1994:52–53.

PART II: *JACKAL in Sino-Tibetan and Indo-European*

II.1 Introduction

The English word *jackal*[30] derives ultimately from Sanskrit **śṛgāla-**, via a tortuous history: < French **chacal** < Turkish **chacāl** < Persian **shaghāl** < Middle Indic **shagāl** < Sanskrit **śṛgāla-.** This word is found only in Indo-Aryan,[31] and nowhere else in Indo-European, so that specialists agree that it is a loan from some non-IE source.

I would like to claim that an excellent candidate for this non-IE source is a PTB root reconstructible as ***s-k-ywal** 'wild dog / jackal / dhole / wolf'.[32]

II.2 Loloish

The Loloish languages have inherited the latter part of the PTB etymon, pointing to a Proto-Lolo-Burmese form ***wan¹** (the superscript numeral refers to the PLB *tone):

> Lahu **vè** 'dhole' (*Cuon javanicus*); Akha **xhà-jɛ́** (Hansson), **k'aˌyeȟ** (Lewis 1986:248) 'dhole/Asian wild dog' (*Cuon alpinus*);[33] Jino **ø⁴** 豺狗 'jackal' (Gai Xingzhi 1986:67).

The Burmese word for *jackal* is a curious compound, WB **wam-pu'-lwe**, literally "flute-bear" (**wak-wam** 'bear', **pu'-lwe** 'flute'), perhaps because of the plaintive sound it makes at night.

If the second syllable of Bisu **hɔŋ-tšón** is cognate to the other forms, it might point to a prototype with a velar-plus-y cluster that became an affricate: ***kywan** > **tšón**. It is more likely that the second syllable is cognate, and not the first, since the regular Bisu reflex of PLB Tone *1 is the high tone / ´ /.

The first syllable of the Akha form is not part of the root, but rather the celebrated "velar animal prefix" (see Section I.3) that occurs frequently with animal names in Akha, e.g. **k'aˌdzeȟ** 'hawk', **k'aˌguˌ** 'pigeon', **k'aˌhmȟ** 'Asiatic black bear', **k'aˌjaȟ** 'sharp-tailed munia', **k'aˌlaˌ** 'tiger', **k'aˌpaˌ** 'frog', **k'aˌtsehˆ** 'sambar deer', **k'aˌzuiˌ** 'leopard', etc. (Lewis 1986:243–248). Thus the Akha word for 'dhole' may be derived from a prototype like ***k-wan**.

The PTB rhyme ***-al** has either disappeared without trace in Lolo-Burmese (e.g. 'frog' PTB ***s-bal** > PLB ***ʔba²**), or else merged with ***-an**, as in this case.

---

[30] The OED (1971:1497) remarks that "the English word was formerly (as still in some dialects) stressed on the second syllable; the current form, and the obsolete *jack-call*, show association with the proper name *Jack*, and names of animals containing it." We might add that the animal's propensity for howling at night may have encouraged an association with the verb *call*.

[31] Medieval and modern Indo-Aryan forms include Prakrit **si(g)āla-**, Kashmiri **śāl**, Nepali **siyāl**, Bengali **siyāl**, Assamese **xiyāl**. In recent times several TB languages of India and the Himalayas have reborrowed this word from Indo-Aryan. These *back-loans* include Limbu **syaʔl** (< Nepali), Kanauri **shyáles ~ shyálí** (prob. < Kashmir), Mizo **shihal** (< Assamese), and Mikir **hijai** (prob. also < Assamese).

[32] This reconstruction is somewhat elaborated from the form *kywal presented in Matisoff 1985:#17 and 2003 (HPTB:261, 407, 423, 449). The reasons for reconstruction of final ***-l** and prefixal ***s-** will be gradually justified below.

[33] Lewis remarks that this word is used to translate 'wolf' in the Bible.

II.3 Jingpho

A key form for our PTB reconstruction is Jingpho **chyăhkyawn** 'a fox, wolf or wild dog' (Hanson 1906/1954:96), **chyahkyon** [tʃă³³khyon³³] 狼 'wolf' (Dai Qingxia et al. 1983:82), **čəkhyōn** (HPTB:407).

PTB final ***-r** and ***-l** both regularly become **-n** in Jingpho, e.g. 'star' PTB ***s-kar** > WT **skar-ma**, Jg. **šəgān**; 'body hair' PTB ***s-mul** > Mizo **hmul**, Jg. **mūn** (HPTB:386).

