

Word families, allofams, and the comparative method

Hannes A. Fellner (Vienna University) and
Nathan W. Hill (SOAS, University of London)

Linguists researching the Trans-Himalayan family do not have a self perception as working outside the mainstream of historical linguistics.¹ Nonetheless, ‘word families’ and ‘allofams’ are important elements in the theorizing of some Trans-Himalayanists (LaPolla 2003: 26, Schuessler 2007, *inter alia*) despite the absence of these terms in the wider discipline. Does the Trans-Himalayanist have a gem of methodological wisdom here that could benefit the Indo-Europeanist, or are do those phenomena treated as word families in Trans-Himalayan admit of superior analyses in more traditional terms? The evidence appears to point to the latter conclusion. We begin with a brief reminder of how ‘word family’ and ‘allofam’ made their appearance, and then turn to examine how Indo-Europeanists explain the word-family-like features of the languages they study. The bulk of the paper attempts to explain would-be Trans-Himalayan allofams in light of the techniques available within traditional historical linguistics. A concluding section considers the past and future of Neogrammarian methodology in Trans-Himalayan.

Bernhard Karlgren introduced the term ‘word family’ into East Asian linguistics, despite devoting only scant lines to its elaboration (1933: 9).² Wolfenden is more explicit; with a floral metaphor he contrasts Walter Simon's ‘horizontal’ (i.e. Neogrammarian) and Karlgren's ‘vertical’ method.³

In pursuing comparative studies of the vocabularies of the Sino-Tibetan languages we are today possessed of two methods of approach. The first of these, and the older, is that of setting up simple word equations from language to language; the second that of comparison by word families only, taking the family as our smallest operating unit. The first method passes from language to language lifting single words from each, without delving down in any way into what we might call the soil beneath them, so that we might, in fact, term such surface operations the “horizontal” method. We have, as it were, plucked a flower without looking to see from what bush we took it. The second method, on the other hand, seeks, in the first place not to set up equations between single words in two or more languages, but first of all to

¹ As do the Neo-linguist followers of Matteo Bartoli (1873-1946) or the Neo-idealist followers of Karl Vossler (1872–1949).

² Karlgren writes: “It is not allowable to identify Chinese 目 Arch. *m̄iôk* ‘eye’ with Tibetan *mig* ‘eye’ so long as we have not first established the word family to which *m̄iôk* belongs. Akin to *m̄iôk* is undoubtedly the word 眸 *m̄iôg* ‘pupil of the eye’ and it is just as likely that it is this *m̄iôg* which corresponds directly to Tibetan *mig*” (1933: 9). It may appear that in this passage Karlgren simply warns that the semantically most direct comparison may not be most genuine; English ‘town’ goes not with German *Dorf* ‘small town’ but *Zaun* ‘fence’. However, Wolfenden (1937: 625) understands Karlgren as saying that 目 *mjuwk* < *C.muk (14-24a) ‘eye’ and 眸 *mjuw* < *mu (13-77c) ‘pupil of the eye’, must somehow be compared to མིག *mig* together as a family. Direct comparison of Chi. 目 *mjuwk* < *C.muk (14-24a) ‘eye’ and Tib. མིག *mig*, albeit not allowable, has continued unabated, even among word family enthusiasts (e.g. Matisoff 2003: 327, Schuessler 2007: 393), but any connection between 目 *mjuwk* < *C.muk (14-24a) ‘eye’ and 眸 *mjuw* < *mu (13-77c) ‘pupil of the eye’ remains to be explained (cf. Schuessler 2007: 393). Observing that in early Chinese documents (e.g. Mencius) the word for ‘pupil’ was 眸子 rather than simple 眸, a reviewer suggests the etymology 眸子 *mu-tsə? < 目子 *C.muk-tsə? ‘child of the eye’.

³ Although Wolfenden thinks only a “fraction of the equations” which Simon proposes “will prove to be true” (1936: 401 note 1) many of Simon's comparisons remain in circulation (cf. e.g. Gong 1980, 1995).

gather the word families of each separate language, and only then, after we have gained a clearer view of the general background of the words composing them, to begin comparative work. This method, from the fact that we dig down into the soil, as it were, from which the individual words have sprung, we might, if we so wished, call the “vertical” method. We have then not only the flower but the actual bush on which it is growing. (Wolfenden 1938: 165-166).

As it happens Wolfenden’s specific proposals—meant to showcase the benefits of the ‘vertical’ word family approach over the ‘horizontal’ approach—have not stood the test of time.⁴ The more recent ‘allofam’ is a coinage of James A. Matisoff, who believes we “need a word to refer to the relationship among the various individual members of the same word-family” (1978: 16-17). He draws attention to variation in living languages to defend the reconstruction of variation in a proto-language.

[W]e must assume that the proto-language itself was awash with allofamic variation, both systematic and unsystematic. Why should a proto-language be any more monolithically invariant than any living language that we can observe with our own eyes? No language is ever perfectly regular at any stage of its history. (Matisoff 1978: 18-19)

Phonetic change does at times appear irregular and proto-languages, like living languages, do exhibit apparently inexplicable internal variation. Nevertheless, proposals such as Matisoff’s that Proto-Trans-Himalayan has 21 variant forms of ‘lung’ strain credulity (1978: 123).⁵ Are cases like this precedented in better studies language families?

In keeping with Matisoff identification of ‘doublets’ in Indo-European with ‘word families’ in Asian linguistics (1978: 18), a reviewer of an earlier version of this paper proposes that alternate ablaut grades and distinct root

⁴ Three examples suffice: (1) Wolfenden puts Tib. མཚུབ་མོ་ *mdzub-mo* ‘finger’ in a family with རྩུག་ (pres. རྩུག་ *hɔdzug*) ‘insert’ (1928: 278). Pointing to མཚུག་གུ་ *mdzug-gu* ‘finger’ as another form of the word, he writes “that there is little doubt that the form *mdzub-mo* is secondary” (1928: 278) and that it “owes its final -b to the influence of the following suffix” (1937: 627 note 2). The suffix -*gu* is an allomorph of the diminutive -*bu* that typically occurs after velars but can after open syllables (Uray 1952: 196). We suggest མཚུག་གུ་ *mdzug-gu* < མཚུབ་མོ་ *mdzub-mo* via **mdzub-bu* > **mdzubu* > **mdzuβu* > **mdzuɣu* > **mdzugu* > མཚུག་གུ་ *mdzug-gu*. Each of these steps has parallels. As Tibetan does not recognize geminate consonants (e.g. ཡིན་ *yino* ~ ཡིན་ནོ་ *yin-no* ‘is’), **mdzub-bu* > **mdzubu* and **mdzugu* > *mdzug-gu* are unproblematic. On **ubu* > **uβu* > **uyu* > **ugu* see Uray 1952: 186-187 and Hill 2011a. (2) Wolfenden makes a family of √*cun* (pres. རྩུན་ *hɔjun*) ‘subdue’, √*cun* (pres. རྩུན་ *hɔhun*) ‘be tamed’, རྩུན་པ་ *zun-pa* ‘melted’, and √*zu* (pres. རྩུ་ *hju*) ‘melt, digest’. He further adds Jinghpaw *tun* ‘dissolve, melt’ (1937: 640). The Tibetan word family itself is specious. The -n in √*cun* ‘subdue / be tamed’ is an inalienable part of the root; the -n in རྩུན་པ་ *zun-pa* ‘melted’ is an affix that derives nominals from verbal stems also seen in pairs such as √*ce* (pres. ཅེ་ *che*) ‘be large’, ཅེན་པོ་ *chen-po* ‘big’ and √*rku* (pres. རྩུ་ *rku*) ‘steal’, རྩུན་མ་ *rkun-ma* ‘thief’. Similarly, the *ɟ* seen in the present stem of √*zu* (i.e. རྩུ་ *hju*) arises due to excrescence after the prefix *h-* according to Conrady’s law (Conrady 1896: 59, Hill 2013: 194) and has no connection to the *ɟ* in the present of √*cun* ‘subdue’, which derives from √*cun* through voicing alternation (cf. Hill 2014). (3) Wolfenden denies that Tib. རྩུན་ *brgyad* ‘eight’ and Chi. 八 *peat* < **p^hret* ‘id.’ (20-14a) are cognate. Noting that 八 *peat* also writes a word with the meaning ‘divide’, he puts ‘eight’ in the same word family as 別 *bjet* < **N-pret* ‘be separated (intr.)’ (20-15a) and half a dozen other words (1938: 167). He connects Tib. རྩུན་ *brgyad* ‘eight’ to a distinct set of words to do with cutting, such as Chi. 鑿 *khet* < **k^het* ‘cut, carve’ (20-01f) (1938: 168). Wolfenden’s exercise to divide Tib. རྩུན་ *brgyad* ‘eight’ and Chi. 八 *peat* < **p^hret* ‘eight’ (20-14a) into separate word families, he presents as the paragon of the word family methodology. Nonetheless, most avowed practitioners of the word family method continue to see these two words as cognate without compunction (cf. Matisoff 2003: 151, Gong 2002[1995]: 107, Schuessler 2007: 152).

⁵ He reconstructs: **s-wa-p*, **s-m-wap*, **k-s-wap*, **l-wa-p*, **r-wa-p*, **k-l-wap*, **p-wa-p*, **tsi*, **tsit*, **tsit*, **tsut*, **pwar*, **pu*, **sin-wap*, **pwap-swap*, **pwar-wap*, **swap-prwap*, **sin-prwap*, **tsi-wap*, **tsi-pwap*, and **tsi-pu*.

affixations in Indo-European are precisely allofams. To explore this proposal witness the following family of words from the appendix of the *American Heritage Dictionary* pointed to by the reviewer.

*wek^w-. To speak. 1. O-grade form *wōk^w- in: a. Latin vōx (stem vōc-), voice: VOCAL, VOICE, VOWEL; b. Greek ops, voice: CALLIOPE. 2. Suffixed o-grade form *wok^w-ā- in Latin vocāre, to call: VOCABLE, VOCATION, VOUCH; ADVOCATE, AVOCATION, CONVOKE, EQUIVOCAL, EVOKE, INVOKE, PROVOKE, REVOKE, UNIVOCAL. 3. Suffixed form *wek^w-es- in Greek epos, song, word: EPIC, EPOS; EPOPEE, ORTHOEPY.

These reconstructed forms (in our notation * uek^w , * uok^w -, * $\text{u}\text{ok}^w\text{-eh}_2$, * $\text{u}\text{ek}^w\text{-es-}$) and the English words that flow from them constitute a word family *sensu lato* since they ultimately go back to the same root. However, it is not obvious that anything is gained if the reconstructed forms and the English words are called allofams; the relation between the proto-forms and the attested forms remains obscure, as does the relation among the different proto-forms. Instead of revealing the system(s), mechanism(s) and change(s) that generated these variously related forms, to dub them a word family risks to lump the words together in neglect of their individual etymologies and systematic relations.

