

The historical phylogeny of Albanian, Part I

Olav Hackstein, Ludwigs-Maximilians-Universität München

olav.hackstein@lmu.de

Roundtable “Albanian among the Indo-European Languages: origins and antiquity”

Tirana, Academy of Sciences, 4 November 2024

Abstract

The paper deals with the possibility of prehistoric contact phenomena and areal convergence between adjacent dialects of Proto-Indo-European. The Uniformitarian Hypothesis implies that the reconstructed Proto-Indo-European parent language, as a natural language, must have had dialects that were in contact with each other. Prehistoric areal contacts between these dialects can therefore be posited. Recent scholarship has identified successive layers of such prehistoric contacts.

My contribution intends to shed light on the earliest layers of areal contact, when the Proto-Indo-European dialects of the Pre-Palaeo-Balkan languages were in contact with the Pre-Proto-Tocharian dialect. I will point to a number of non-trivial, pervasive innovations that distinguish Pre-Proto-Tocharian and Pre-Palaeo-Balkan Indo-European from each other.

The linguistic evidence thus supports the “Anatolian-first and the Tocharian-second Hypothesis”, and further substantiates that Balkan-Indo-European sides *pari passu* with the remaining Indo-European phyla.

1 PIE dialects and early areal contacts: hypotheses and methodology

The purpose of the present article is to discuss the latest work on the phylogenetic models of the diversification of the Indo-European language family, focusing in particular on the models proposed by Don Ringe and his collaborators (Ringe et al. 2002, Ringe 2017, CPHL archive), and by Paul Heggarty et al. (2023).

Both the the “Ringe tree” (as dubbed by Heggarty et al. 2023b: 71) and Heggarty’s model (Heggarty 2023a) converge in the three-step model of the Indo-European branches splitting off from Proto-Indo-European, assigning the first split-off to Anatolian, placing Tocharian in second position, and lastly having the remaining Indo-European languages begin their migration.

(1) Compare the “Ringe tree” (so dubbed by Heggarty et al. 2023: 71), in a binary branching structure (Ringe 2017: 69):

(Anat (Toch ((Ital, Celt) ((Gmc, Alb) ((Gk, Arm) (Ilr, BSl)))))),
and as a cladistic tree, reported in Ringe, Warnow and Taylor 2002, fig. 8:

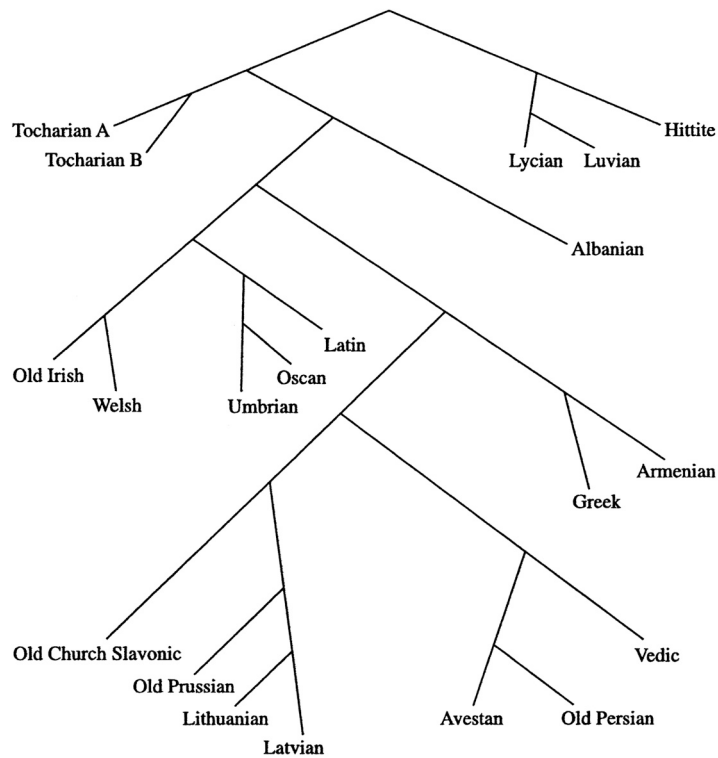
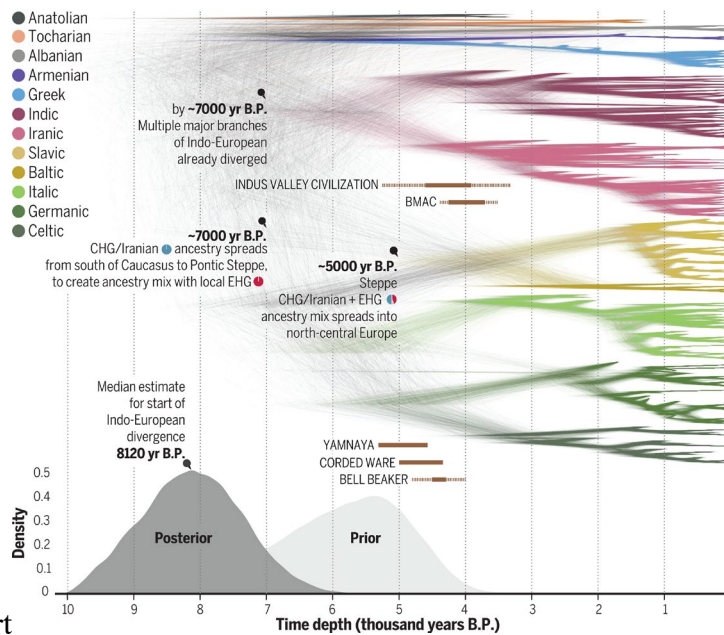


Figure 8. One of the best trees with Germanic omitted.



(2) Heggart

y et al. (2023a):

Compare further the binary tree in Heggarty et al. (2023b: 71):

(((((Indo-Iranic, Balto-Slavic) (Greek, Armenian)) (Italic, Celtic))), Tocharian), Anatolian)

“As presented in Fig. 2 in (5), the Ringe topology does not in fact include a definite position for Albanian or Germanic, however. We therefore implemented a series of backbone

constraints to enforce a topology of ((((((Indo-Iranic, Balto-Slavic) (Greek, Armenian)) (Italic, Celtic))), Tocharian), Anatolian), while allowing Germanic and Albanian to move freely.”

The trees devised by Ringe’s and Heggarty’s respective teams feature Tocharian and Balkan Indo-European as consecutive branching clades, thus assigning a particular antiquity to the diversification of the three Balkan Indo-European languages Albanian, Armenian, Greek (in alphabetical order, with no implication for a hierarchy in antiquity) from the remaining Indo-European languages.