The Jingpho vowel **-o-**, like Written Tibetan **-o-**, frequently reflects PTB ***-wa-** (a sequence which is preserved as such in Written Burmese), e.g.:

	PTB	WT	Jg.	WB
'be free/loose' (v.i.)	***g-lwat**	glod-pa	lòt	lwat � kywat
'free/release' (v.t.)	***s-lwat**	hlod-pa	šəlòt	hlwat ⋉ khywat

Hanson (94, 221) characterizes the prefix **chyă-** /čə-/ as "a preformative, mostly used in the formation of (abstract or) verbal nouns, where the transitive form of the verb takes **jă-** or **shă-**". As an example he gives **htùm** 'be ended', **jăthùm** 'bring to a close', **chyăhtùm** 'an end' (tone marks added). There is, however, no verb "**hkyawn**" in the language. The causative prefixes **jă-** and **shă-** are also irrelevant in this case.[34]

I conclude that in this case, Jg. **chyă-** /čə-/ is a reflex of the TB animal prefix ***s-**, ultimately derived from the independent noun ***sya** 'flesh/game animal/animal' (see Section I.1 above). The Jingpho form for 'wild dog' thus justifies the reconstruction ***s-kywal**.

II.4 Northern Naga *C-khyual

More evidence for the so far hardly proven final ***-l** in our reconstruction is provided by Proto-Northern Naga, as reconstructed in W. T. French (1983II:580), on the basis of six languages: Moshang, Yogli, Nocte, Wancho, Konyak, Phom, and Chang.[35]

These languages do not preserve final consonants very well, but French has managed to figure out their reflexes of the voiced apical codas ***-n**, ***-r**, and ***-l**:[36]

PNN	Moshang	Yogli	Nocte	Wancho	Chang	Konyak	Phom
*-n	-n	-n	-n	-n	-n	-n	-n
*-r	-r	-l	-n	-n	-n	-Ø	-Ø
*-l	-l	-l	-n	-n	-Ø	-Ø	-Ø

---

[34] The distribution of these quite productive prefixes is phonologically determined: **jă-** occurs before verb roots with initial sibilants or affricates, while **shă-** appears elsewhere. This morpheme is easily seen to be a reflex of the well-attested PTB causative prefix ***s-**.

[35] These languages, sometimes referred to as the "Konyak" group, have long been grouped with Bodo-Garo and Jingpho, in what the *Linguistic Survey of India* referred to as *Bodo-Naga-Kachin* (Grierson & Konow 1903–1928, Vol. III, Part 2), an idea resurrected by R. Burling (1983) under the name of the *Sal* languages. In any case these "Northern Naga" languages do not seem particularly closely related to the many other languages called "Naga".

[36] The following chart is based on the data presented in prose in W.T. French, Vol. I:347.

Unfortunately no forms meaning 'wild dog' are available from Moshang or Yogli, which preserve final liquids as such. In fact all French had to go on were forms from the following four languages: Wancho **šan**; Konyak **šo**; Phom **šo**; Chang **šo** 'wolf' ≽ **šuo** 'wild dog'. However this data is already sufficient (zeroes in Chang, Konyak, and Phom) to rule out *-**n** and *-**r**, leaving only *-**l** as a possibility. This reconstruction has apparently been strongly reinforced by a Nishing (Tani group) cognate **yal** 'renard' (Jacquesson 1998:102).[37]

The proto-rhyme of this etymon does not seem to have been simply *-**al** or *-**a:l**, which would have given Konyak -**e**, Phom -**a**, and Chang -**a** (French I:392, 403). However, the back rounded vowel in Konyak, Phom, and Chang **šo**, as well as the medial -**u**- in the Chang allofam **šuo** 'wild dog' makes it plausible to reconstruct the rhyme with a labial element, *-**ual** (French I:426).

There are no initial consonant clusters in modern N. Naga languages, leading French (I:187) to admit that "in this area we are particularly dependent upon external evidence as a guide to the shape of the reconstructed elements". PNN *-**khy**- is reconstructed when Yogli, Moshang, and Nocte have **kh**-, Wancho and Chang have **k**-, and Konyak has **š**- (194), while the former presence of a now lost prefix before this initial, symbolized by "C-", led to Phom, Wancho, and Chang **š**- (195), as in the set for 'wild dog', thus leading French to reconstruct initial *-**C-khy**-, giving us PNN *-**C-khyual**.[38] It is reasonable to guess that this proto-prefix was actually the "animal prefix" *-**s**-.