Let us consider this Indo-European ‘word family’ with a look first at the proto-forms, before turning to the etymologies of the individual English words. The root * uek^w (attested in the verbal system as a reduplicated thematic aorist * $\text{ue}\text{-u}\text{k}^w\text{-e/o-}$ > Skt. *āvocat*, Gk. *εἶπον eîpon* ‘said’) forms a root noun of the o/e-type * $\text{u}\text{ok}^w\text{-}/\text{*}\text{u}\text{ek}^w\text{-}$ ‘voice, word’ with the rectus stem attested in Lat. *vōx*, Gk. *ὄψ ops* etc. and obliquus * $\text{u}\text{ek}^w\text{-}$ in Av. *vac*.⁶ The form * $\text{u}\text{ok}^w\text{-eh}_2$ is a so-called transponat, a mechanical back-projection into the sound system of PIE, based on the denominative verb Latin *vocāre* which presupposes a noun **wok^w-ā-* ‘saying’ for Proto-Italic. Since there is otherwise no trace of a **wok^w-ā-* ‘saying’ it should not be reconstructed for PIE solely on the basis of Italic, however, the morphological devices of **R(o)-eh₂-nouns* and denominative verbs thereof are of Indo-European provenance. The form * $\text{u}\text{ek}^w\text{-es-}$ belongs to an originally proterokinetic s-stem * $\text{u}\text{ék}^w\text{-os-}/\text{*}\text{u}\text{ek}^w\text{-és-}$ attested in Skt. *vácas-* and Gk. *ἔπος épos* ‘speech’. Therefore, the only “variants” of the root * uek^w in the PIE nominal system are found in two different and well-defined morphological contexts: on the one hand the root noun, and on the other hand, the s-stem. Turning to the English words (← means borrowed): VOCAL ← OFr. *vocal* ← Lat. *vōcālis*; VOICE ← OFr. *voiz* < Lat. *vōcem* VOWEL ← OFr. *voieul* < Lat. *vōcālis*; CALLIOPE ← latinized Gk. *Καλλιόπη Kalliópē*; VOCABLE ← MFr. *vocable* ← Lat. *vocābulum*; VOCATION ← OFr. *vocacion* ← Lat. *vocātiōnem*; VOUCH ← Anglo-French *voucher* ← OFr. *vocher* < Gallo-Roman *voticare* metathesized from Lat. *vocitāre*; ADVOCATE ← Lat. *advocātus*; AVOCATION ← Lat. *avocātiōnem*; CONVOKE ← MFr. *convoquer* ← Lat. *convocāre*; EQUIVOCAL ← Late Latin *aequivocus*; EVOKE ← Fr. *évoquer* ← Lat. *ēvocāre*; INVOKE ← OFr. *invoker* ← Lat. *invocāre*; PROVOKE ← OFr. *provoker* ← Lat. *prōvocāre*; REVOKE ← OFr. *revoquer* ← Lat. *revocāre*; UNIVOCAL formed on the basis of *vocal* and Latin *ūnivocus*; EPIC ← Lat. *epicus* ← Gk. *ἐπικός epikós*; EPOS ← Gk. *ἔπος épos*; EPOPEE ← Fr. *épopée* ← Gk. *ἐποποιία epopoía*; ORTHOEPY ← Gk.

⁶ This section assumes familiarity with the *termini technici* of Indo-European linguistics. For the reader’s convenience an overview of Indo-European ablaut and the comparative method is included in an appendix.

ὀρθόπειᾶ *orthoépeia*. Latin *vōcālis, vocābulum, (a)vocātiōnem, advocātus* as well as *ūnivocus* and *aequivocus* are nominal and *vocītāre, convocāre, ēvocāre, invocāre, ēvocāre, prōvocāre, revocāre* are verbal derivatives ultimately from Lat. *vocāre*. The derivational processes giving rise to these forms are completely transparent in Latin synchronic grammar (see Weiss 2009: 266-324), their development from Latin in the Romance languages is reconstructable with the comparative method (see, e.g., Jensen 1999) and the history of their transfer to modern English is philologically recoverable. The same is true for Gk. ἐπικός *epikós*, ἐποποιία *epopoía* and ὀρθόπειᾶ *orthoépeia*, all of which are derived from ἔπος *épos*. Strictly speaking, from the Proto-Indo-European point of view we are dealing with exactly two variants in this 'word family', namely, the PIE s- and a (likely post-PIE) *R(o)-eh₂- stem from the root * h_2ek^w . All the apparent allofamic variation disappears once linguistic systematicity, the comparative method, and philological rigor are applied.

Before prematurely concluding that allofams are chimera of Trans-Himalayan slovenliness, one must concede that Indo-European offers up yet more thorny doublets. Consider English *feather* < * pét-r-eh_2 - and Greek πτερόν *pterón* 'feather, wing' < * pt-er-ó .⁷ The two forms of this root, namely * pét-r- and * pt-er- , can be explained along the following lines without recourse to the notion of allofamy: Indo-European had an original proterokinetic heteroclitic noun that took the form * pét-r in the rectus stem and * pt-én- in the obliquus stem, cf. Hitt. *pettar, pettan-* 'wing, feather'. English 'feather' derives from * pét-r-eh_2 - 'collection of feathers' with the *-eh₂ collective suffixed to the inherited rectus stem * pét-r- .⁸ Greek πτερόν *pterón* continues * pt-er-ó 'feathery thing', an *-o- possessive derivative of a stem * pt-er- , which itself is an analogically renewed obliquus stem, i.e. * pér-tu- (ON *fjgrðr*)⁹: * $\text{p}_r\text{-téu-}$ (Lat. *portus*, English *ford*) 'crossing' :: * $\text{pét-r} : X = \text{*pt-ér}$, or the like. Thus, although it is correct to reconstruct both * pét-r- and * pt-er- in the proto-language *sensu lato*, it is not the case that any one *état de langue* is hypothesized to have contained both. To have left matters at saying * pt-er- and * pét-r- are allofams would be unsatisfactory. Cases like 'feather' in Indo-European can approach Matisoff's 'lung' in complexity. For example, Alan Nussbaum (1986) sets up 12 related Indo-European words 'head', 'horn', 'skull' etc., but across 293 dense pages, he explains each of them one by one as ultimately deriving from the single word * kér-h_2 'head-bone (material)', using combinations of derivation, analogy, and semantic change, as the case may be.

These Indo-European case studies make clear that in order to advance the kind of explanations that account for apparently intractable proto-variation, it is necessary for the morphological alternations posited in the protolanguage to have left sufficient, even if very fragmentary traces in attested languages. The terms of the Indo-European explanations (proterokinetic and heteroclitic, collective and possessive derivative) refer to attested declensional and derivational patterns. One reason for the prevalence of allofamy in the reconstruction of Trans-Himalayan is probably that morphologically innovative languages (such as Lahu) have played a more substantial

⁷ This pair is Trask's example of an 'oblique cognates' (2000: 235).

⁸ Cf. also * pet-n-o > OIr. *én* 'bird' and * pet-n-eh_2 > Lat. *penna* 'feather'.

⁹ ON *fjgrðr* via Scottish → English *firth* and via Norwegian → English *fjord*.

role than those (such as Limbu) that retain more inherited patterns. To sum up the immediately foregoing discussion, Matisoff's proposal for 21 variants of 'lung' is not totally unprecedented and unreasonable, but until the forms are rigorously accounted for in an elaborated theory of Trans-Himalayan historical phonology and morphology, they are better called 'problems' rather than 'allofams' and once they are explained to calling them 'allofams' is superfluous.

.1 Seven types of allofams

When Matisoff introduces allofams he is judicious in specifying that it is a cover term for different types of phenomena from well understood morphology to imagined relationships where none exist.

[W]ord-families are groups of forms which bear a non-fortuitous phonological and semantic relationship to each other. The sound/meaning relationships among the allofams of a word-family may follow a more or less productive pattern, so that in favorable cases the variations may be traced back to systematic (or at least plausible) alternations in the proto-language itself (often involving proto-affixations). In many cases, however, the synchronically observable intra- or inter-lingual allofamy follows no particular pattern that repeats itself elsewhere. This situation may result from conflicting or overlapping morphological processes that obscure each other's outputs, unsystematic or sporadic increments to roots, interference or contamination from genetically unrelated forms, dialect mixture—or of course it is always possible that the forms in question were never co-allofams at all, and their resemblance is entirely specious. (Matisoff 1978: 18)

Looked at carefully each allofam tells a story that combines in some measure regular phonology, borrowing, and analogy. Matisoff himself names (1) dialect variation, (2) morphology, and (3) specious resemblance as three types of allofams. As four additional sources of allofams I identify (4) areal words, (5) overlooked sound laws, (6) contamination, and (7) mistaking recent developments for proto-variation. In instances of all seven categories the identification of allofams has prevented the correct identification of the underlying phenomenon.

.1.1 Dialect variation misanalysed as proto-variation

Dialect mixture is a challenge for historical linguistics and often leads to the impression of proto-variation. In English all words that begin with v- are loanwords. For example, the word 'very' comes from Old French *verai* 'true' and 'vicar' comes from Latin *vicarius* 'substitute, deputy' via Old French *vicaire*. There are however three apparent exceptions to this rule of thumb, 'vane', 'vat', and 'vixen'; these words have the good German cognates *Fahne*, *Fass* and *Füchsin*. With the evidence of the modern languages one might reconstruct *v-/*f- variation in proto-Germanic. Aside from the inelegance of introducing a phonemic contrast to explain three words, Old English confirms an original f- in *fana*, *fet* and *fyxan*. The reason for the exceptional v- onsets is that these words came into standard English from a southern English dialect which had changed all f's to v's through a regular sound change. In both standard English and the southern English dialect source of these three words sound change operates regularly, but inter-dialect borrowing yields an impression of mercurial sound change and proto-variation.

Sanskrit offers another illustration of dialect borrowing. Sanskrit r- and l- both correspond to both r- and l- in

other Indo-European languages. Thus, in the Sanskrit word *raji-* 'direction' an r- corresponds to r- in other languages (Gk. ὀρέγω *orégō* 'reach', Lat. *regō* 'guide, steer', Eng. 'right'), but in the word *rócate* 'it shines' r- corresponds to l- in other languages (Hitt. *lukezzi* 'kindles', Toch. *lyuketär* 'lights up'; also Gk. λευκός *leukós* 'bright', Latin *lūx* 'light', Eng. 'light'). In the Sanskrit *loka* 'world, realm' an l- corresponds to an l- in other languages (Lat. *lūcus* 'sacred grove', Lith. *laūkas* 'field, land'). In the word *lohá* 'copper' an l- corresponds to an r- in other languages (Lat. *rūdus* 'lump [of bronze]', OCS *ruda* 'metal'). One could posit proto-variation between *r- and *l- to account for the chaotic correspondences, but this merely disguises a Sanskrit problem as a feature of Indo-European. One could reconstruct four separate origins in Indo-European to explain the four separate correspondences, but such a mechanical approach does not explain why it is only Sanskrit that disagrees with the other languages.

There are also l- and r- doublets ('allofams') within Sanskrit, such as *riptá, líptá* 'smeared', *róman-, lóman-* 'hair' (cf. Irish *ruaimneach* 'long hair'), *róhita-, lóhita-* 'red' (cf. Gk. ἐρυθρός *eruthrós*, Lat. *ruber*, Eng. 'red'). Such doublets suggest that Sanskrit is of no value for distinguishing inherited *r- and *l-. Even without further evidence one would be entitled to posit some kind of dialect mixture. A close examination of textual attestations shows that r- is more prevalent in older and more western documents whereas l- is more typical outcome of the East and in younger documents (Fortson 2010: 204 §10.8, 211 §10.34).

Such cases of dialect mixture masquerade as a challenge to the exceptionlessness of sound change, but Campbell makes the fundamental methodological point with reference to dialect mixture in Norman French.

While these words are exceptions to strict exceptionlessness of sound change, we cannot explain their individual histories, that they are exceptions, without reference to the sound change. (Campbell 2004: 215).