1.1 Anatolian-first and Tocharian-second

The debate surrounding the “Anatolian-first and Tocharian-second” hypothesis is still running, but despite claims to the contrary, there is firm evidence in favor of it, and the evidential basis may even be further corroborated.

Anatolian and Tocharian share a host of morphological and syntactic parameters that are absent from Nuclear Indo-European, including Albanian and the other Indo-European languages:

(3) Morphological features (5+): *h₂e*-conjugation (Jasanoff 2003); rarity and limited development of simple thematic verbs (Ringe 2000, Jasanoff 2018: 234); collective inflection (Hackstein 2012: 154–5); agglutinative case morphology in the ablative and allative cases (Hackstein 2007, 2018); and an important addendum: Anatolian and Tocharian lack the long-vowel subjunctive.

Syntactic features (6+): SOV word order; postpositions (Hackstein 1997); postposed apposition (Hackstein 2003b, 2010: implication SOV → linearization ((hyponym) hyperonym), cf. Cinque 2023: 38–9, 124–5); correlative construction PIE **k^wo ... so-/to- ...* (Watkins 1976: 312–5); lack of compound verbs (Hackstein 2023 and see below).

Occasional lexical correspondences, such as,¹ e.g., **(a)** in the term for ‘hunger’, Hitt. *k/gašt* c. ‘hunger; famine’ & Toch. B *kest*, A *kašt* as opposed to Lat. *sitis* ‘thirst’, Gk. *phthisis*, Ved. *kṣiti-* ‘waning, decline, destruction’; **(b)** Hitt. *eku-* & Toch. AB *yok-* ‘drink’ (PIE **h₁eg^{wh}-*) as opposed to innovative Nuclear IE **peh₃-* (Krause/Thomas 1960: 57, Weiss 2018: 373 contra Malzahn 2016: 287); **(c)** Anatolian-Tocharian lacks the semantic development of **b^huh₂-* ‘grow, become’ to a light (and eventually an auxiliary) verb, Toch. AB *pyutk-* (< **b^heyh₂-t-(s)ke-* < **b^heyh₂-d-ske-* with *d*-enlargement as in Slavic, OCS *bŏdŏ*); but cases like the foregoing can only claim significance as part of cumulative evidence.

¹ Malzahn 2016 argued that there is no lexical evidence that would prove an early split-off of Tocharian. But cases like Toch. B *kest*, A *kašt* or Toch. AB *yok-* ‘drink’ definitely point into the opposite direction. Of course, the caveat applies that lexical isoglosses, if inherited and claimed to be archaisms, provide only but weak evidence for subgrouping. It cannot be excluded that the other language branches have lost the selfsame lexical items so that the alleged evidence for lexical archaisms becomes an *argumentum e silentio*. But an accumulation of such evidence of course substantiates the evidence and reduces the chance of a fallacious *argumentum e silentio*.

Remarkable phonological archaisms of Anatolian and Tocharian include **(a)** the preservation of the dental plus velar onset in the lexeme for ‘earth’, Hitt. *tekan* & Toch. A *tkam* and the “joint failure of Anatolian and Tocharian to form thorn clusters” (Jasanoff 2017: 234); **(b)** the Tocharian preservation of the dental in Toch. *antapi* ‘both (sides)’ as against its loss in Nuclear Indo-European, e.g., OHG *umbi*, OIr. *imb-*, Umbr. *amb-*, Gk. ἀμφί etc. (Jasanoff 1976: 125).

None of the aforementioned features is shared by Balkan-Indo-European including Albanian.

1.2 Tocharian-second and agreements with Nuclear Indo-European

But, on the other hand, a further host of features sets off Tocharian from Anatolian.

(4) Tocharian diverging from Anatolian and siding with the Nuclear Indo-European languages as a “Classical Indo-European” language.

Contrasting with Anatolian, Tocharian shows the following Nuclear Indo-European features:

- activization of the Anatolian stative *nt*-participle, contrasting with mediopassive *mh₁no*-participle, in in Tocharian and Nuclear Indo-European (Fellner & Grestenberger 2018).
- the resultative-stative IE perfect (pace Jasanoff 2018: 234 not a “Nuclear IE innovation”, see Hackstein 2024, e.g., Toch. B *aiku* ‘having become acquired & being known’ from PIE **h₂e-h₂ik̑-uōs* ‘id’ x **h₂e-h₂oik̑-e* → ***h₂o-h₂oik̑-*, **h₂o-h₂ik̑-*; A **kakanu* > *kaknu*, B *kekenu* ‘having been generated and extant (with); provided with’; Homeric Greek γένε- ‘produce’ γεγάώς, γεγονώς ‘having become, being’).

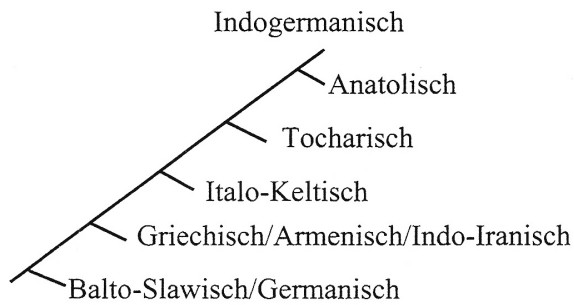
See further the inventories compiled by Jasanoff 2018: 234 and Weiss 2018: 373.

To sum up, Tocharian assumes a second position intermediate between Anatolian on the one hand and Nuclear Indo-European on the other.

1.3 Janus-faced Tocharian: Tocharian-second and Tocharian & Balkan Indo-European isoglosses

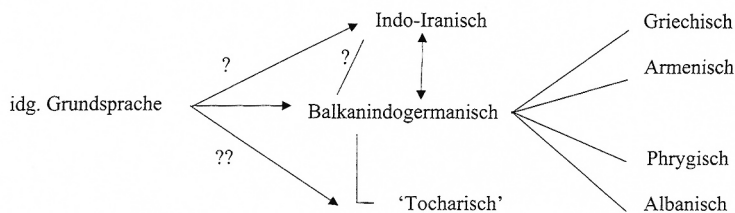
Having dealt with the respective cladistic positions of Anatolian and Tocharian, a third question concerns the cladistic position of Balkan-IE, comprising Greek, Armenian, Albanian and Phrygian.