The word for 'jackal' in the Central Chin language Mizo (Lushai) is **kâwl-ui** (**ui** 'dog' < PTB *-**kʷəy**), but the first syllable, phonemically /kɒɒl/, has nothing to do with our etymon *-**s-k-ywal**, since **kâwl** means 'the Burmese; a Burman', and the literal meaning of the compound is "Burmese dog" (Lorrain 1940:235). No doubt the jackal is so called as a contemptuous way of referring to its perceived undesirable qualities.

II.5   Tibetan and Chinese

The Written Tibetan word **wa** 'fox' is quite unusual, since it is practically the only word in the language with initial **w**-. Jäschke (1881/1958:470) attributes its strange pronunciation to onomatopoeia ("the name corresponding to the sound of barking").[39] However, Benedict reconstructs the word as PTB *-**gwa**, pointing out that an Amdo dialect in Gansu has the form **gaa** 'fox', and that closely related Himalayish languages have similar forms: Chamba Lahuli **gŭa**, Bunan **goa-nu** ~ **gwa-nu**.[40] Benedict goes on to compare the Tibetan form with Chinese 狐 (OC *-**g'wo** [GSR #41i], Mand. *hú*) 'fox' (STC:166, 186), since it shows the regular development of PST *-**a** > OC *-**o** after velar initials.[41]

---

[37] Nishing evidently dropped (or never acquired) the velar component of the initial, in a way very reminiscent of Loloish (above, Section II.2), perhaps indicating that our etymon should be reconstructed as *-**s-k-ywal**, with both the *-**s**- and *-**k**- animal prefixes. Of course it is also quite possible that this Nishing form is a backloan from an Indo-Aryan source. (See note 31.)

[38] The reflexes of PNN *-**sy**- and *-**C-sy**- (e.g. Phom -**Ø** and Chang **s**-) are quite different from those of the velar clusters (I:229).

[39] Jäschke (*loc. cit.*) remarks that "the fox is the riding-beast of the goblins; whenever his barking is heard, it is in consequence of his receiving lashes from his rider". The specific meaning 'jackal' is conveyed by the compound **wa-spyaŋ** (glossed "jackall"). The compound for 'wolf' is **spyaŋ-ki**, where the second element is doubtless a variant of **khyi** 'dog'; cf. also the Western Tibetan form **k'yi-čáṅ** 'jackal', lit. "dog-wolf" (Jäschke:332–333).

[40] STC: n.111, p. 34. The Amdo form is cited from Przhevalski 1880.

[41] See also HPTB:167, 173.

It is at least possible that PST/PTB **gwa** 'fox' has some sort of allofamic relationship with our etymon ***s-k-ywal** 'wild dog'.

<p align="center">***</p>

We conclude that the most likely explanation of the Sanskrit word for 'jackal' is that it is an ancient borrowing from some Tibeto-Burman language, transmitted from the TB area via an unknown route. Perhaps it is time to add JACKAL to our list of Eurasian animalian *Wanderwörter,* along with HORSE and BEE/HONEY.

One final thought: Recent research has demonstrated that the dog was first domesticated some 11,000 to 14,000 years ago in Southern China, where dogs still show greater genetic diversity than anywhere else.[42] Might it not be possible that PIE ***kwon** 'dog' is an ancient loanword from Sino-Tibetan? On the Tibeto-Burman side we not only have ***s-k-ywal** 'jackal', but also ***kwəy** 'dog'. On the Chinese side, I have already tried to group 犬 'dog' (Mand. *quǎn*; OC ***k'iwən** [GSR #479a-d], ***khwîn** [Schuessler 2007:437]) and 狗 'dog' (Mand. *gǒu*; OC ***ku** [GSR #108d], ***kəu** [Schuessler 2007:257]), in a word-family reconstructed as ***kʷəy-n**.[43] Yet ***s-k-ywal** also seems particularly close to OC ***k'iwən** or ***khwîn**, and I am now tempted to believe that all these ST forms are interrelated, constituting a word-family ***kʷəy-n ≍ *s-k-ywal**, and that the PIE etymon was borrowed from some allofamic variant of this ST word-family in prehistoric times.

---

[42] See the article "In taming dogs, humans may have sought a meal" in the New York Times, Sept. 8, 2009, p. D1.
[43] See HPTB:448.