It is only a methodological commitment to exceptionlessness that allows potential exceptions to be discovered and explained. The neo-grammarians hypothesis of exceptionless sound change is less a hypothesis which itself must be tested and more a research axiom which allows one to test hypotheses about the histories of languages. This axiom is rooted in sound *laws* that are scientific laws *stricto sensu* since they are governed by physical entities of the auditory and articulatory apparatus as well as the linguistic computational system.¹⁰

Dialect variation and exceptionless phonetic changes are two ways of looking at the same phenomenon. If a wave of linguistic change passes through a community in an exceptionless fashion, then we have an exceptionless sound law. If a wave of change merely laps up against a language community, sporadically changing the pronunciation of a few words, then it is not a sound law at all, but merely the borrowing of a few words from a neighboring dialect (in which the sound change did occur without exception). The first one is linguistic change *stricto sensu*, the second one is diffusion. Strictly speaking the three English words that have v- rather than f- are not part of the inherited lexicon of Standard English, but rather are loanwords from a southern dialect of English which had changed 'f' to 'v'. The contradiction between the wave model and family tree model is only apparent. When doing

¹⁰See Hale 2007: 124-144.

dialect geography it is useful to locate the geographic diffusion of each feature, but from the perspective of the synchronic linguistic state (*état de langue*) all sound change is mechanical and exceptionless. Thus, Matisoff's invocation of the principle 'chaque mot a son histoire' (Matisoff 1975: 164) to defend the scarcity of exceptionless sound laws in Benedict's work does not answer Miller's criticism." Each word has its own history; some words are inherited, some borrowed, some affect by this sound change, some by that one, some analogically altered, etc. The duty of the historical linguist is to trace these histories one by one with exceptionless sound changes as his primary investigatory tool. Words have their own histories because each suffers the vicissitudes of history as befits it, not because each can idiosyncratically exempt itself from the inexorable flow of time, yielding the "sporadic survival of rare, unshifted sounds" (Matisoff 1975: 164 approvingly quoting Miller's 1974: 199 paraphrase of Benedict).

Within the Trans-Himalayan family Chinese provides a good example of inter-dialect borrowing. A Chinese character generally contains two components, one that gives a hint to the pronunciation of the character and one that helps to distinguish this character semantically from other characters with the same pronunciation. The phonetic components refer to the pronunciation at the time of the script's invention or codification. At times the same phonetic component is used both for words that end in Middle Chinese with -j and those that end in Middle Chinese with final -n (e.g. 24-35d 難 *nan*, and 24-35l 鬮 *nej*). One must presume that at the time the characters were coined both readings ended with the same consonant. The original final cannot be either -n or -j because there are series of characters with shared phonetic components where all of the readings are -n (e.g. 24-23a 亶 *tanX* < *tʰanʔ, 24-23g 禮 *trjenX* < *tranʔ, 24-23j 鱸 *trjen* < *tran, 24-23l 氈 *tsyen* < *tan) and others where all of the readings are -j (e.g. 18-18a 麻 *mae* < *C.mʰraj, 18-18c 糜 *maH* < *mʰajs, 18-18f 磨 *ma* < *mʰaj, 18-18g 糜 *mje* < *C.maj). Starostin's solution is that in these cases the original final consonant is -r and that in Western Chinese it changed to -n but in eastern China -r changed to -j (1989: 400). This dialect variation is explicitly discussed in Han dynasty scholastic works. For example, in a commentary on the 呂氏春秋 *Lǚshì chūnqiū* the scholar 高誘 *Gāoyòu* (*floruit* c. 200 CE) wrote that "the people of 兗州 *Yǎnzhōu* [in modern-day Shāndōng province] all pronounce the clan-name 殷 *Yīn* as 衣 *Yī*" (Baxter 1992: 295, bracketed comments are Baxter's). Any comparison of Chinese to other languages which relies on Old Chinese reconstructions that do not make use of Starostin's insight will discover variation between -n and -j, but this variation is of no comparative value.

Dialect variation within Tibetan has led both Conrady and Matisoff astray. Conrady reconstructs Tibetan ཉ་ *rta* 'horse' as *rtag, pointing to the Gtsañ dialect form ཉིག་གི་ *rtig-gi* 'foal' (Jäschke 1881: 212), which he reconstructs as *rtag-bu, i.e. *rtag followed by the diminutive suffix -bu (1896: xii). In fact, ཉིག་གི་ *rtig-gi* is an un-etymological respelling of ཉེ་ *rtehu* based on the pronunciation of the latter in this dialect (Hill 2009: 122, Hill 2011a: 119-120).

¹¹This phrase is normally credited to Jules Gilliéron, but Campbell credits it to Hugo Schuchardt (2004: 212-213) and Joseph traces it even further back to Grimm (2002: 50).

Matisoff cites Tibetan རྟོག་ *lyag* 'good' (2003: 51) as evidence for a reconstructed root *l(y)ak ~ *l(y)aŋ ~ *m-d(y)ak; Jäschke, Matisoff's source, explicitly gives this word as a vulgar pronunciation of རྟོག་ *yag* 'good' (1881: 174). Matisoff has projected variation among Tibetan dialects into proto-Trans-Himalayan.

.1.2 Morphology misanalysed as proto-variation

English exhibits an alternation between -s- and -r- in pairs of words such as 'was' and 'were', 'lose' and 'forlorn', or 'freeze' and 'frore'.¹² German normally confirms the variant with -r- (*war, verlieren, frieren*), but rarely shows variation itself, as in the stems of the verb *erkiesen* (archaic), *erkor, erkoren* 'choose', for which English has only -s-. If today's German and English were the only two attested Germanic languages such cases or -r-/s- variation would be prime candidates for allofams in proto-Germanic. These alternations are however not evidence of proto-variation, but instead are the final remnants of a once regular pattern. The pattern is still clear in the principal parts of Old English verbs, thus *freosan, freas, froren* 'freeze', *ceosan, ceas, coren* 'choose', *forlēosan, forlēas, forluron* 'lose'. The alternations originally arose as an effect of Verner's law (Fortson 2010: 341, 345-346). An exceptionless phonetic change gave rise to a morphological pattern which through late analogical leveling has led to the impression of proto-variation.

This example showcases the importance of synchronically irregular morphology for restoring the regular morphological system of an earlier stage. Whereas for phonology it is regular phonological correspondences that are paramount, "les procédés particuliers d'expression de la morphologie" are "ce qui est probant pour établir la continuité entre une « langue commune » et une langue ultérieure" (Meillet 1925: 25). Many allofams are probably remnants of ancient morphology.

Guillaume Jacques has taken up the search for irregular morphology among the Trans-Himalayan daughter languages. In 2007 he identified a pattern in personal pronouns. Chang Naga and Southern Qiang exhibit the same pattern of stem suppletion in the first person pronoun. In Chang Naga the casus recti are formed from *ŋò* and the casus obliqui are formed from *kɣ-* in Southern Qiang the casus recti are formed from *ŋa*⁵⁵ and the casus obliqui from *qa*⁵⁵ (Jacques 2007). This discovery provides a useful model for explaining the existence of two first person singular stems in Tibetan ཎ་ *na* and *kho-*.¹³

As pointed out by Rolf Stein, prefixes and suffixes are morphology, and as morphology they have functions that can be examined through the close examination of textual passages; they are not wild-cards.

Le rôle des pré- et suffixes dans la sémantique est certain, mais la recherche devra considérer chaque cas en particulier. Aussi faut-il avoir recours non pas seulement au collationnement des lexiques, mais encore à l'analyse des textes. (Stein 1941: 206)

Until more work is done on Trans-Himalayan morphology on the basis of languages such as those of the Kiranti

¹²A now archaic adjective (cf. Shelley's "Expiring in the frore and foggy air", 'The Revolt of Islam,' Canto 9, XXV, 1818).

¹³The stem *kho-* is seen in མོ་ མོ་ མོ་ *kho-bo* 'I' (male speaker) and མོ་ མོ་ མོ་ *kho-mo* 'I' (female speaker).

and Rgyalrongic subbranches, one may expect to see much apparent proto-variation as the residue of moribund or defunct morphology. Handel makes a similar point in his characterization of the state of the art.

One of the great challenges in S[ino-]T[ibetan] linguistics is to uncover the various morphological processes operating in the different languages and subgroups, thus identifying root morphemes. This will, on the one hand, allow us to better explain what is now simply identified as 'variation' and, on the other hand, permit a more precise specification of sound laws and sound correspondences. (Handel 2008: 425-426)

To paraphrase, the sooner the discipline replaces a search for allofams with a search for sound laws, the more it will flourish.

.1.3 Non-etymological resemblance misanalysed as proto-variation

Two words sometimes appear similar in meaning and in phonetic shape, but no known process accounts for their relationship. A well-known example is Latin *deus* 'god' and Greek *θεός* *theós* 'god', which seem as if they must be cognate, but they are not (Fortson 2010: 25). Similarly, Latin *habeo* and English *have* are nearly identical in form and meaning, but they reconstruct to separate Indo-European roots **keh₂p* (Lat. *capio*, Eng. 'have') and **g^heHb^h* (Lat. *habeo*, Eng. 'give'). The temptation may remain to regard **keh₂p* and **g^heHb^h* as allofams, but no known process in Indo-European relates them. Whether the phonetic and semantic similarity of these roots is pure coincidence, or whether it is due to some obscure morphological process is not relevant until this process emerges into the light. Both roots are supported by cognates in daughter languages and nothing would be clarified by presenting them together as **keh₂p* and **g^heHb^h*. If some scholar is able to relate them in a way that also solves seemingly unrelated problems—the ultimate test of a good hypothesis—then let such a scholar step forward. Until such a time, they are separate roots, plain and simple.

The method of allofams permits non-etymological resemblances to pass as cognates. Matisoff chides Miller for proposing that Burmese *နီဝ်* *nā* 'ill', *နီလ်* *nū* 'leprosy', and *နီဝ်* *nī* 'red' are related and he facetiously proposes an English word family of 'spit', 'split', 'slit', 'shit' (1978: 46). Although Matisoff makes clear that he does not accept Miller's proposed word family, he does not discuss in methodological terms how an investigator who employs allofams among his conceptual tools is able to avoid collecting the wrong would-be allofams. Miller's proposals are not *prima facie* more ridiculous than Wolfenden's association of 'subdue' and 'melt' or 'eight' with words for cutting (*vide supra*). I am aware of no explicit discussions in the literature concerning how to adjudicate proposals of word family membership.

Matisoff himself has at times proposed invalid word family relationships on the basis of non-etymological resemblance. For example, he reconstructs a word **l-tsyak* 'iron' to account for Tibetan *ལྷགས་* *lčags* 'iron' and Burmese *လက်* *lak* 'bit (bridle)' (2003: 317). Native Burmese words do not begin with voiced stops, the semantics of the comparison are weak, and 'iron' is too recent a technology to be of Trans-Himalayan provenance (cf. Wagner

1993: 52-96).¹⁴ Exceptionless sound laws provide a measure by which proposed cognates may either be accepted or rejected; allofams provide no such explicit means to exclude non-etymological resemblance.

1.4 Areal words misanalysed as proto-variation

When a new product or technology is discovered or popularized the word for it is often borrowed and re-borrowed quickly and widely such that it is difficult to establish exactly which language borrowed it from where. This phenomenon is well known in modern languages, as the forms of 'tea' and 'metro' in European languages attests; the same phenomenon has occurred since the dawn of time. Although Indo-European languages have very similar looking words for 'wine' the reconstructible ancestors of various languages differ in detail (Gk. (Ϝ)οἶνος (*w*)oĩnos < *uóih₁no, Lat. *vīnum* < *uīh₁no,¹⁵ Hitt. *wiyana-* < *uīh₁ono, Arm. *gini* < *uóih₁nijo-). This apparent variation, particularly in light of similar words in Semitic (Arabic *wayn*, Hebrew *yayin*) and Kartvelian (Georgian *gvino*, Laz *gūini*) suggests that these many forms reflect an areal term that spread along with the technology (Beekes 1987, Rubio 1999: 8, Hock & Joseph 2009: 483, Fortson 2010: 42).