A loose connection of Balkan Indo-European with Tocharian has been suspected since Pokorny 1919: 38, with Klingenschmitt 1994: 244 diagnosing a certain overlap of Tocharian with this Balkan Indo-European group (“gewisse Übereinstimmungen mit der Balkangruppe”). Matzinger 2005a: 383, Schindler apud Matzinger 2012 devised a time-space model, involving three stages of split-offs: Anatolian first, Tocharian second, Italo-Celtic third, Greek and Armenian later:



(Schindler apud Matzinger 2012:143)

This model assumes “a Tocharian tinge” for historical Balkan Indo-European:



(Matzinger 2012: 144)

Innovations shared by Tocharian with Balkan-Indo-European

(5) An **affiliation of Tocharian with Balkan-Indo-European** can be substantiated on the basis of several shared complex morphological innovations, as mentioned in Hackstein 2024:

(5+ features, e.g.): (a) Toch.B *klowotär*, A *klawatär* ‘is called’ & Toch. Alb. *quhet* non-active 3sg. ‘is called’ < PIE $*\hat{k}lu\bar{u}-eh_1-\bar{i}e/o-$, “uridg. Essivverb”; cf. Matzinger 2005b: 46, Schumacher & Matzinger 2013: 993; (b) **innovative factitive suffix $*-\bar{n}-\bar{i}e/o-$** in the Tocharian subjunctive [7] and present [12] & in Albanian and Greek, Alb. *bën* ‘does, makes’, Greek φαίνει ‘brings to light, causes to appear’ < post-IE $*b^h h_2-\bar{n}-\bar{i}e/o-$ (Schumacher & Matzinger 2013: 49); (c) **nominal collective suffix $*-a\bar{i}-$ [$*-eh_2i-$]** in Tocharian A (voc. *klai*, poss. in THT1283a7, Itkin 2022: 98) & Alb. *grā* ‘women’, Greek γύναι ‘oh guild of women’; (d) **polylexical patterns:** Toch.B *kuse* ‘who, what?’ < PIE $*k^w is + so$ & Alb. *kush* ‘who?’ < PIE $*k^w os + so(-s)$, Greek τίς δ’ ὅδε; (e) **preverbal reflexive particle:** Alb. *u* < PIE $*sue/suoi$ (Schumacher & Matzinger 2013: 45, 48) & Toch. A *su* (< PIE $*sue$) (Hackstein 2024: 154);

In the innovative clustering of pronominal stems, Tocharian accords in nontrivial fashion with Greek and Armenian (see Hackstein 2002, 2012: 142):

(a) Toch. B complementizer *kuce* < $*k^w id\ te h_1$ and *ce* < $*te h_1$ ‘thus’ & Armenian *e+ t^ce* ,that’ < $*k^w e(h_1)\ te(h_1)$ ‘how so > that’, Myc. *ho-te*, Homeric ὅ-τ(ε) < $(H)io-te(h_1)$.

(b) Toch. A *ku-pre* ‘when’ < $*k^w i-b^h ro-\bar{i}$ & Arm. *e-rb* ‘when?’ < $*k^w e-b^h ro-$; Homeric ὥρα ‘as long as’ < $(H)io-b^h r-t$.

1.4 Contrary observation: Balkan Indo-European and Albanian are classical Indo-European languages

Conversely, however, the Balkan-Indo-European languages do not bear the slightest resemblance to the Anatolian and Tocharian type of archaic morphosyntax, but instead resemble classical or —especially in the case of Albanian — even post-classical Indo-European languages.

The most important morphological and syntactic features that distinguish Anatolian and Tocharian from Nuclear Indo-European listed above are not shared by Balkan Indo-European. Rather, Balkan Indo-European patterns with Nuclear Indo-European and the classical Indo-European languages in this regard.

What is more, there are significant pervasive innovations that clearly establish that Balkan Indo-European patterns with Nuclear Indo-European. I name two examples.

1.4.1 Pervasive innovations: rise of verbal compounds and productivity *s*-mobile roots (*s*-prefix) in Nuclear Indo-European, including Balkan Indo-European

A prominent isogloss that clearly marks the Anatolian and Tocharian branches as archaic and **sets them off from all other Indo-European branching, including Balkan-Indo-European** has so far gone unnoticed: the rise of verbal compounds and the productivity of the so-called Proto-Indo-European *s*-mobile roots exhibiting the accretion of an *s*-prefix.

The formation of verbal compounds (i.e., preverb-verb stems involving the prefixation of local particles) is an innovation in the IE languages. It is in its infancy and **unproductive in Anatolian and Tocharian**, but tremendously productive in the post-Anatolian and post-Tocharian Nuclear Indo-European branches including Celtic, Germanic, Baltic, Slavic, Italic, Albanian, Greek, Armenian and (later) Indo-Iranian.

Crucially there is a correlation between prefixation of local particles (LP prefixation) which has spread in Nuclear Indo-European and the productivity of “*s*-mobile”. “*s*-Mobile” denotes a morphophonemic anomaly primarily concerning verbal roots which may feature the onset accretion of an extrasyllabic phoneme /s/.

Languages like Anatolian and Tocharian, with unproductive LP-prefixation, exhibit fewer cases of mobile *s*-. The inverse holds for languages like Albanian, Greek, Germanic and Slavic, with a wealth of *s*-mobile roots.

The dichotomy between (Anatolian and) Tocharian **non-s-roots**, versus Albanian and Nuclear Indo-European ***s*-mobile roots** can be illustrated by the following non-exhaustive list of cases:

SCARCITY/ ABSENCE OF MOBILE <i>S</i> -PREFIX IN TOCHARIAN	PRODUCTIVITY OF MOBILE <i>S</i> -PREFIX IN NUCLEAR INDO-EUROPEAN
VELAR ONSETS	
PIE <i>*kedh₂</i> - ‘spread out’: Toch.AB <i>kātā</i> -	Gk. σκίδνῃμι, σκεδάννυμι, Alb. <i>hedh</i> ‘throw away/off’ < <i>*skouđ-éje-</i> , and OGeg. <i>hjedh</i> ‘id.’ < <i>*skedh₂-e/o-</i> < <i>*s-kedh₂-e/o-</i> (aliter <i>s</i> -mobile: Schumacher and Matzinger 2013: 976, Neri 2021a in DPEWA s.v. <i>hedh</i> ~ <i>hjedh</i>)
PIE <i>*kel(h₃)</i> - ‘sting, prick’: Toch.B <i>kälts</i> - ‘oppress’ (Malzahn 2010: 591–2)	PIE <i>*s-kel(h₃)</i> - : Alb. <i>halë</i> f. < PIE <i>*skol-ne-h₂</i> - f. ‘sharp item, bone, awn’, <i>*skl̥-(m)n-ó-</i> ‘splitting, cutting, sharp’, <i>*skel-</i> ‘split’ (LIV ² 552–3, Demiraj, Imberciadori 2023a in DPEWA s.v. <i>halë</i>); Alb. <i>hell</i> ‘pike’ < PIE <i>*skólh₃o-</i> ‘tip’, PIE <i>*s-kelh₃</i> - ‘split off’ (Demiraj, Neri 2021d in DPEWA s.v. <i>hell</i> assuming <i>s</i> -mobile); Alb. <i>hollë (i)</i> ‘thin’ < PIE <i>*skēl(h₃)-ó-</i> ‘a little thin, split’, PIE <i>*s-kelh₃</i> - ‘split off’ (Demiraj, Imberciadori 2023b in DPEWA s.v. <i>hollë (i)</i>)
PIE <i>*kers-</i> ‘cut’: Toch. <i>kärsā</i> - ‘know’	PIE <i>*s-ker-</i> : NHG <i>scheren</i> ‘shear’
PIE <i>*g^wes-</i> ‘be extinguished, wane’: Hitt. <i>kašt</i> , Toch. A <i>kašt</i> , B <i>kest</i> ‘hunger’, AB <i>käs-</i>	PIE <i>*s-g^wes-</i> ‘extinguish’: Gk. σβέννυμι