Abbreviations

A �curlyeqprec B	A and B are co-allofams; A and B are members of the same word-family	**MK**	Mon-Khmer
Ak.	Akha	**OC**	Old Chinese
BSLP	*Bulletin de la Société de Linguistique de Paris* (Paris)	**OED**	Oxford English Dictionary
		OPWSTBL	Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics
Cant.	Cantonese	**P**	Proto-
CISTL	Kitamura et al. 1994	**PIE**	Proto-Indo-European
GSR	Karlgren 1957	**PLB**	Proto-Lolo-Burmese
GSTC	Matisoff 1985	**PMK**	Proto-Mon-Khmer
HPTB	Matisoff 2003	**PNN**	Proto-Northern-Naga
IA	Indo-Aryan	**PST**	Proto-Sino-Tibetan
ICSTLL	International Conference on Sino-Tibetan Languages and Linguistics	**PTB**	Proto-Tibeto-Burman
		Si.	Siamese
IE	Indo-European	**SiL**	*Studies in Linguistics* (Berkeley)
Jg.	Jingpho (= Kachin)	**ST**	Sino-Tibetan
L	Luce 1986	**STC**	Benedict 1972
LB	Lolo-Burmese	**TB**	Tibeto-Burman
Lh.	Lahu	**UCPL**	University of California Publications in Linguistics
LTBA	*Linguistics of the Tibeto-Burman Area* (Berkeley; Chico, CA; Melbourne)	**VSTB**	Matisoff 1978a
		W	Weidert 1987
Mand.	Mandarin Chinese	**WB**	Written Burmese
MC	Middle Chinese	**WT**	Written Tibetan

References

Benedict, Paul K. 1943. "Secondary infixation in Lepcha." *SiL* 1.19:1–2.

Benedict, Paul K. 1972. *Sino-Tibetan: a Conspectus* (Princeton-Cambridge Studies in Chinese Linguistics #2). James A. Matisoff, Contributing Editor. New York: Cambridge University Press. ("STC")

Burling, Robbins. 1983. "The *sal* languages." *LTBA* 7.2:1–32.

Burling, Robbins. 1999. "On Kamarupan." *LTBA* 22.2:169–171.

Cowles, Roy T. 1965. *The Cantonese speaker's dictionary*. Hong Kong: Hong Kong University Press.

Dai Qingxia 戴庆厦, et al. 1983. 景汉辞典 *Jǐng-hàn cídiǎn* [*Jingpho-Chinese dictionary*]. Kunming: Yunnan Nationalities Press.

Diffloth, Gérard. 1980. "The Wa languages." *LTBA* 5.2:1–182.

Driem, George van. 1987. *A grammar of Limbu* (Mouton Grammar Library #4). Berlin, New York, Amsterdam: Mouton de Gruyter.

French, Walter T. 1983. *Northern Naga: A Tibeto-Burman mesolanguage*. 2 volumes. Ph.D. Dissertation, City University of New York.

Gai Xingzhi 盖兴之. 1986. 基诺语简志 *Jīnuò-yǔ jiǎnzhì* [*Brief description of the Jinuo language*]. Beijing: Nationalities Press.

Grierson, Sir George Abraham & Sten Konow (eds.). 1903–1928. *Linguistic survey of India*. 13 volumes. Calcutta: Office of the Superintendent of Government Printing. Reprinted (1967, 1973), Delhi: Motilal Banarsidass.

Hanson, Ola. 1906. *A dictionary of the Kachin language*. Rangoon. Reprinted (1954, 1966), Rangoon: Baptist Board of Publications.

Hansson, Inga-Lill. 1989. "A comparison of Akha, Hani, Khatu and Pijo." *LTBA* 12.1:6–91.

Haudricourt, André-Georges. 1954. "Comment reconstruire le chinois archaïque." *Word* 10:351–361.

Haudricourt, André-Georges. 1966. "The limits and connections of Austroasiatic in the Northeast." In Norman H. Zide (ed.), *Studies in comparative Austroasiatic linguistics*, 44–56. The Hague: Mouton & Co.

Jacquesson, François. 1998. "L'évolution et la stratification du lexique: contribution à une théorie de l'évolution linguistique." *BSLP* 93.1:77–136.

Jäschke, Heinrich August. 1881. *A Tibetan-English dictionary, with special reference to the prevailing dialects*. London. Reprinted (1958), London: Routledge and Kegan Paul.

Jones, Robert B. 1961. *Karen linguistic studies: Description, comparison, and texts* (UCPL #25). Berkeley & Los Angeles: University of California Press.