Matisoff reconstructs a Sino-Tibetan word *s/m-ra-ŋ 'horse' to account for Bur. འཇམ་མཚོ་ *mranh*, OTib. མར་ *rmañ*, and Chi. 馬 *maeX* < *m^hra? (01-73a) 'horse' (2003: 177). With the notation *s/m-ra-ŋ Matisoff appears to propose that *sraŋ, *sra, *mraŋ, *mra, *ra and *raŋ were all valid pronunciations for 'horse' in the proto-language; six somewhat similar words for horse showing no semantic differentiation among them in a single synchronic system is to my knowledge unprecedented. This proposed cognate set violates the normal correspondence of Chinese *mr- to Tibetan *br-*, known as 'Simon's law' (e.g. Tib. སྤྲུང་ *sbrañ* 'fly, bee' and Chi. 蠅 *ying* < *m.rəŋ [06-24a] 'fly' or Tib. སྤྲུལ་ *sbrul* 'snake' and Chi. 虺 *xjwǐjX* < *[m̥r]uj? [27-19a], cf. Hill 2011b: 448-449) and also shows an irregular correspondence among the finals; Chinese rising tone (-X < *ʔ) normally corresponds to Tibetan open syllables (e.g. Tib. ལྔ་ *l̥ia* 'five' and Chi. 五 *nguX* < *C.ŋ^sa? [01-29a] 'five', Tib. ཁ་ *kha* 'bitter' and Chi. 苦 *khuX* < *k^ha? [01-01u] 'bitter', or Tib. ལྟོ་ *lta* 'look at' and Chi. 睹 *tuX* < *t^sa? [01-38c] 'see'). These phonetic irregularities in the word for 'horse' alone suggest that it may be a *Wanderwort* and not a cognate set reconstructible to the proto-language. Considerations of material culture change the supposition that 'horse' is a *Wanderwort* into a certainty. Because horses appear in the Chinese archeological record around 1250 BCE during the 商 *Shāng*, China's first historic dynasty (Anthony 2007: 456-457; Li & Chen 2012: 111-113), it is not possible that the proto-language had a word for 'horse'. The irregular sound correspondences observed in *Wanderwörter* leads to an appearance of

¹⁴Tib. ལྟོ་ *l̥ags* 'iron' is probably related to Chi. 鐵 *thet* < *ʃ^sik (1256b) 'iron' in so far as they are both instances of the same *Wanderwort* that is also present in Tai (cf. Schuessler 2007: 497). Chang Kun (1972) also made the mistake of proposing that 'iron' was an inherited word.

¹⁵A reviewer wondered whether Lat. *vīnum*, might also descend from *uóih₁no-, but the Latin form together with Faliscan *vinu*, Umbrian *vinu*, and Volscian *vinu* all point to Proto-Italic *wīno- < *uīh₁no-. In particular, Lat. *vīnum* is not from *uóih₁no- because of the Sabellic cognates. Proto-Sabellic preserves Proto-Italic diphthongs. The *i* in Umbrian cannot go back to *oi, which would have given *ō and in turn *u*, cf. *unu* 'one' (Lat. *ūnus*) vel. sim..

variation in the proto-language, if one fails to distinguish such *Wanderwörter* from inherited vocabulary.

.1.5 Overlooked sound laws

The interchange between Sagart and Matisoff over the question of whether Burmese *w-* corresponds to Tibetan *g-* is an illustrative example of how recourse to allofams prevents one from noticing sound laws, and thereby impedes progress. Sagart presents the evidence of Figure 1 (slightly amended), writing that Matisoff “has missed a sound correspondence” (2006: 211).

<i>Tibetan</i>	<i>Meaning</i>	<i>Burmese</i>	<i>Meaning</i>
གོ <i>go</i>	space	အဝ <i>awa</i>	space
ဂြိုမာ <i>gro-ma</i>	potentilla anserina	ဝ <i>wa</i>	tuber
နွေမို <i>sgor-mo</i>	round	ဝနံ့ <i>wanh</i>	round

Figure 1: Sagart's examples of Tib. *g-* corresponding to WBur. *w-*

After dismissing 'round' on the grounds that he prefers to compare Written Burmese ဝနံ့ *wanh* to Mizo *val*, Matisoff admits that the comparison of Tibetan ဂြိုမာ *gro-ma* and Written Burmese ဝ *wa* “looks good, but the only way to make sense of it is to assume a form like *grwa (173–174), which was treated by W[ritten]B[urmese] as if it had a double prefix, *g-r-” (2007: 437-438); Sagart paraphrases “Burmese has 'treated g and r as prefixes' (understand: has lost g- and -r- as a result of untold, random processes), so that only w- remains” (2008: 154). Matisoff's account is manifestly not the only way to make sense of the comparison of Tibetan ဂြိုမာ *gro-ma* and Written Burmese ဝ *wa*; an alternative is to posit a segment, such as a labiovelar, or *w-* itself that regularly changed into *g-* in Tibetan and *w-* in Burmese.¹⁶ To dismiss Tibetan ဂေါ *go* 'space', Matisoff simply writes that it “is clearly part of the same word-family as *sgo* 'door' and that it “seems likely that this is somehow related to *m-ka “door”, no doubt via an intermediate form like *kwa or *gwa” (2007: 438). By positing ဂေါ *go* as a member of a word family and speculating about how it may be related to a number of starred forms, the justification for which is not provided, Matisoff dismisses Sagart's proposed sound law, concluding “there is no reason at all to think that *w- > W[ritten]T[ibetan] gr- or g-” (2007: 438). The same method, positing different word families for the two members of a comparison and thereby dismissing the comparison, is also what led Wolfenden to dismiss Tibetan འབྲུག *brgyad* and Chinese 八 *peat* < *p^hret (20-14a) 'eight' as cognates (*vide supra*). At least in this interchange between Sagart and Matisoff, allofams appear to serve the function off shutting down scholarly inquiry rather than advancing it.¹⁷

¹⁶Consider Proto-Indo-European **u* > *g* in Armenian (e.g. Armenian *gorc* 'work' versus English *work*, Armenian *get* 'river' versus English *water*, cf. Schmitt 1981: 69) or **w* > *gu* in Romance borrowings from Germanic, e.g. French *guerre* < Frankish **werra* 'war', French *guise* < Frankish **wisa* (cf. Old High German *wisa*, Old English *wise*) 'manner'.

¹⁷Matisoff complains that Sagart “puts my term 'allofams' ... in shudder-quotes throughout, perhaps because he thinks the concept ... is unworthy of serious discussion” (2007: 439). This objection is surprising given Matisoff's own use of quotes to

Let us stay for a moment with 'eight'. Indo-European $*h_1oktōu$ becomes Old English *eahta* 'eight', Latin *octō*, and Sanskrit *aṣṭāu*, according to regular and explicitly named sound laws.¹⁸ In contrast, it is unclear how Matisoff's reconstruction $*b-r-gyat$ 'eight' yields such forms as Tibetan འཕྱེད་ *brgyad*, Written Burmese ရဲဝှ်ရဲဝှ် *rhac*, Jingpho *matsát*, Chokri Naga *tatha*, and Mizo *pariat*.¹⁹ Conversely, it is unclear what leads Matisoff to reconstruct $*b-r-gyat$, rather than $*m-t-ryat$, $*par-g-tsat$ or one of many other potential proposals. Matisoff's own remark is that the "reflexes of this phonologically complex numeral are predictably varied" (2003: 151).

There are sound laws at work in the story of Sino-Tibetan 'eight'. In particular, Li Fang-Kuei, comparing Tibetan འཕྱེད་ *brgyad* 'eight' to Chinese 八 *peat* < $*p^{s}ret$ (20-14a) 'eight', proposed that the *-g-* witnessed in Tibetan is epenthetic (Li 1959: 59; Hill 2011: 447-448). Comparison to Old Burmese ရဲဝှ်ဝှ် *rhyat* (Nishi 1999: 47) rather than Written Burmese ရဲဝှ် *rhac* further improves the situation.

Reconstructed words of a reconstructed language consist of reconstructed segments. The comparison of a word in one language (e.g. English 'day') with a word in another language (e.g. German *Tag*), with knowledge of the sound laws relating the two languages allows one to reconstruct their ancestor (Germanic $*dagaz$). In contrast, by "the mere collation of a 'word family' in one language with one in another we will fail to establish correspondences between the languages compared, and we will not succeed in elucidating the character of their proto-language" (Nishida 1977: 9). To give a concrete example, if like Wolfenden one compares all at once the Tibetan words $\sqrt{c}un$ 'subdue' / 'be tamed', རྩུ་པ་ *zün-pa* 'melted', and $\sqrt{z}u$ 'melt, digest' with Jinghpaw *tun* 'dissolve, melt' (1937: 640), it is not possible to decide whether the proto-language had a final $*-n$ or not (see note 4). It is impossible to decide whether the ancestral lexeme began with *t-*, *č-* or *z-*. Only by pairwise equations of words in cognate languages is it possible to identify correspondence patterns and "gerade Lautentsprechungsgesetze sind doch die Grundlage jeder Weiterforschung und darum das Ziel, dem vor allen Dingen zuzustreben ist" (Conrady 1896: viii).

.1.6 Contamination

Words of closely associated meaning may phonetically influence each other in a way that is not easy to model with

caste suspicion on the much better established "regular sound correspondences" (2007: 437).

¹⁸ The journey from $*h_1oktōu$ to OEng. *eahta* 'eight' entails a stop at proto-Germanic $*ahtōu$ (cf. Gothic *ahtau*) and Old English *eahta*.¹⁸ In observing the procession from Indo-European to proto-Germanic, one may first consider the consonants. The laryngeal $*h_1$ before vowels disappears without a trace. The palatal $*k̑$, after merging with $*k$, becomes $*h$ following Grimm's law; the $*t$ remains unchanged because it is the second of two adjacent voiceless stops. Turning to the vowels, the change of 'o' to 'a' in Germanic is regular and unconditioned; the vowel $*ō$ remains, merging with \bar{a} . These changes arrive at the reconstruction of $*ahtōu$ in proto-Germanic. Old English brightened 'a' and 'ā' to 'æ' and 'ǣ' (e.g. Old English *ræt* versus German *Ratte* – but not before nasals cf. OEng. *dranc* < West Germ. *drank*); then, front vowels including 'æ' and 'ǣ' broke into diphthongs before *h*, *w*, or *r*, plus another consonant, for our purposes changing $*aht-$ to *eaht-*. Turning attention from the front of the word to its back, unstressed diphthongs were monophthongized on their way to Old English ($*-tōu$ > $*-tō$). Subsequently unstressed 'ō' became 'a' (Hogg 2011: 233 §6.28) and through these steps Germanic $*ahtōu$ yields Old English *eahta*. This footnote relies on Forston (2010) and Ringe (2006).

¹⁹The rule 'remove the hyphens to yield the Tibetan form' works in this and other cases (e.g. Tib. སྐུ་མཚན་ *sman* < $*s$ -man, Matisoff 2003: 37), but not always (e.g. Tib. སྐུ་མཚན་ *śi* 'die' < $*səy$, Matisoff 2003: 34, Tib. རྩུ་པ་ *legs* < $*l(y)ak$, Matisoff 2003: 51).

four-part analogy. A famous example is the irregular development of Modern English *female* from Middle English *femele* under the influence of *male*. In another example, Indo-European *k^wetwóř 'four' irregularly became Proto-Germanic *petwóř > *fedwōř (Go. *fidwōř*, OEng. *fēower*) under the influence of the *p- in *pénk^we 'five' (Gk. πέντε *pénte*, Skt. *pāñca*, Lithuanian *penki*). In neither of these two cases of contamination is variation in the proto-language the correct explanation.