Labial onsets

PIE <i>*perH-</i> ‘fly’: Toch.B <i>prere</i> , A <i>pär</i> ‘arrow’	PIE <i>*s-perH-</i> ‘fly’ (Schumacher & Matzinger 2013: 233): Alb. <i>fier</i> ‘far’ (< PIE <i>*sperH-om</i>)
PIE <i>*pek̥-</i> ‘observe’: Toch. B <i>pāk-</i> ‘look for, intend’	PIE <i>*s-pek̥-</i> ‘look out’: Lat. <i>speciō</i>
PIE <i>*pelH-</i> ‘speak’: Toch.B. <i>pällātär</i> ‘praise’, <i>pele</i> ‘language’; and as denominative derivative (from PIE <i>*pel-ueh₂</i> ‘loud call’) Hitt. <i>palūae-</i> ‘recite’	PIE <i>*s-pelH-</i> ‘speak out (aloud)’: Goth. <i>spill</i> ‘tale’, Eng. <i>spell</i> ‘incantation’, Alb. <i>fjalë</i> ‘word’ (Demiraj, Neri 2020 in DPEWA s.v. <i>fjalë</i> , Arm. <i>arā-spel</i> ‘legend, saying, riddle’ (Hackstein 2023: 21 ex. 26))
PIE <i>*penh₁-</i> ‘span’: Toch. B <i>peññ-</i> , A <i>panw-</i> (Hackstein 1995:23, 26–7)	PIE <i>*s-penh₁-</i> : E. <i>span</i> , G. <i>spannen</i> (LIV ² 578)

Coronal onsets

PIE <i>*teh₂</i> - ‘take a step’: Toch.AB <i>tā(kā)</i> -, Hittite <i>tiya/e^(mi)</i> ‘come to a stand, place oneself’, Old Irish <i>ad-tá</i> ‘is (there)’	PIE <i>*s-teh₂</i> :- Alb. <i>shton</i> ‘*stand out in size; augment, increase’, Lat. <i>stāre</i> , Gk. ἵσταμαι ‘I
--	--

	step’ < PIE * <i>s-teh</i> ₂ - (Hackstein 2023: 96–7, cf. 2023: 22–4 ex. 29)
PIE * <i>teup</i> - ‘press’: Toch.A impv. <i>p-täpsäs</i> ‘destroy’	PIE * <i>s-teup</i> -: Alb. <i>shtyp</i> ‘crumble, beat, push, print (out)’, Lat. <i>stupēre</i> ‘be struck’ < PIE * <i>s-tup</i> - (Schumacher & Matzinger 2013: 216, Hackstein 2023: 95–6, 2023: 24 ex. 33, Neri 2021c in DPEWA s.v. <i>shtyp</i>).
PIE * <i>tu-n-d</i> -: Toch.B <i>tämts</i> - ‘beat, scatter’ (Hackstein 2003c: 183, Malzahn 2010: 651)	PIE * <i>s-tu-n-d</i> - ‘beat off’: Alb. <i>shtyn</i> ‘push away’ < PIE * <i>s-tud-n</i> - (Schumacher & Matzinger 2013: 999, Neri 2021d in DPEWA s.v. <i>shtyn</i> , Hackstein 2022: 95)
PIE * <i>tep</i> - ‘step’: Toch.B * <i>tep(p)</i> -, <i>cepy</i> - ‘step forward’ (Malzahn 2010: 636–7, Adams 2013: 323)	PIE * <i>s-tep</i> - ‘step out’: E. <i>step</i> -, NHG <i>stapfen</i> ‘trudge’
PIE * <i>turH-s</i> - ‘begin to stir up’: Toch.A <i>trus</i> - ‘tear up’ < * <i>truH-s</i> - < PIE * <i>turH-s</i> -	PIE * <i>s-tur</i> -: E <i>stir</i> -, NHG <i>stören</i> from Gmc. * <i>stur-ja-na</i> ‘stir up and out’, PIE * <i>s-tur</i> -

Liquid onsets

PIE * <i>lejd</i> ^h - ‘slide’: Toch.A <i>litkoş</i> ‘sunk/slid down’	PIE * <i>s-lejd</i> ^h - ‘slide away’: E. <i>slide</i> , NHG <i>schlittern</i> (Krisch 1990)
PIE * <i>neh</i> ₂ - ‘bath’: Toch.B <i>nāsk</i> - ‘bath away, swim’	PIE * <i>s-neh</i> ₂ -: Ved. <i>snā</i> - (Hackstein 1995: 172–3), OIr. <i>snaid</i> ‘swims’
PIE * <i>neh</i> ₁ - ‘sew’: Toch.A <i>nāskā</i> - (<i>nāske</i> ‘I’ve sewn’, <i>nānāsku</i> , Pinault 2001: 132–3)	PIE * <i>s-neh</i> ₁ - ‘sew (together)’: Toch.A <i>šñor</i> ‘sinew’, Gk. (* <i>v</i>)νητός ‘sewn, heaped up’ (Hom. <i>Od.</i> 2.338), OIr. <i>sniid</i> ‘sews’

Demiraj, Bardhyl, Sergio Neri. 2020. *ffjal|ë -a*. In: DPEWA. URL: <https://www.dpwa.gwi.uni-muenchen.de/dictionary/?lemmaid=14091> (accessed 01 November 2024).

Demiraj, Bardhyl, Giulio Imberciadori. 2023a. *hálë -a*. In: DPEWA. URL: <https://www.dpwa.gwi.uni-muenchen.de/dictionary/?lemmaid=22375> (accessed 20 March 2024).

Demiraj, Bardhyl, Giulio Imberciadori. 2023b. *hóllë (i)*. In: DPEWA. URL: <https://www.dpwa.gwi.uni-muenchen.de/dictionary/?lemmaid=22214> (accessed 20 March 2024).