Karlgren, Bernhard. 1957. *Grammata serica recensa*. Stockholm: Museum of Far Eastern Antiquities, Publication 29. ("GSR")

Kitamura Hajime 北村甫, Nishida Tatsuo 西田龍雄 & Nagano Yasuhiko 長野泰彥 (eds.). 1994. *Current issues in Sino-Tibetan linguistics*. Osaka: Organizing Committee of the 26ᵗʰ ICSTLL. ("CISTL")

Lewis, Paul W. 1986. *Akha-English-Thai dictionary*. Bangkok: Darnsutha Press.

Lorrain, J. Herbert. 1940. *Dictionary of the Lushai language* (Bibliotheca Indica #261). Calcutta: Royal Asiatic Society of Bengal. (Reprinted 1965, 1976).

Luce, Gordon H. 1986. *Phases of pre-Pagán Burma: Languages and history*. 2 volumes. Oxford & New York: Oxford University Press. ("L")

Marrison, Geoffrey E. 1967. *The classification of the Naga languages of northeast India.* 2 volumes. Ph.D. dissertation, School of Oriental and African Studies, University of London.

Matisoff, James A. 1969. "Lahu and Proto-Lolo-Burmese." In Alton L. Becker (ed.), *OPWSTBL*, Vol. I, 117–221. Ann Arbor: University of Michigan.

Matisoff, James A. 1972. "Tangkhul Naga and comparative Tibeto-Burman." *Tōnan Ajia Kenkyū* 10.2:1–13.

Matisoff, James A. 1973. "The Mon-Khmer substratum in Tibeto-Burman." Talk presented at the First International Conference on Comparative Austroasiatic Linguistics (Honolulu, 1973).

Matisoff, James A. 1975. "Benedict's Sino-Tibetan: A rejection of Miller's conspectus inspection." *LTBA* 2.1:155–172.

Matisoff, James A. 1978a. *Variational semantics in Tibeto-Burman: The 'organic' approach to linguistic comparison* (OPWSTBL #6). Philadelphia: Institute for the Study of Human Issues. ("VSTB")

Matisoff, James A. 1978b. *Mpi and Lolo-Burmese microlinguistics* (Monumenta Serindica #4). Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Matisoff, James A. 1979. "Problems and progress in Lolo-Burmese: Quo vadimus?" *LTBA* 4.2:11–43.

Matisoff, James A. 1985. "God and the Sino-Tibetan copula, with some good news concerning selected Tibeto-Burman rhymes." *Journal of Asian and African Studies* (Tokyo) 29:1–81. ("GSTC")

Matisoff, James A. 1994. "Regularity and variation in Sino-Tibetan." In *CISTL*, 36–58.

Matisoff, James A. 1999. "In defense of Kamarupan." *LTBA* 22.2:173–182.

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction* (UCPL #135). Berkeley & Los Angeles: University of California Press. ("HPTB")

Matisoff, James A. 2008. *The Tibeto-Burman reproductive system: Toward an etymological thesaurus* (UCPL #140). Berkeley & Los Angeles: University of California Press. ("TBRS")

Mazaudon, Martine. 1993–1994. *Problèmes de comparatisme et de reconstruction dans quelques langues de la famille tibéto-birmane*. Ph.D. dissertation, Université de la Sorbonne Nouvelle, Paris.

Oxford English Dictionary, compact edition. 1971. 2 volumes. Oxford: Oxford University Press. ("OED")

Przhevalski, Nikolai M. 1880. *Mongolie et pays des Tangoutes*, trans. by G. de Laurens. Paris: Librairie Hachette.

Purnell, Herbert C. 1970. *Toward a reconstruction of Proto-Miao-Yao*. Ph.D. Dissertation, Cornell University.

Schuessler, Axel. 2007. *ABC etymological dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.

Schuessler, Axel. 2009. *Minimal Old Chinese and Later Han Chinese: A companion to Grammata serica recensa (ABC Chinese dictionary)*. Honolulu: University of Hawai'i Press.

Shorto, Harry L. 2006. *A Mon-Khmer comparative dictionary*. Paul Sidwell, Editor. Canberra, Australia: Pacific Linguistics.

Smith, Kenneth D. 1975. "The velar animal prefix relic in Vietnam languages." *LTBA* 2.1:1–18.

Sun, Jackson Tianshin 孙天心 . 1993. *A historical-comparative study of the Tani (Mirish) branch of Tibeto-Burman*. Ph.D. Dissertation, University of California, Berkeley.

Weidert, Alfons K. 1987. *Tibeto-Burman tonology: A comparative account* (Current Issues in Linguistic Theory #54). Amsterdam & Philadelphia: John Benjamins. ("W")