Matisoff reconstructs Tibeto-Burman *l/b-ŋa 'five', with proto-variation between the initials *l- and *b- (2003: 149); it is more likely that only the lateral initial is correct (Tib. ལྷ་ *lha*, Dakpa *le³ŋe⁵³*) and that forms such as Mizo *pà-ngá* 'five' are contaminated by the labial initial inherited in the word for 'four' (Mizo *pà-lí*, Tib. འཇི་ *bži* < *blyi, Kuröp *ble*). In this way, contamination explains away another allofam and rids the proto-language of unnecessary variation.

1.7 Misanalysed language internal developments

The philological study of classical languages has served as the bedrock of Indo-European comparative linguistics. Indo-Europeanists acquire a reading knowledge of most of the older languages in the family, including Greek, Latin, Sanskrit, Hittite, Old Church Slavonic, Old Irish and so on. In the Trans-Himalayan family Tibetan, Burmese, Chinese, Tangut, Newar, Meithei, and Lepcha count among the early attested languages. Benedict and Matisoff cite words from Tibetan, Burmese, and Chinese but make essentially no use of the other early attested languages. Their use of Tibetan, Burmese, and Chinese does not extend to the consultation of primary documents.²⁰

Despite the availability of good philological studies and a number of glossaries of Old Tibetan, Matisoff does not make use of data from this early stage of the language. This negligence has concrete effects; Matisoff proposes a root *b-rəy 'draw / write' "on the basis of Tibetan allofams like *hbri-ba* 'draw, write', *bris* 'a picture' on the one hand, and *ris* 'figure, form, design' and *ri-mo* 'id.' on the other" (2003: 132). The Old Tibetan conjugation of the verb is འདྲི་ *hdri* (present), རྩིས་ *bris* (past), རྩི་ *bri* (future), རྩིས་ *ris* (imperative), with root √ri; the present form འདྲི་ *hbri*, which Matisoff cites, is a later development (Hill 2005). The noun རྩིས་ *bris* 'a picture, i.e. that which has been drawn' is a nominalized form of the past stem རྩིས་ *bris*. Thus, no Tibetan evidence supports a *b- in Matisoff's reconstruction *b-rəy.

In another case Matisoff proposes that Tibetan གསལ་ *gsal* 'clear' derives from *sal (2003: 405). However, Classical Tibetan s- reflects a merger of Old Tibetan *sts-* and *s-*. Examples such སྡོག་པ་ཐམས་ཅད་བསྐྱུལ་དྲེ་ *sdig-pa thams-cad bstald* 'clear away all sins (IOL Tib J 751, f. 40v, l. 1) and བར་ཚད་ཐམས་ཅད་ཡོངས་སུ་བསྐྱུལ་དྲེ་ *bar-chad thams-cad yoñs-su bstalte* 'completely clear away all hindrances' (PT 16, f. 29r, l. 2) show that the etymological root for this word is

²⁰Chang says of Benedict 1972 that forms "gleaned at random from dictionaries and taken at face value are the extent of the written Tibetan data" (1973: 336); because of this reliance on dictionaries in an interview Li describes the reconstructions of Benedict and Matisoff as "junk" (1988: 68).

√stsal rather than √sal; the reconstruction *sal cannot be correct.²¹ In sum, some allofams disappear when one considers in the light of philological evidence; a working knowledge of the primary and secondary literature of classical languages is necessary in the work of reconstruction.

.2 Conclusion

The foregoing discussion has shown that allofamy is not necessary to describe the inheritance of cognate linguistic forms from a reconstructed ancestor, either in Indo-European or in Trans-Himalayan. Things that have been called allofams can be classified into at least seven different types according to the historical processes that gave rise to the apparent variation. The apparent cases of allofams must be scrutinized, one by one, to identify the specific historical process that gave rise to each. The conclusion that Trans-Himalayan linguistics should follow the Neogrammarian programme is not novel. In a passage already quoted earlier, Conrady, the founder of our discipline, writes that “Lautensprechungsgesetze sind doch die Grundlage jeder Weiterforschung und darum das Ziel, dem vor allen Dingen zuzustreben ist” (Conrady 1896: viii). Also, a number of researchers over the decades have expressed concern that Trans-Himalayan lacks the rigorous standards of the better studied language families. In a review of Benedict’s *Conspectus* Roy Andrew Miller draws writes that between “these two extremes—the comparative method with its sound correspondences on the one hand and Benedict and his “generalizations” on the other—we must, in all honesty, recognize the existence of a considerable chasm” (1974: 198-199). Chang Kun similarly cautions that “systematic phonological reconstruction is the first requisite for historical work in linguistics” and points out “its general absence here” (1973: 336).²² Nishida (1977: 9) and G. Starostin (2009) specifically take aim at word families as incompatible with traditional methodology.²³

It is only fair to practitioners of word family linguistics to admit that their research has benefited the field enormously. Matisoff’s *Sino-Tibetan Etymological Dictionary and Thesaurus* (2015) takes Carl Darling Buck’s *Dictionary of Selected Synonyms in the Principal Indo-European Languages* (1949) as a model, but as the very first effort to systematically draw together the roots of the Trans-Himalayan proto-language a more apt comparison is

²¹Better Burmese philology also has a role to play in improving reconstruction. For example, on the basis of Written Burmese ရှစ် *rhac* ‘eight’ and ရပ် *rap* ‘stand’, Matisoff reconstructs Lolo-Burmese as *ʔ-rit^l and *ʔ-rap^l (2003: 56, 315), but Old Burmese offers ရှိတ် *rhyat* and ရှိတ် *ryap* (Nishi 1974, 1975).

²² Matisoff’s reply to Miller evinces a lack of enthusiasm for the principle of the exceptionlessness of sound laws (*Ausnahmslosigkeit der Lautgesetze*); he complains that Miller “likes everything to be super-tidy and super-rigorous” (1975: 165). In 1978 Matisoff devoted a monograph to elaborating an alternative theory of the ‘allofams’ to justify the absence of sound laws in Tibeto-Burman. A quarter of a century later Matisoff writes in his *Handbook*, “I follow Benedict in caring little for a chimerical methodological purity” (Matisoff 2003: 9). In a reply to Sagart’s review of this book, Matisoff complains that Sagart “can’t seem to accept the fact that not everything can be known about ‘regular sound correspondences’ at this stage. I have earned the right to make occasional educated guesses” (2007: 437).

²³ Haudricourt’s (1954a and 1954b) explanation of tonogenesis in the history of Chinese is a famous example of traditional methodology with reference to East Asian languages, and shows that application of the method does lead to interesting and important discoveries.

August Friedrich Pott's *Etymologische Forschungen auf dem Gebiete der indogermanischen Sprachen* (1833-6). Reporting a comparison inherited from his teacher Franz Bopp, Pott included Latin *deus* and Greek *θεός theós* as cognates of Sanskrit *deva* 'god', although noting the phonological irregularity (Davies 1998: 173-174). Scholars such as Theodor Benfey (1837) and Georg Curtius (1862) slowly brought the *opinio communis* to reject this proposal, with Max Müller holding on to the defunct comparison as late as 1875. The power of the comparative method is to show that obvious looking cognates such as these are in fact impossible; *deus* and *θεός theós* are now the textbook example. The allofams approach, by accepting word families as given before turning to cross linguistic comparison, will never reach the point of being able to reject obvious looking cognates. As a discipline and a community we can recognize this methodological failing and together slowly move from the phase of Pott to the phase of Buck and Pokorny (1959), without this transition in any way implying ingratitude or disrespect for our Pott.

Irregularities in sound change and internal variation in proto-languages are there to be beheld as brute facts, they are not methodological insights worthy of incorporation into the practice of historical linguistics. Research on Indo-European and other well studied language families has proceeded unimpeded without recourse to allofams. The reason is that 'problems' is a better name than 'allofams' and one which encourages rather than discourages scrutiny.²⁴ The medical establishment finds it easier to tell parents that their newborns have fallen pray to 'sudden infant death syndrome' than to admit that they have suddenly died without explanation. In a similar fashion, reconstructions such as **(s-/ʔ-)p-/b-/w-u/i(·)-k* 'belly' provide no insight into the systematic relationships among the daughter languages presumed to continue it. If the current state of knowledge does not permit one to reconstruct proto-forms that lead regularly to attested forms, then it is irresponsible to reconstruct proto-forms at all. As Wittgenstein tells us, that which one cannot speak about must be passed over in silence. Matisoff warns that "the analyst must be on his guard lest he be seduced into facile speculation about possible allofamic relationships" (1978: 45); one should go further and encourage the analyst to desist from allofamic speculation altogether.

.3 Appendix: a précis of the comparative method and Indo-European ablaut

The comparative method is the primary research tool in historical linguistics, although many handbooks amply teach and discuss this method,²⁵ the technique's lack of prominence in Trans-Himalayan research is reason enough to describe the method again here in more detail. The well-known history of Grimm's law is a clear and traditional means to exhibit the method and its benefits.²⁶ This law, named after Jakob Grimm (1785–1863), although earlier

²⁴Malkiel (1954), who uses the term 'word family' in reference to the Romance languages without citing any predecessors for this terminology, appears exactly to regard the 'member of a word family' as a synonym for 'problem'. Rhetorically asking how to classify variants he answers that the "aim should be to arrange these variants in such a way as to reconcile their assumed sequence in time and space with the known hard facts of historical grammar, which, to be valid, must be based on reliable equations" (1954: 267).

²⁵For recent treatments see Weiss 2014 and Hale 2014.

²⁶Among many excellent discussions see Bloomfield 1933: 14, Hock 1991: 37-47, Campbell 2004: 156-164, Fortson 2010: 339-341, and especially Pedersen 1962: 248-262.

noted by Rasmus Rask (1787-1832), explains the development of the Indo-European stop consonants in the Germanic languages: Proto-Indo-European (PIE) voiceless stops (*p, *t, *k, *k^w) become fricatives (*f, *þ [θ], *h [x], *h^w [x^w]), voiced stops (*b, *d, *g, *g^w) become voiceless stops (*p, *t, *k, *k^w), and voiced aspirated stops (*b^h, *d^h, *g^h, *g^{wh}) become plain voiced stops (*b, *d, *g, *g^w) in Proto-Germanic (PGmc). The following table illustrates this with examples for correspondences of stops in initial position:²⁷

Sanskrit	Cl. Greek	Latin	English	PGmc	PIE
<i>pát, pad-</i>	πούς <i>poús</i> , ποδ- <i>pod-</i>	<i>pēs, ped-</i>	foot	*fōt	*pod-, ped-
<i>tráyas</i>	τρεις <i>treís</i>	<i>trēs</i>	three	*þrīz	*trejēs
<i>śná, śún-</i>	κύων <i>kúōn</i>	<i>canis</i>	hound	*hundaz	*k _u ó-, *k _{un} -
<i>katará</i>	πότερο <i>pótero</i>	<i>quo-</i>	whether	*h ^w aþeraz	*k ^w o-tero-
<i>daśan</i>	δέκα <i>déka</i>	<i>decem</i>	ten	*tehun	*dék _m t
<i>jānāti</i>	γινώσκω <i>gi-gnōskō</i>	<i>gnosco</i>	know	*knēana	√*g _{neh} ₃
<i>jīva-</i>	βίος <i>bíos</i>	<i>vīvus</i>	quick	*k ^w ik ^w az	*g ^w i _h ₃ uo-
<i>bhrātr-</i>	φράτηρ <i>phrātér</i>	<i>frāter</i>	brother	*brōþēr	*b ^h rēh ₂ tēr
<i>da-dhāti</i>	τίθημι <i>ti-thēmi</i>	<i>faciō</i>	do	*dō-	√*d ^h eh ₁
<i>haṃsá</i>	χήν <i>khén</i>	<i>ānser</i> (< *hānser)	goose	*gans	*g ^h ans-
<i>hanti</i>	θεινῶ <i>theinō</i>	<i>-fendō</i>	bane	*banō	√*g ^{wh} en

The agreement of Sanskrit, Greek, Latin, etc. on the value of a given segment against the testimony of languages such as German and English reveals the characteristic of Germanic as a family, hence Grimm's law is also called the first Germanic sound shift.