Neri, Sergio. 2021a. *hedh ~ hjedh*. In: DPEWA. URL: <https://www.dpwa.gwi.uni-muenchen.de/dictionary/?lemmaid=18478> (accessed 07 March 2024).

Neri, Sergio. 2021b. *hënë*. In: DPEWA. URL: <https://www.dpwa.gwi.uni-muenchen.de/dictionary/?lemmaid=18264> (accessed 23 March 2024).

Neri, Sergio. 2021c. *shtyp*. In: DPEWA. URL: <https://www.dpwa.gwi.uni-muenchen.de/dictionary/?lemmaid=14208> (accessed 18 March 2024).

Neri, Sergio. 2021d. *shtyn*. In: DPEWA. URL: <https://www.dpwa.gwi.uni-muenchen.de/dictionary/?lemmaid=18458> (accessed 18 March 2024).

1.4.2 Pervasive innovations: innovative polylexical negation

In addition to the scarcity of *s*-mobile roots in Anatolian and Tocharian as against their frequent occurrence in Nuclear Indo-European, including Albanian, **a further significant innovation that distinguishes Anatolian and Tocharian from Albanian and Nuclear Indo-European concerns innovative polylexical negators.**

As illustrated in Table 1, Anatolian and Tocharian preserve three morphemes (negation **ne*, **h₁ara*, **h₂ieui*). Greek and Albanian clip the negation (**Ø h₁ara h₂oju*). Additionally, Albanian, like the rest of Nuclear Indo-European, drops the particle **h₁ara* (**h₂oju-k^wid*).

Table 1: Temporal and areal transition in the expression of complex negation (PIE **ne* + **h₁ara* + **h₂ieui(i)/h₂oju*) (Hackstein 2024: 157).

ANATOLIAN	Hitt. <i>nāwi</i>	<i>*ne</i>	Ø	<i>h₂ieui</i>
	Hitt. <i>natta āra</i>	<i>*ne</i>	<i>h₁ara</i>	Ø
TOCHARIAN	Toch.A <i>māryu</i>	<i>*ne</i>	<i>h₁ara</i>	<i>h₂ieui</i>
BALKAN INDO-EUROPEAN (GREEK, ALBANIAN)	Gk. ἄρα οὐ	Ø	<i>h₁ara</i>	<i>h₂oju</i>
	Gk. οὐκί, Alb. <i>as, s-</i>	Ø	Ø	<i>h₂oju-k^wid</i>
NORTHWEST INDO-EUROPEAN (Germanic, Baltic, Slavic)	OHG <i>ni ju</i> , Lith. <i>niaĩ</i> , <i>nejaĩ(gi)</i> , OCS <i>ne</i> <i>(j)u(-že)</i>	<i>*ne</i>	Ø	<i>h₂ieui</i>
	Goth. <i>ni aiw</i> , OHG <i>ni</i> <i>io</i> , Mod. Germ. <i>nie</i>	<i>*ne</i>	Ø	<i>h₂ieuiom</i>

Albanian thus occupies an intermediate position between Greek and the rest of Nuclear Indo-European:

- 1 Anatolian
- 2 Tocharian
- 3a Greek
- 3b Albanian**
- 4 Rest of Nuclear Indo-European

2 Resolving the contradiction: Indo-European dialects and prehistoric areal contacts

How are we to reconcile the “Tocharian-second” model with the convergences between Tocharian and North-West Indo-European languages, on the one hand, and the convergences between Tocharian and the Balkan-Indo-European languages, Greek and Albanian, on the other?

The convergences between Tocharian and Balkan IE are remnants of **earlier prehistoric areal contact phenomena**. Here, the PIE dialect continuum model (Hackstein 2005) is more adequate.

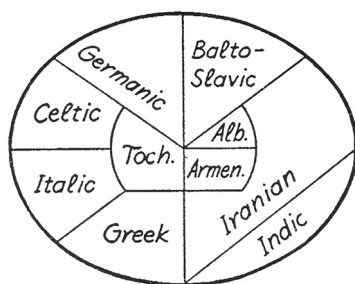
Proto-Indo-European, if assumed to be a natural language, can be safely assumed to have had dialects, and it is likewise natural for adjacent dialects to develop and exhibit areal contact phenomena. This accords with the Uniformitarian Principle: “The general processes and principles which can be noticed in observable history are applicable in all stages of language history” (Hock 2021: 908).

Linguistic diversification does not preclude prehistoric areal contacts.

It would be fallacious to take every linguistic convergence to imply a shared node/clade and phylum. Compare, for instance, the shared feature of the *have*-perfect in Standard Average European: the Scandinavian languages, Albanian and Modern Greek all share this feature, though it certainly does not originate from a Scandinavian-Balkan Indo-European clade.

Meillet 1908

The idea of setting up a prehistoric Balkan Indo-European (sub)group has been entertained for a long time. Already in 1908, Meillet cautiously observed some linguistic similarities between Albanian, Armenian, Greek, and Tocharian.



(Meillet 1908: 134)

According to Meillet, it is likely that Tocharian takes an intermediate position between Italic-Celtic, Greek, Armenian and Slavic, but it remains difficult to render this more precise or ascertain it:

“Il est vraisemblable que le tokharien occupe, parmi les dialectes indo-européens, une place intermédiaire entre l’italo-celtique, le grec, l’arménien et le slave; mais on ne saurait rien préciser ni rien affirmer.” (Meillet 1912:142)

Adams 1984

Adams' 1984 investigation led to "the grouping of Tocharian with Meillet's "Northwestern" group in the first instance, with particularly close ties with Germanic. The ties Tocharian has with both Greek and Indic outside the Northwestern group reflect later contacts on the part of the pre-Tocharians with the pre-Greeks and pre-Indics as the pre-Tocharians moved progressively eastward in the late Proto-Indo-European world." (Adams 1984: 395)

"it might not be too out of the line to suggest that that variety of late Proto-Indo-European that was to become Proto-Tocharian first took shape and became in some measure distinguishable from other varieties of Proto-Indo-European in the course of the fourth millenium B.C. amidst the other members of Meillet's Northwestern Group (**to which we should also add Albanian ...**)" (Adams 1984: 401)

Tocharian and Northwest-Indo-European prehistoric areal contacts: a non-exhaustive list

Tocharian exhibits some connections to the Western group (Sieg 1908, Porzig 1954: 213-4, Ringe 2018: 63). Here are 7 examples:

Toch. B *ente* 'where, when' and OHG *huanta* 'whence, why' (from **k^wo-m-d^heh₁*, **k^wo-m-d^hoh₁*, Hackstein 2017).