In 1862 Carl Lottner (1834-1873) catalogued the exceptions to Grimm's law. Two readers of his paper discovered patterns among these irregularities. The exceptionlessness of the resulting restatement of Grimm's law stunned the world and resulted in a generation of younger scholars, the Neogrammarians (*Junggrammatiker*), who embraced the exceptionlessness of sound laws as a principle of their discipline, and brought historical linguistics to a new height of precision and rigor. Among Lottner's exceptions Hermann Grassmann (1809-1877) noticed cases where Sanskrit and Greek lack the aspiration that would be predicted on the basis of Germanic forms (1863). Thus, the English word 'grab' predicts a Sanskrit form beginning with *g^h-, but the actual Sanskrit word is *grabh-*, and the English word 'bid (at an auction)' predicts a Greek form with *b^h- (> p^h), but the actual Greek word is *πεύθομαι peúthomai*. The explanation is that Greek and Sanskrit, independently of each other, do not permit roots with two aspirates, and when these would occur the aspiration is lost from the earlier consonant in the word (i.e. pre-Sanskrit *g^hrab^h > Sanskrit *grabh*, pre-Greek *b^hε_μd^h > *p^hε_μd^h > Greek *πεύθομαι peúthomai*). Among Lottner's exceptions that remained after Grassman's amelioration, Karl Verner (1846-1896) noticed words where the Germanic consonant was voiced rather than voiceless as predicted by Grimm's law:

Sanskrit	Cl. Greek	Latin	English	PIE
<i>upári</i>	ὑπέρ <i>h-upér</i>	<i>s-uper</i>	over /'əʊvə(ɪ)/	*upér
<i>pitár</i>	πατήρ <i>patér</i>	<i>pater</i>	father /'fa:ðə(ɪ)/	*ph ₂ tér
<i>śatám</i>	ἑκατόν <i>he-katón</i>	<i>centum</i>	hundred	*k _m tóm

The explanation is that the unexpected voicing occurs if the segment appears internal to a word and is proceeded

²⁷Note that the segment *b is notoriously weakly attested in the phonological system of the Indo-European proto-language and it is therefore difficult to find equations with all the classical Indo-European languages. Comparanda with non-initial *b are PIE √lep : Latin *labium* : PGmc. *lep- in English lip.

by an unstressed vowel, as seen in the Sanskrit cognates (1877).²⁸ The statement of this conditioning environment is subtle; it relates voicing in Germanic to Sanskrit accentuation. The success with which Verner solved the remaining exceptions to Grimm's serves as a paragon toward which researchers strive, as much now as in Verner's day.

If Grassmann and Verner had been content to remark on the interesting 'internal variations' Proto-Indo-European must have possessed, they would not have achieved these discoveries. Another case in point of neogrammarian exceptionless rigor concerning morphophonology of proto-languages is the systematization of varying allomorphs, especially in the sphere of ablaut. The PIE root *sed 'sit', for example, is reflected in e-grade full-grade *sed- in English 'sit' (cf. Latin *sedere* 'to sit', Greek ἕδρα *hédra* 'seat'), in the o-grade full-grade *sod- in English 'sat', in the lengthened (or long) e-grade *sēd- in English 'seat' (cf. Latin *sēdēs* 'seat'), in the lengthened (or long) o-grade *sōd- in Old English *sōt* > English 'soot' (stuff that sits on something), and, finally, the zero-grade *sd- in *ni-sd-o- > English 'nest' (cf. Skt. *nīḍa-* 'resting place' and *ni sad* 'to settle down'). Each of these ablaut grades have specific functions and morphophonological environments that can be systematically reconstructed using the comparative method for the Indo-European (and Germanic) proto-language based on variation within and among the daughter languages. For example, verbal adjectives formed with the suffix *-to- always trigger zero grade of the root (*g^wm-tó- 'having come' > Skt. *gatá-*, Gk. βατός *bátos*, Lat. *ventus*), whereas singular forms of the perfect always trigger o-grade of the root (*de-dork-e 'has seen' > Skt. *dadárśa*, Gk. δέδορκα *dédorka*). The apparent "variation" in the root vowel can therefore be explained in terms of specific morphological functions.

Ablaut as alternation of vowels in the same morpheme that correlates with difference in meaning were well known already at the time of the neogrammarians from languages as different as Sanskrit and Germanic, especially from the verbal system:²⁹

Sanskrit	OHG ³⁰	ModGerm	PGmc	PIE	Gloss
<i>vártā-mi</i>	<i>wird-u</i>	<i>werd-e</i>	*werþ-ō	*uért-ō	I become
<i>va-várt-a</i>	<i>ward</i>	<i>war</i>	*warþ-a	*ue-uórt-e	I became
<i>va-vṛt-ānás</i>	<i>gi-wort-an</i>	<i>ge-word-en</i>	*wurd-anaz	*urt-onós	pret.ptcp. become

From a PIE perspective, the 1sg.prs. has e-grade, 1sg. pret. o-grade, and the pret. ptcp. zero-grade.³¹ The neogrammarians also knew about ablaut in the nominal system which is evidenced by Brugmann's *Grundriss*. The discovery that ablaut in the nominal system, e.g., nom.sg. Gk. πούς *poús* < *pod-s and Lat. nom.sg. *pēs* < *ped-s, is not due to something like allofamy in the proto-language, but the result of regular morphophonological processes is one of the great feats of Indo-European linguistics in the 20th century. Based on earlier insights by Saussure (1878), Pedersen (1926), Kuryłowicz (1935), and Kuiper (1942), scholars around the Indo-European departments of Erlangen, Vienna, and Harvard systematized ablaut in the nominal system in the second half of the 20th century.³² The basic model is recognized by most scholars of Indo-European linguistics (and is, of course, being further explored and amended) today.³³ The basic word structure of Proto-Indo-European nominals is as follows:

²⁸The rule can be formulated as follows: The Proto-Germanic voiceless fricatives *f, *þ, *h, *h^w (< PIE *p, *t, *k, *k^w) and *s when immediately following an unstressed syllable in the same word, underwent voicing and became the fricatives *b [β], *d [ð], *g, *g^w and *z (with the latter rhotazing to r in most of the Germanic daughter languages, see in the main text 4.2.).

²⁹For an introduction see Fortson (2010: 75-85 for PIE morphophonology in general, 88-110 for verbal morphology).

³⁰The Old English equivalents are *weorþe*, *wearþ*, *geworden*.

³¹The alternation of the root consonant, called *Grammatischer Wechsel*, is, of course, due to Verner's law.

³²Foundational works are Rix (1965), Eichner (1973, 1974), Schindler (1967-1994), Nussbaum (1986).

³³Cf. Fortson (2010: 113-138).

reduplication	root	suffix	ending
(redupl.)	R	(S _n) ³⁴	E
stem			
word			

In principle the slots root, suffix, and ending are able of ablaut depending on the class. The different classes reconstructed for PIE are called accent-ablaut classes since stress and ablaut interact in various ways.³⁵ The main division within a paradigm of a PIE nominal is between rectus (or strong) stem, the basis for nominative (sg, du, pl), accusative (sg, du) and vocative (sg, du, pl), and oblique (or weak) stem, the basis for accusative plural and genitive, ablative, dative, instrumental (sg, du, pl).³⁶ Apart from reduplicated nouns (which are rare),³⁷ there are three basic nominals: root nouns without an overt suffix, thematic nouns with suffixes of the structure $-(C_n)o-$ and athematic nouns with suffixes of the structure $*(C)(e)C-$. Different accent-ablaut classes exist in root nouns and athematic nouns.³⁸ The major classes of root nouns can be schematized in the following way:

	o/e-type	e/ø-type	ē/e-type
rectus stem	R(ó)-E(ø)	R(é)-E	R(é̄)-E(ø)
obliquus stem	R(é)-E(ø)	R(ø)-E(full)	R(é)-E(ø)

o/e-type:³⁹ *dom-/*dem- 'house', nom.sg. *dom-s > Arm. *town*; gen.sg. *dem-s > Av. *dāng paiti-*, Gk. *δεσπότης des-pótēs* 'master of the house'; renewed gen.sg. *dm-és > Av. *nəmō*, Arm. *tan*; archaic endingless locative *dēm > Av. *dām*.

e/ø-type: *h₂ner-/*h₂nr-, 'man', rectus in nom.sg. Gk. *ἀνὴρ anér*, acc. sg. Skt. *náram*; obliquus in gen.sg. Gk. *ἀνδρός andrós*, inst.pl. *nṛbhīh*.

ē/e-type (rare): *h₃rēǵ-/*h₃reǵ-, rectus in nom.sg. *h₃rēǵ- > Skt. *rāj*, Lat. *rēx*, OIr. *rí* 'king'; lengthened grade usually generalized throughout the paradigm.

Turning to athematic nouns, there are four major accent-ablaut classes:⁴⁰

	acrostatic	proterokinetic	hysterokinetic	amphikinetic
rectus stem	R(ó)-S(ø)-E(ø)	R(é)-S(ø)-E(ø)	R(ø)-S(é)-E(ø)	R(é)-S(o)-E(ø)
obliquus stem	R(é)-S(ø)-E(ø)	R(ø)-S(é)-E(ø)	R(ø)-S(ø)-E(é)	R(ø)-S(ø)-E(é)

Acrostatic:⁴¹ *uód-r/*uéd-n- 'water' is a so-called heteroclitic noun where not only ablaut varies between rectus and

³⁴Brackets indicate optionality. The little n is variable since potentially there can be more than one suffix in a nominal.

³⁵The exact phonological origin of this interaction is still a matter of debate.

³⁶There are archaic endingless locatives that are distinguished from the rectus and obliquus stem by a different ablaut grade, usually full or lengthened grades where the respective obliquus has zero or full grade.

³⁷Cf. Masculine *k^we-k^wl-os 'wheel' > Gk. *κύκλος kúklos* : PGmc *h^weh^wlaz (> OE *hwēol* > English 'wheel') and neuter plural collective *k^we-k^wl-éh₂ 'wheelage' > Skt. *cakrá* : PGmc. *h^weula- (> ON *hjól*) : Gk. *κύκλα kúkla* (with short -a < *h₂ in analogy to the athematic neuters). Note the different PGmc outcomes due to Verner's law.

³⁸Ablaut in thematic nouns is a topic of ongoing investigations, see especially Nussbaum 2017.

³⁹The o/e-ablaut is limited to roots with a stop after the ablauting vowel *-ET(T), whereas roots with a resonant after the ablauting vowel *-ER(C) replaced o/e- with o/ø-ablaut, see in the main text.

⁴⁰Concerning the terminology, Kuiper's statement from his classic paper still has something to be said for it: "The terms are little pompous to my taste, but it will be best to retain them." (Kuiper 1942: 4)

⁴¹There is a second acrostatic type R(é)-S(ø)-E(ø) / R(é)-S(ø)-E(ø), *ǵérh₂-s-/*ǵérh₂-s-, rectus stem in Gk. *γῆρας géras* 'old age', obliquus stem in Gk. *γῆρας géras* 'gift of honor'. This is an example of a paradigmatic split in which each levelling of an ablaut grade produced two different paradigms with two separate meanings, cf. Old English *stæf*, pl. *stafas* [stavas] > English *staff* and *stave*.

obliquus stem, but also the suffix. This alternation of ablaut and stem-consonant is preserved within one paradigm with the nom.sg. *uód-ŕ in Hitt. *wātar* and gen.sg. *uéd-ŋ-s in Hitt. *witenaš*. Another example is *nók^w-t-/*nék^w-t- 'night' with the rectus stem nom.sg. *nók^w-t-s in Gk. νύξ *núks*, Lat. *nox* and the obliquus stem gen.sg. *nék^w-t-s in Hitt. *nekuz* (*mēhur*) '(time) of the evening'.