Toch. B *trāpp-*, A *trāp-* 'anstoßen, straucheln' (Krause/Thomas 1960: 67 [§45,3a], 68 [§46,2]) and Gmc., NHG *treffen* 'hit', OIc. *drepa* 'strike'.

Toch. B *pārsk-* 'be afraid' (Krause/Thomas 1960: 68, Hackstein 1995: 192-194), deponent verb from **p_ṛk-ske-tor*; PIE root perhaps **perg-* 'beat' (*EWAh* 3, 480, cf. Arm. *harkanem* 'beat (up)'); **p_ṛk-tó-* 'beaten; afraid', **p_ṛk-teh₂* 'fear' > OHG *forahta* (f./ō), *forahten*. Contrast PIE **d_ʰuej(-s)-* (Hom. δέιδω, arm. *erkeay*), **b^hejH-* (Lith. *bijoti(s)*, OCS *bojati se*, Ved. *bháyate*) in other IE languages.

Toch. B *ñakte*, A *ñkät* 'god' and Gmc., E. *god* 'deity which receives a libation' (**ni-ǵhu-tóm*, *ni-ǵhu-teh₂*; cf. Watkins 1974, Hackstein 2006, and recently Jackson & Oettinger 2014: 122).

Toch. B *pilta*, A *pält* 'leaf', Germ. *Blatt* (Krause & Thomas 1960: 52, Adams 2013: 415).

Toch. *klänk-* 'bend, turn oneself', *klenke* '(vehicle) which is steered' < **klongo-* (Adams 2013:245, Hilmarsson 1996:149; on the morphological type, see Malzahn 2013:67), NHG *lenken*, *Gelenk* '(flexible) joint', MHG *lenken* 'bend', OHG *ir(h)lenken* 'twist, dislocate (a joint)'.

Productivity of *-n*-stem inflection shared by Tocharian and Germanic (Adams 1988: 398–9).

A non-exhaustive list of Greek-Tocharian isoglosses & prehistoric areal contacts

Tocharian likewise exhibits connections to Greek and Balkan Indo-European (Adams 1984, Nikolaev 2022, Hackstein 2024).

Gk. and Toch. genitive dual in **-o_ṛsi* > Homeric Gk. *-ouv*, Toch. *-(n)ais* (Winter 1962: 126, Adams 1984: 398, 1988: 164–5).

Dor. Gk. υἱός 'son' and Toch. B *soy* 'son' (PIE **suh_{2/3}-iu-*) with suffix **-iu-* instead of **-nu-* (as found in Gmc., Slav. Ilr.; Ringe 1996: 61–3).

This isogloss is highly significant: “Since the *borrowing* of inflectional morphology, to say nothing of inflectional categories (as opposed to their *persistence* in wholesale language shift) is at best extremely rare (cf., e.g., Sankoff 2002: 658), these two innovations are clinching” (Ringe 2008: 63).

There is an abundance of Tocharian-Greek lexical isoglosses.

Gk. παντ-, Toch. *pont-* ‘all, every, each’ (Imberciadori 2024/5: 147–55).

Gk. ὄει, Toch. AB *swā-s-*.

Gk. κεφαλή, Toch. A *špāl* ‘head’, OHG *gibil* ‘braincase’, *gibilla* ‘skull, head’ (Krause/Thomas 1960: 66, Imberciadori 2024/5: 727–9).

Gk. κίς, κίως to be connected to PIE **kih₂-* ‘move’ and Toch. A *šāyū* ‘creepy crawler’ (PIE **kih₂-iō-*, Toch. A *šāyā-* ‘crawling leg’ + possessive suffix **-uōnt-?*, Pinault 2004: 258).

A recent dossier of Tocharian-Greek isoglosses is provided by Nikolaev 2022: 25n.16 (“retentions from PIE”). Here are 9 examples:

Gk. βλωθρός ‘(grown) high’ (< **m̥h₃dhro-*) and Toch. A *malto* ‘at first, in first place’, Toch. B *melte* ‘pile’ (< **mol(h₃)dho-* ‘top’); see Fraenkel 1932: 6, Imberciadori 2024/5: 347–9.

Gk. ἀταλός ‘tender’, ἀ(τι)τάλλω ‘to bring up a child’ and Toch. A *ātāl* ‘man, male’; see Klaus T. Schmidt *apud* Balles 1999: 3 n. 2, Hackstein 2003a: 56 A10 (the reconstruction **h₂et-l̥*, **h₂etelo-* is further supported by Proto-Germanic **apala-n* / **apilan* ‘descent, lineage’ and HLuwian *atala-* ‘brother’).

Gk. θέλω ‘to enchant, beguile’ (< **d^helg-*) and Toch. A *talke*, B *telki* ‘sacrifice’ (< **d^holgo-*), IEW 247.

Gk. λυγαλέος ‘wretched, unhappy, sore’, λυγρός ‘mournful’ and Toch. B *lakle* ‘pain, suffering’ (< **luglo-* or **lugro-* with assimilation of *l...r* to *l...l*).

Gk. λιμὴν ‘harbor; market place’ and Toch. A *lyām*, B *lyam* ‘sea’ (< **limen-*).

Gk. μάχομαι ‘fight’ and Toch. B *māk(ā)-* ‘to run’ (see Malzahn & Peters 2010).

Gk. νάρκη ‘numbness, paralysis’ and Toch. B *nerke* ‘hesitation’, *närk-* ‘to keep away’ (see Malzahn 2010: 686).

Gk. ὀλίγος ‘small; few’ and Toch. A *lykäly*, B *lykaške* ‘small’ (< Proto-Tocharian **lyäk-* < **h₃lig-* ‘little’, further supported by Armenian *alk* ‘at’ ‘poor’; Imberciadori 2024/5: 196–201).

Gk. ὄρχος ‘row of vines or trees’ and Toch. B *waršiye* ‘(grove of) trees’, *warke* ‘garland’ (< **uerg^h-*, possibly supported by Albanian *varg* ‘chain, string of beads’).

Albanian is part of Nuclear Indo-European

Like the convergences between NW Indo-European, Greek, and Tocharian, so too can the convergence between Albanian, Greek and Tocharian be the result of prehistoric areal contacts. The pervasive innovations discussed above in 1.4.1 and 1.4.2 identify the Balkan Indo-European languages, and thus Albanian, as classical, i.e., Nuclear Indo-European languages. Albanian may even appear to be fairly innovative; cf. (a) the reduction of the inventory of participles inherited from Proto-Indo-European (present **-nt-*, mediopassive **-mh₁no-*, stative **-uōs/uos/us-*) is complete in Albanian, contrasting with Greek and many

other branches of Indo-European which have kept the *nt*-participle down to the present day; **(b)** ablaut is no longer productive in Albanian (Matzinger 2016: 83).

References

- Adams, Douglas Q. (1984). The position of Tocharian among the other Indo-European languages. *Journal of the American Oriental Society* 104: 395–402.
- Adams, Douglas Q. (1988). *Tocharian Historical Phonology and Morphology*. New Haven: Connecticut.
- Adams, Douglas Q. (2013). *A Dictionary of Tocharian B. Revised and greatly enlarged*. Leiden: Rodopi.
- Cinque, Guglielmo (2023). *On Linearization: Toward a Restrictive Theory*. Cambridge, Massachusetts: The MIT Press.
- CPHL = *Computational Phylogenetics in Historical Linguistics*; project and archive by Don Ringe, Tandy Warnow, Luay Nakhleh, Steven N. Evans; <https://www.cs.rice.edu/~nakhleh/CPHL/>
- DPEWA = *Digitales Philologisches Wörterbuch des Altalbanischen, Digital Philological-Etymological Dictionary of Old Albanian* (15th-18th c. CE) [Fjalori digjital filologjik-etimologjik i shqipes më të vjetër në shek. XV-XVIII], under the supervision of Bardhyl Demiraj, Olav Hackstein as principal investigators, and the collaboration of Giulio Imberciadori, Sergio Neri, Anila Omari et al. Online <https://www.dpwa.gwi.uni-muenchen.de/>.
- Fellner, Hannes, Laura Grestenberger (2018). Die Reflexe der *-nt- und *-mh₁no-Partizipien im Hethitischen und Tocharischen. In *100 Jahre Entzifferung des Hethitischen. Morphosyntaktische Kategorien in Sprachgeschichte und Forschung*. Akten der Arbeitstagung der Indogermanischen Gesellschaft vom 21. bis 23. September 2015 in Marburg. Herausgegeben von Elisabeth Rieken unter der Mitwirkung von Ulrich Geupel und Theresa Maria Roth. Wiesbaden: Reichert. 63–82.
- Hackstein, Olav (1995). *Untersuchungen zu den sigmatischen Präsensstammbildungen des Tocharischen*. Göttingen: Vandenhoeck & Ruprecht.
- Hackstein, Olav (1997). Präverb, Post- und Präposition im Tocharischen: Ein Beitrag zur Rekonstruktion urindogermanischer Syntax. *Tocharian and Indo-European Studies* 7: 35–60.
- Hackstein, Olav (2002). A New Study of Classical Armenian *eʿe* (*tʿe*). Paper presented to the 21th East Coast Indo-European Conference, Philadelphia, University of Pennsylvania, 15 June 2002.
- Hackstein, Olav (2003a). Zur Entwicklung von Modalität in Verbaladjektiven. In *Indogermanisches Nomen. Derivation, Flexion, Ablaut*. Herausgegeben von Eva Tichy, Dagmar S. Wodtko, Britta Irslinger. Bremen. 51–66.
- Hackstein, Olav (2003b). Apposition and word-order typology in Indo-European. In: Brigitte Bauer & Georges-Jean Pinault (eds.), *Language in Space and Time. A Festschrift für Werner Winter on the Occasion of his 80th Birthday*. Berlin, New York. 131–152.
- Hackstein, Olav (2003c). Review of “Douglas Q. Adams, A Dictionary of Tocharian B. Amsterdam 1999. (Leiden Studies in Indo-European 10).” *Indo-Iranian Journal* 46,2: 177–189.
- Hackstein, Olav (2007). Ablative Formations. In *Verba Docenti. Studies in historical and Indo-European linguistics presented to Jay H. Jasanoff by students, colleagues, and friends*. Edited by Alan J. Nussbaum. Ann Arbor: Beech Stave Press. 131–153.
- Hackstein, Olav (2012). The evolution of finite complementation in Tocharian. In: *Tocharian and Indo-European Studies* 13. 117–148.
- Hackstein, Olav (2017). The morphological and constructional evolution of OHG *huuanta* and Dutch *want*. *Indo-European Linguistics* 5. 1–30. <https://doi.org/10.1163/22125892-00501004>.

- Hackstein, Olav (2018). Allative Formations: Homeric Greek ἄλλυδις ἄλλη, Old Latin *alii aliā* and Congeners. In: *Vina Diem Celebrant: Studies in Linguistics and Philology in Honor of Brent Vine*. Edited by Dieter Gunkel, Stephanie W. Jamison, Angelo O. Mercado, and Kazuhiko Yoshida. New York, Ann Arbor: Beech Stave Press. 93–106.
- Hackstein, Olav (2023). When words coalesce II: Preverb incorporation in Indo-European. *Indo-European Linguistics* 11: 1–40. <https://doi.org/10.1163/22125892-bja10027>.
- Hackstein, Olav (2024a). Albanian and Balkan Indo-European. In *The Albanian Language Area and its Surroundings from Late Antiquity to the High Middle Ages. Proceedings of the 7th German-Albanian Cultural Studies Conference (21.-22. April 2023, Hubmersberg/ Pommelsbrunn)*. Herausgegeben von Bardhyl Demiraj in der Reihe Albanische Forschungen, Bd. 50. Wiesbaden: Harrassowitz. 143–166.
- Hackstein, Olav (2024b). The East Tocharian Śroṇa-Koṭikarṇa-Avadāna: Linguistic Remarks. Talk presented to the *Kyōto Symposium of Silk Road Studies—from India to China via Turfan*. Kyōto University. 5 October 2024.
- Heggarty, Paul et al. (2023a). Language trees with sampled ancestors support a hybrid model for the origin of Indo-European languages. *Science* **381**, eabg0818. DOI: [10.1126/science.abg0818](https://doi.org/10.1126/science.abg0818).
- Heggarty et al. (2023b). Supplementary Materials for Language trees with sampled ancestors support a hybrid model for the origin of Indo-European languages. *Science* 381, eabg0818 (2023). DOI: [10.1126/science.abg0818](https://doi.org/10.1126/science.abg0818).
- Hock, Hans Henrich (2021). *Principles of Historical Linguistics*. Third edition, revised and updated. Berlin, New York: Mouton de Gruyter.
- Imberciadori, Giulio (2024/5, forthcoming). *Etymologische Untersuchungen zum System der tocharischen Adjektive*. 2 Bände. Inauguraldissertation zur Erlangung des Doktorgrades der Philosophie an der Ludwigs-Maximilians-Universität München 2023. Under revision for print.
- Jackson, Peter, Norbert Oettinger (2014). Gott. *Münchener Studien zur Sprachwissenschaft* 68/1: 119–125.
- Jasanoff, Jay H. (1976). Gr. ἄμφο, lat. *ambō* et le mot indo-européen pour l'un et l'autre. *Bulletin de la Société de Linguistique* 71,1: 123–131.
- Jasanoff, Jay H. (2003). *Hittite and the Indo-European Verb*. Oxford: Oxford University Press.
- Jasanoff, Jay H. (2018). The impact of Hittite and Tocharian: Rethinking Indo-European om the 20th century and beyond. In *Handbook of Historical and Comparative Indo-European Linguistics*, edited by Jared Klein, Brian Joseph, Mark Wenthe. (= HSK 41.3). Berlin, Boston: De Gruyter, Mouton. 220–238.
- Itkin, Ilya (2022). Nouvelles formes tokhariennes B extraites des fragments de Berlin. I. Quelques formes nominales nouvelles ou « oubliées ». *Tocharian and Indo-European Studies* 21: 97–105.
- Klingenschmitt, Gert (1994). Die Verwandtschaftsverhältnisse der indogermanischen Sprachen. In: Jens E. Rasmussen (ed.), *In honorem Holger Pedersen*. Wiesbaden: Reichert. 235–251.
- Krause, Wolfgang, Werner Thomas (1960). *Tocharisches Elementarbuch. Band I Grammatik*. Heidelberg: Carl Winter.
- Krisch, Thomas (1990). Zur Etymologie von nhd. *gleiten* (mit Bemerkungen zu weiteren Etymologien). *Historische Sprachforschung* 103: 116–131.
- LIV² = *Lexikon der indogermanischen Verben. Die Wurzeln und ihre Primärstammbildungen*. Zweite, erweiterte und verbesserte Auflage bearbeitet von Martin Kümmel und Helmut Rix. 2001. Wiesbaden: Reichert.
- Malzahn, Melanie (2010). *The Tocharian Verbal System*. Leiden, Boston: Brill.
- Malzahn, Melanie (2016). The second one to branch off. The Tocharian lexicon revisited. In *Etymology and the European lexicon. Proceedings of the 14th Fachtagung der Indogermanischen Gesellschaft 17-22 September*