Proterokinetic: *péh₂-uŕ/*ph₂-uén- 'fire', the nom.sg. *péh₂-uŕ is directly attested in Hitt. *pahhur* and the gen.sg. *ph₂-uén-s in Hitt. *pahhuenaš*. Another example is *g^wén-h₂/*g^wn-éh₂- 'woman' with the rectus stem nom.sg. *g^wén-h₂ in Skt. *jáni-*, Toch. B *šana*, OIr. *ben* and the obliquus stem *g^wn-éh₂- in Skt. *gná-*, Toch. B *snoy*, OIr. *mná*. Note that the paradigms in Tocharian B and Old Irish are synchronically irregular and that there is a paradigm split in Sanskrit.

Hysterokinetic: *(h₃)udn-éj/*(h₃)udn-i- 'land', nom.sg. *(h₃)udn-éj is directly attested in Hitt. *utnē* and the gen.sg. *(h₃)udn-ĭ-és in Hitt. *utniyaš*. Another example is *ph₂-tér-/*ph₂-tr- 'father', nom.sg. *ph₂-tér (< *-tér-s) in Skt. *pitá*, OAv. *ptā*, Gk. πατήρ *patér*, Lat. *pater*, obliquus stem in dat.sg. Skt. *pitré*, OAv. *piθrē* and gen.sg. Gk. πατρός *patrós*, Lat. *patris*.

Amphkinetic: *d^hég^h-om/*d^hg^h-m- 'earth', nom.sg. *d^hég^h-om is directly attested in Hitt. *tēkan* and the gen.sg. *d^hg^h-m-és in Hitt. *taknaš*. Another example is *pént-oh₂-/*pnt-h₂- 'path', nom.sg. *pént-oh₂-s in Skt. *pánthās*, OAv. *pañtā* and gen.sg. *pnt-h₂-és in Skt. *pathás*, OAv. *paθō*. Note the synchronically irregular alternation of the medial consonants in both languages that are diachronically due to the different ablaut grades producing different phonological environments.

These accent-ablaut classes mostly survive as archaisms or irregularities within the synchronic grammars of the Indo-European daughter languages, but are recoverable with the comparative method from variation found among them.⁴² Tracing *pod- and *ped- or *uód-ŕ and *uéd-ŋ- back into the proto-language as variant forms without at least trying to systematically connect them would be multiplying entities beyond necessity.

The commen

⁴²The accent-ablaut classes are morphologically linked by the process of internal derivation. This is derivation by switching the accent-ablaut pattern, and not by the addition of overt morphology. Internal derivation is synchronically found in Anatolian, Indo-Iranian and Greek and can be reconstructed for PIE: akrostatic : proterokinetic (e.g., *póh₂-u- / *péh₂-u- in Gk. πῶν *pôu* 'flock' : *péh₂-i-u- / *ph₂-i-éu- in Skt. *pāyú-*, *pāyán-* 'guardian'); akrostatic : amphikinetic (e.g., *sóċ-ŕ / *sék-ŋ- in Hitt. *šakkar* 'excrement' : *sék-ōr / *sċ-n- in Gk. σῶρ 'dung'); proterokinetic : hysterokinetic (e.g., *sċéh_x-mŋ / *sċuh_x-mén- in Skt. *syúma* 'band' : *sċuh_x-mén / *sċuh_x-mn- in Gk. ὑμήν *humén* 'membrane'); proterokinetic : amphikinetic *h₁éj-tŕ / *h₁i-tén- in Lat. *iter*, *itin-* 'way' : *h₁éj-tōr / *h₁i-tn- in Toch. A *ytār* 'way, road'); hysterokinetic : amphikinetic (e.g., *ph₂-tér, *ph₂-tr- in Gk. πατήρ *patēr* 'father' : *péh₂-tōr, *ph₂-tr- in Gk. ἀ-πάτωρ *apátōr* 'having no father'). Internal derivation in Indo-European linguistics recognized since Schmidt (1889) and elucidated by Jochem Schindler in 1970s (see Nussbaum 2014: 238ff. with literature) is typologically not uncommon (cf. English *record* /ˈɪkɔːd/ : *record* /ɪˈkɔːd/; Arabic *šārib* 'drinker' : collective *šarb*; Lith. *áukštas* 'high' : *aũkštas* 'growth'; Mandarin 好 *hǎo* 'good' : 好 *hào* 'be fond of').

Acknowledgements

HF would like to acknowledge the generous support of Austrian Science Found (FWF): project number Y1044-G30.

NH would like to acknowledge the generous support of the European Research Council, under the auspices of 'Beyond Boundaries: Religion, Region, Language and the State' (ERC Synergy Project 609823 ASIA).

NH and HF would like to thank Laura Grestenberger and Randy LaPolla for comments on earlier drafts of this paper.

References

- Abondolo, Daniel (1998). 'Introduction.' *The Uralic languages*. Daniel Abondolo (ed.), London: Routledge.
- Anthony, David W. (2007). *The horse, the wheel, and language: How bronze-age riders from the Eurasian steppes shaped the modern world*. Princeton: Princeton University Press.
- Baxter, William H. (1992). *A handbook of Old Chinese phonology*. Berlin: Mouton de Gruyter.
- Baxter, William H. and Laurent Sagart (2014). *Old Chinese: A new reconstruction*. Oxford: Oxford University Press.
- Beekes, Robert S. P. (1987). 'On Indo-European 'wine'.' *Münchener Studien zur Sprachwissenschaft* 48: 21–26.
- Benedict, Paul K. (1972). *Sino-Tibetan: a Conspectus*. Cambridge: at the University Press.
- Benedict, Paul K. (1975). 'Where it all began: memories of Robert Shafer and the 'Sino-Tibetan Linguistics Project.' *Linguistics of the Tibeto-Burman Area* 2.1: 81-92.
- Bloomfield, Leonard (1933). *Language*. New York, H. Holt and Company.
- Brugmann, Karl (1886-1893). *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen*. Vol. 1-2. Straßburg: Trübner.
- Chang Kun (1972). 'Sino-Tibetan `iron': *qhleks'. *Journal of the American Oriental Society* 92: 436-46.
- Chang Kun (1973). Review of Benedict 1972. *The Journal of Asian Studies* 32.2: 335-337.
- Campbell, Lyle (2004). *Historical linguistics: an introduction*. Edinburgh: Edinburgh University Press.
- Campbell, Lyle and William J. Poser (2008). *Language Classification: History and Method*. Cambridge: Cambridge University Press.
- Conrady, August (1896). *Eine indochinesische Causativ-Denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten*. Leipzig: O. Harrassowitz
- Davies, Anna Morpurgo (1998). *History of Linguistics. Vol IV: Nineteenth-Century Linguistics*. London: Longman.
- van Driem, George (2012). 'The Trans-Himalayan phylum and its implications for population prehistory.' *Communication on Contemporary Anthropology* 5: 135-142.
- van Driem, George (2018). 'Linguistic history and historical linguistics.' *Linguistics of the Tibeto-Burman Area* 41.1: 106–127.
- Eichner, Heiner (1972). 'Die Etymologie von heth. *mēhur*.' *Münchener Studien zur Sprachwissenschaft* 31: 53-107.
- Eichner, Heiner (1974). 'Zu Etymologie und Flexion von vedisch *strī* und *pūmān*.' *Die Sprache* 20: 26-42.
- Fortson, Benjamin W. (2010). *Indo-European language and culture: an introduction*. Malden: Blackwell.
- Gong Hwang-cherng (1980). 'A comparative study of the Chinese, Tibetan, and Burmese vowel systems.' *Bulletin of the Institute of History and Philology* 51.3: 455-490. (reprinted in:) *Collected Papers on Sino-Tibetan Linguistics*. Taipei: Zhongyang yanjiuyuan yuyanxue yanjiusuo choubichu (中央研究院語言學研究所籌備處), 2002, 1-30.
- Gong Hwang-cherng (1995). 'The system of finals in Proto-Sino-Tibetan.' *The ancestry of the Chinese language*. William S. Y. Wang (ed.), Berkeley: Project on Linguistic Analysis, University of California: 41-92. (reprinted in:) *Collected papers on Sino-Tibetan linguistics*. Taipei: Zhongyang yanjiuyuan yuyanxue yanjiusuo choubichu (中央研究院語言學研究所籌備處), 2002: 79-124.
- Grassmann, Hermann (1863). 'Ueber die aspiraten und ihr gleichzeitiges vorhandensein im an- und auslaute der wurzeln.' *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete des Deutschen, Griechischen und Lateinischen* 12.2: 81-110.
- Hale, Mark (2007). *Historical linguistics: Theory and method*. Oxford. Blackwell
- Hale, Mark (2014). 'The Comparative Method: theoretical issues.' *The Routledge Handbook of Historical Linguistics*. Claire Bower and Bethwyn Evans (eds.), Abingdon-on-Thames: Routledge. 146-160.
- Handel, Zev (2008). 'What is Sino-Tibetan? Snapshot of a Field and a Language Family in Flux.' *Language and Linguistics Compass* 2: 422–441.
- Haudricourt, André-George (1954a). 'De l'origine des tons en vietnamien.' *Journal Asiatique* 242: 69–82.
- Haudricourt, André-George (1954b). 'Comment reconstruire le chinois archaïque.' *Word* 10.2-3: 351–364.
- Hill, Nathan W. (2005). 'The verb 'bri 'to write' in Old Tibetan.' *Journal of Asian and African Studies* 68: 177-182.
- Hill, Nathan W. (2009). 'Review of Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction. By James A. Matisoff. Berkeley: University of California Press, 2003.' *語言暨語言學 / Languages and Linguistics* 10.1: 173-195.