- 2012, *Copenhagen*. Edited by Bjarne Simelkjaerr Sandgaard Hansen, Benedicte Nielsen Whitehead, Thomas Olander, Birgit Anette Olsen. Wiesbaden: Reichert. 281–292.
- Matzinger, Joachim (2005a). Phrygisch und Armenisch. In *Sprachkontakt und Sprachwandel*. Edited by Gerhard Meiser and Olav Hackstein. Wiesbaden: Reichert. 375–394.
- Matzinger, Joachim (2005b). Messapisch und Albanisch. In *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 2: 29–54.
- Matzinger, Joachim (2012). Zwischensprachen: Areallinguistische Bemerkungen aus dem Bereich des Balkanindogermanischen. In *Iranistische und Indogermanistische Beiträge in Memoriam Jochem Schindler (1944-1994)*, edited by Velizar Sadovski and David Stifter. Wien: Österreichische Akademie der Wissenschaften. 137–159.
- Matzinger, Joachim (2016). *Die sekundären nominalen Wortbildungsmuster im Altalbanischen bei Gjon Buzuku. Ein Beitrag zur altalbanischen Lexikographie*. Wiesbaden: Harrassowitz.
- Nikolaev, Alexander (2022). Greek μάρπτο and Tocharian A märk-. *Cuadernos de Filología Clásica. Estudios griegos e indoeuropeos* 32: 23–28.
- Pinault, Georges-Jean (2001). Nouveautés lexicales et morphologiques dans le manuscrit de Yanqi du Maitreyasamiti-Nātaka en tokharien A. In *Norm und Abweichung. Akten des 27. Deutschen Orientalistentages (Bonn, 28. September bis 2. Oktober 1998)*, hrsg. von Stefan Wild und Hartmut Schild, Würzburg, Ergon Verlag. 121–136.
- Pinault, Georges (2004) Zum Tocharischen in der Turfanforschung. In *Turfan Revisited – The First Century of Research into the Arts and Cultures of the Silk Road*, ed. by Desmond Durkin-Meisterernst et al. Berlin: Dietrich Reimer Verlag. 256–263.
- Pokorny, Julius (1919). Die Stellung des Tocharischen im Kreise der indogermanischen Sprachen. *Berichte des Forschung-Institutes für Osten und Orient* 3. Wien. 24–57.
- Porzig, Walter (1954). *Die Gliederung des indogermanischen Sprachgebiets*. Zweite, unveränderte Aufl. 1974. Heidelberg: Carl-Winter.
- Ringe, Don (1996). *On the Chronology of Sound Changes in Tocharian. Volume I: From Proto-Indo-Europeans to Proto-Tocharian*. New Haven: Connecticut.
- Ringe, Don (2000). Tocharian class II presents and subjunctives and the reconstruction of the Proto-Indo-European verb. *Tocharian and Indo-European Studies* 9: 121–142.
- Ringe, Don (2018). Indo-European dialectology. In *Handbook of Historical and Comparative Indo-European Linguistics*, edited by Jared Klein, Brian Joseph, Mark Wenthe. (= HSK 41.3). Berlin, Boston: De Gruyter, Mouton. 1716–1731.
- Schumacher, Stefan, & Joachim Matzinger (2013). *Die Verben des Altalbanischen. Belegwörterbuch, Vorgeschichte und Etymologie*. Wiesbaden: Harrassowitz.
- Sieg, Emil, Wilhelm Siegling (1908). Tocharisch, die Sprache der Indoskythen. Vorläufige Bemerkungen über eine bisher unbekannte indogermanische Literatursprache. *Sitzungsberichte der Berliner Akademie der Wissenschaften* 39: 915–932.
- Southern, Mark R. V. (1999). *Sub-Grammatical Survival: Indo-European s-mobile and its Regeneration in Germanic*. Washington: Institute for the Study of Man.
- Watkins, Calvert (1976). Towards Proto-Indo-European syntax: problems and pseudo-problems. In S. Steever (ed.), *Papers from the Parasession on Diachronic Syntax*. Chicago: Chicago Linguistic Society. 305–326.
- Weiss, Michael (2018). Tocharian and the West. In *Priscis Libentius et Liberius Novis. Indogermanische und sprachwissenschaftliche Studien. Festschrift für Gerhard Meiser zum 65. Geburtstag*. Herausgegeben von Olav

Hackstein und Andreas Opfermann unter Mitarbeit von Harald Bichlmeier und Sabine Häusler. Hamburg: baar Verlag. 373–381.

Winter, Werner (1962). Nominal and pronominal dual in Tocharian. *Language* 38: 111–134.

Handout including full bibliography available at:

URL <https://www.indogermanistik.uni-muenchen.de/personen/professoren/hackstein/publikationen/index.html>

QRC