- Hill, Nathan W. (2011a) 'Alternances entre *h* et *b* en tibétain ancien et dans les langues tibétaines modernes.' *Revue d'Etudes Tibétaines* 20: 115-122.
- Hill, Nathan W. (2011b). 'An inventory of Tibetan sound laws.' *Journal of the Royal Asiatic Society of Great Britain & Ireland* (Third Series) 21.4: 441-457.
- Hill, Nathan W. (2013) 'Relative order of Tibetan sound changes affecting laterals.' *Language and Linguistics* 14.1: 193-209.
- Hill, Nathan W. (2014) 'A Note on Voicing Alternation in the Tibetan Verbal System.' *Transactions of the Philological Society* 112.1: 1-4.
- Hock, Hans Heinrich. (1991). *Principles of historical linguistics*. Berlin: Mouton de Gruyter.
- Hock, Hans Heinrich and Brian D. Joseph. (2009). *Language history, language change and language relationship: an introduction to historical and comparative linguistics*. Berlin: Mouton de Gruyter.
- Hogg, Richard M. (1992). *A grammar of Old English. Volume 1: Phonology*. Oxford: Blackwell.
- Jacques, Guillaume (2007). 'A shared suppletive pattern in the pronominal systems of Chang Naga and Southern Qiang.' *Cahiers de linguistique - Asie orientale* 36.1: 61-78.
- Jäschke, Heinrich August (1881). *A Tibetan-English dictionary, with special reference to the prevailing dialects*. London: Routledge and Kegan Paul.
- Jensen, Frede (1999). *A Comparative Study of Romance*. Bern: Lang.
- Joseph, John E. (2002). *From Whitney to Chomsky: essays in the history of American linguistics*. Amsterdam: John Benjamins.
- Karlgren, Bernhard (1933). 'Word families in Chinese.' *Bulletin of the Museum of Far Eastern Antiquities* 5: 9-120.
- Kroeber, Alfred A. (1938). 'Editorial preface.' *Sino-Tibetica* 1: 1-3.
- Kuiper, Franciscus B. J. (1942). *Notes on Vedic Noun-Inflection*. (Mededeelingen der nederlandse akademie van wetenschappen, afdeling letterkunde, nieuwe reeks, deel 5, No. 4.) Amsterdam.
- Kuryłowicz, Jerzy (1935). *Études indoeuropéennes*. Kraków: Gebethner & Wolff.
- LaPolla, Randy J. (2003). "Overview of Sino-Tibetan morphosyntax." *The Sino-Tibetan Languages*. Graham Thurgood and Randy J. LaPolla, eds. London: Routledge, 22-42.
- LaPolla, Randy J. (2016). 'Once again on methodology and argumentation in linguistics: Problems with the arguments for recasting Sino-Tibetan as "Trans-Himalayan".' *Linguistics of the Tibeto-Burman Area* 39.2: 282-297.
- Li Fang-Kuei (1959). 'Tibetan Glo-ba-'dring.' *Studia Serica Bernhard Karlgren dedicata*. Søren Egerod and Else Glahn (eds.), Copenhagen: Ejnar Munksgaard, 55-59.
- Li Fang-Keui (1988). *Linguistics East and West: American Indian, Sino-Tibetan, and Thai. With an Introduction by George Taylor. Interviews conducted by Ning-Ping Chan and Randy LaPolla*. Berkeley: The Regents of the University of California.
- Li Liu and Xingcan Chen (2012). *The Archeology of China: From the Late Paleolithic to the Early Bronze Age*. Cambridge: Cambridge University Press.
- Lottner, Carl. (1862). 'Ausnahmen der ersten lautverschiebung.' *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete des Deutschen, Griechischen und Lateinischen* 11.3/4: 161-205.
- Malkiel, Yakov (1954). 'Etymology and the Structure of Word Families.' *Word* 10.2/3: 265-274.
- Matisoff, James A. (1972). *The Loloish tonal split revisited*. Berkeley: Center for South and Southeast Asia Studies, University of California.
- Matisoff, James A. (1975). 'Benedict's Sino-Tibetan: a rejection of Miller's Conspectus inspection.' *Linguistics of the Tibeto-Burman Area* 2.1, 155-72.
- Matisoff, James A. (1978). *Variational Semantics in Tibeto-Burman: the 'organic' approach to linguistic comparison*. (Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics, Volume VI.) Philadelphia: Publication of the Institute for the Study of Human Issues (ISHI).
- Matisoff, James A. (1982). 'Proto-languages and proto-Sprachgefühl.' *Linguistics of the Tibeto-Burman Area* 6.2.1-64.
- Matisoff, James A. (1990). 'On megalocomparison.' *Language* 66.1 (1990): 106-120
- Matisoff, James A. (2000). 'On 'Sino-Bodic' and other symptoms of neosubgroupitis.' *Bulletin of the School of Oriental and African Studies* 63.3: 356-69.
- Matisoff, James A. (2003). *Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibeto-Burman reconstruction*. Berkeley: University of California Press.
- Matisoff, James A. (2007). 'Response to Laurent Sagart's review of *Handbook of Proto-Tibeto-Burman: System and*

- philosophy of Sino-Tibetan reconstruction.* *Diachronica* 24.2: 435–444.
- Matisoff, James A. (2008). *The Tibeto-Burman reproductive system: toward an etymological thesaurus*. Berkeley: University of California Press.
- Meillet, Antoine (1925). *La méthode comparative en linguistique historique*. Oslo: H. Aschehoug & co.
- Miller, Roy Andrew (1956). 'The Tibeto-Burman ablaut system.' 国際東方学会議紀要 *Kokusai Tōhō Gakusha Kaigi kiyō / Transactions of the International Conference of Orientalists in Japan* 1: 29-56.
- Miller, Roy Andrew. (1974). 'Sino-Tibetan: Inspection of a Conspectus.' *Journal of the American Oriental Society* 94.2: 195-209.
- Nishi, Yoshio 西 義郎 (1974). 'Birumabungo-no-ac-ni tsuite' (ビルマ文語の-ac について) [On -ac in Burmese]. *Tōyō gakuho* (東洋学報) 56.1: 1-43.
- Nishi Yoshio 西 義郎 (1975). 'OB ry- ni Tsuite (1) (OB ry- について (1)) [On OB ry- (1)].' *Kagoshima Daigaku Shiroku* (鹿児島大學史録) 8: 1-16.
- Nishida Tatsuo (1977). 'Some problems in the comparison of Tibetan, Burmese and Kachin Languages.' *Studia phonologica* (音声科学研究) 11: 1-24.
- Nussbaum, Alan J. (1986). *Head and Horn in Indo-European*. Berlin: Walter de Gruyter.
- Nussbaum, Alan J. (2014). 'Greek τέχμαρ 'sign' and τέχμωρ 'sign': Why both?' *Das Nomen im Indogermanischen. Morphologie, Substantiv versus Adjektiv, Kollektivum. Akten der Arbeitstagung der Indogermanischen Gesellschaft vom 14. bis 16. September 2011 in Erlangen*. Norbert Oettinger und Thomas Steer (eds.), Wiesbaden: Reichert, S. 215-260.
- Nussbaum, Alan J. (2017). 'Agentive and Other Derivatives of "τόμος-Type" Nouns.' *Verbal Adjectives and Participles in Indo-European Languages. Proceedings of the conference of the Society for Indo-European Studies (Indogermanische Gesellschaft), Paris, 24th to 26th September 2014*. Claire Le Feuvre, Daniel Petit and Georges-Jean Pinault (eds.), Bremen: Hempen. 233-266.
- Pedersen, Holger (1962). *The discovery of language: linguistic science in the nineteenth century*. John Webster Spargo, Trans. Bloomington: Indiana University Press.
- Pedersen, Holger (1926). *La cinquième déclinaison latine*. (Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser, XI, 5.) København: Høst & Søn.
- Peiros, Ilia and Sergei Starostin. (1996). *A Comparative Vocabulary of Five Sino-Tibetan Languages*. Melbourne: The University of Melbourne, Department of Linguistics and Applied Linguistics.
- Ringe, Donald (2006). *A History of English: From Proto-Indo-European to Proto-Germanic*. Oxford: Oxford University Press.
- Rix, Helmut (1965). 'Lat. *iecur iocineris*.' *Münchener Studien zur Sprachwissenschaft* 18: 79-92.
- Rubio, Gonzalo (1999). 'On the alleged 'Pre-Sumerian substratum'.' *Journal of Cuneiform Studies* 51: 1-16.
- Sagart, Laurent (2006). 'Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibeto-Burman reconstruction. By James A. Matisoff.' *Diachronica* 23.1: 206–223.
- de Saussure, Ferdinand (1878). *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. Leipzig: Teubner.
- Schindler, Jochem (1967a). 'Das idg. Wort für „Erde“ und die dentalen Spiranten.' *Die Sprache* 13: 191-205.
- Schindler, Jochem (1967b). 'Zu hethitisch *nekuz*.' *Zeitschrift für Vergleichende Sprachforschung* 81: 290-303.
- Schindler, Jochem (1969). 'Die idg. Wörter für Vogel und Ei.' *Die Sprache* 15: 144-167.
- Schindler, Jochem (1972a). *Das Wurzelnomen im Arischen und Griechischen*. Unpublished Ph.D. thesis, Würzburg.
- Schindler, Jochem (1972b). 'L' apophonie des noms-racines indo-européens.' *Bulletin de la Société de Linguistique* 67: 31-38.
- Schindler, Jochem (1975a). 'L'apophonie des thèmes indo-européens en -r/n.' *Bulletin de la Société de Linguistique* 70: 1-10.
- Schindler, Jochem (1975b). 'Zum Ablaut der neutralen s-Stämme des Indogermanischen.' *Flexion und Wortbildung. Akten der V. Fachtagung der Indogermanischen Gesellschaft, Regensburg, 9.–14. Sept. 1973*. Helmut Rix (ed.), Wiesbaden: Reichert, 259-67.
- Schindler, Jochem (1975c). 'Armenisch *erkn*, griechisch *ὀδύνη*, irisch *idu*.' *Zeitschrift für Vergleichende Sprachforschung* 89: 53-65.
- Schindler, Jochem (1994). 'Alte und neue Fragen zum indogermanischen Nomen.' *In honorem Holger Pedersen. Kolloquium der Indogermanischen Gesellschaft 1993 in Kopenhagen*. Jens Elmegård Rasmussen (ed.), Wiesbaden:

- Reichert, 397-400.
- Schmidt, Johannes. (1889). *Die Pluralbildungen der indogermanischen Neutra*. Weimar: Böhlau.
- Schmitt, Rüdiger (1981). *Grammatik des Klassisch-Armenischen mit sprachvergleichenden Erläuterungen*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Schuessler, Axel (2007). *Etymological Dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.
- Schuessler, Axel (2009). *Minimal Old Chinese and later Han Chinese*. Honolulu: University of Hawai'i Press.
- Shafer, Robert (1966-74). *Introduction to Sino-Tibetan*. Wiesbaden: Harrassowitz.
- Simon, Walter (1929). 'Tibetisch-Chinesische Wortgleichungen, ein Versuch.' *Mitteilungen des Seminars für Orientalische Sprachen an der Friedrich Wilhelms-Universität zu Berlin* 32: 157-228.
- Starostin, George (2009). Review of Scheussler 2007. *Journal of Language Relationship* 1: 155-162.
- Starostin, Sergej Anatol'evič (Сергей Анатольевич Старостин) (1989). *Rekonstrukcija drevnekitajskoj fonologičeskoj sistemy* [Reconstruction of Old Chinese phonological system] (*Реконструкция древнекитайской фонологической системы*). Moscow: "Наука." Glavnaja redakcija vostočnoj literatury. (Главная редакция восточной литературы "Наука.")
- Stein, Rolf (1941). 'Notes d'étymologie tibétaine.' *Bulletin de l'École française d'Extrême-Orient* 41: 203-231
- Uray, Géza (1952). 'A Tibetan diminutive suffix.' *Acta Orientalia Academiae Scientiarum Hungaricae* 2: 182-220
- Verner, Karl A. (1877). 'Eine Ausnahme der ersten Lautverschiebung.' *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprache* 23.2: 97-130
- Wagner, Donald B. (1993). *Iron and steel in ancient China*. Leiden: E.J. Brill.
- Weiss, Michael (2009). *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor: Beech Stave Press
- Weiss, Michael (2014). 'The Comparative Method.' *The Routledge Handbook of Historical Linguistics*. Claire Bowerman and Bethwyn Evans (eds.), Abingdon-on-Thames: Routledge. 127-145.
- Wolfenden, Stuart N. (1928). 'The prefix m- with certain substantives in Tibetan.' *Language* 4.4: 277-280.
- Wolfenden, Stuart N. (1936). 'On certain alternations between dental finals in Tibetan and Chinese.' *Journal of the Royal Asiatic Society of Great Britain and Ireland (New Series)* 68.3: 401-416.
- Wolfenden, Stuart N. (1937). 'Concerning the variation of final consonants in the word families of Tibetan, Kachin, and Chinese.' *Journal of the Royal Asiatic Society of Great Britain and Ireland (New Series)* 69: 625-655.
- Wolfenden, Stuart N. (1938). 'Concerning the origins of Tibetan brgiad and Chinese 丩 'pwât'.' *T'oung Pao* 34.3: 165-73.