

ARTICLE

Lād “law”: a Bactrian loanword in the Nuristani languages

Jakob Halfmann 

University of Würzburg, Würzburg, Germany
Email: jakob.halfmann@uni-wuerzburg.de

Abstract

This article proposes a new etymology for the Nuristani word family of Katë *lod* ~ *lot*, Nuristani Kalasha *lād*, etc. It is argued that these are best understood as early borrowings from Bactrian λαδο “law”.

Keywords: Bactrian; Nuristani; Middle Iranian; Etymology; Afghanistan

It has long been noted that the Nuristani languages, alongside their inherited vocabulary, contain many early Indo-Aryan loanwords (Morgenstierne 1973: 332–3; Buddruss 1977: 25, 31). Older Iranian loanwords, on the other hand, have generally been reported as almost absent (Buddruss 1977: 25–7). While Iranian lexical influences are indeed less prominent compared to the high proportion of Indo-Aryan loanwords, it is not unlikely that a detailed etymological examination grounded in a better understanding of the historical phonology will reveal further examples. One family of words that may be added to the list of Iranian borrowings is the following:¹

Katë

w *lod*, NE/SE *lot* “peace, settlement [of a dispute]”

Related forms:

w *ladír*, NE/SE *ladér* “mediator(s)”

SE *alót* “unfair” (cf. *a-gúñi* “ungrateful”, *a-bídi* “unwise”, etc.)

SE *lot karōlē* “peacemaker” (*kar-ō-l-ē* do-CAUS-IPFV.PTCP-M)

SE *špā lot* “law enacted by townsmen” (*špa-* “to agree”)

SE *āmři lot* “peace settlement in adultery cases”

SE *pa-lót že-* “to sit in arbitration” (*pa-* “at/to”, *že-* “to sit”)

(SE from Strand 1999; w from Strand 2011 and Grjunberg 1980: 115; NE from Sun-Aro 2016 and own data)

Nuristani Kalasha

z *lāt* ~ *lād*, N *lād* “justice, right(s)”

Related forms:

z *a-lād* “invalid, unjust”

N *lād k-* “to settle a dispute, administer justice” (*k-* “to do”)

¹ On choice of language names, cf. Halfmann 2021.

z *lād a-šéi nišāw* “to carry out justice”
 (lit. “seat justice on the head”, Prs. *haqq ba kursī nišāndan*)
 z *lād-tā* “high court” (lit. “justice place”)
 z *lād-tāw* “court” (lit. “justice place”)
 z *tū-bā lād ō* “you are right” (lit. “it is your right”)
 N *úma lād oy, túba lād noy* “I am right, you are not right” (lit. “it is my right, it is not your right”)
 (z from Tāza 2017; N from Degener 1998)

Ashkun *lāt* “peace” (Morgenstierne 1929: 264; Strand 2008)
ladír “mediator” (Klimburg 1999: 74)

Prasun z/1 *lad* “court, settlement [of a dispute]”
 1 *lad li-* “to hold court, to settle a dispute, to bring about peace”
 (li- “to do”) (Buddruss and Degener 2015)

An etymology for these words was first proposed by Morgenstierne (1929: 264), who connected them with OIA *labdhi-* “obtaining, gain, acquisition” under the assumption of a basic meaning of “peace”, which might have arisen from OIA *labdhi-* through the identification of peace as something that is gained or obtained. Morgenstierne (1929) did not explicitly state whether he considered the Nuristani lexemes to be Indo-Aryan loanwords or rather an independent inheritance from Proto-Indo-Iranian that is merely cognate to the Old Indo-Aryan word.

From the perspective of historical phonology only the assumption of an Indo-Aryan borrowing would in fact be admissible, since the Katë forms show initial *l-* in correspondence to the *l-* of the other languages. This distinguishes the cognate set from those showing an earlier correspondence of NKal., A., Pr. *l* ~ K. *ř*:²

NKal. <i>letr</i> , (A. <i>latrā-</i> “to harvest”)	~	Kt. SE <i>řetr</i>	“harvesting time”
NKal. z <i>wacāla</i> , N <i>oçalá</i> ; A. <i>očālā</i>	~	Kt. W <i>wēcěř</i> , NE <i>učěř</i> , SE <i>vacěř</i>	“calf”
NKal. <i>palá</i> , A. <i>palá</i>	~	Kt. W/NE/SE <i>parě</i>	“apple”
NKal. <i>mül</i> , A. <i>mulí</i> , Pr. <i>mülú</i>	~	Kt. SE <i>muří</i>	“price”
NKal. <i>tul-</i> , A. <i>tol-</i> , Pr. <i>atul-</i>	~	Kt. W <i>tuřé-</i> SE <i>tuřě-</i> Kt. SE <i>tüř</i>	“to weigh sth.” “a weight measure”

This correspondence can be interpreted in various ways:

1. It could be the result of a Proto-Nuristani **l* which changed > *ř* only in Katë; Katë *l* would accordingly come from later borrowings.
2. It could reflect an earlier layer of borrowings before the sound change **l* > *ř* in Katë; Katë *l* would accordingly come from later borrowings.
3. It could reflect a substitution for *l* in earlier borrowings; Katë *l* would accordingly come from later borrowings.

No matter which explanation is preferred, Katë words with *l* can only have come into being after the time of the common ancestor of all Nuristani languages, most likely as

² Data from sources cited above.

a result of borrowing. The cognate set of Kt. *tūr* and *tuřé- ~ tuřé-* shows a typically Indo-Aryan development of PIr. **rHV-* (Av. *tarā-* vs. OIA *tulā-* “scales”). It could therefore at most be possible that the words showing the earlier correspondence were borrowings from Indo-Aryan into Proto-Nuristani. This is the only way in which one could salvage the assumption of a Proto-Nuristani phoneme **l*, which is *a priori* unlikely to have existed, in view of the common Indo-Iranian merger of PIE **l* and **r*.³ But even the assumption of a loan from OIA *labdhi-* or a descendant form runs into difficulties, since the Nuristani forms attest to an earlier long **ā*, and lengthening of a short *a* is not otherwise attested in this context for the Nuristani group.

If one additionally considers the semantics of the Nuristani words the connection with OIA *labdhi-* becomes very tenuous. First of all, it is already clear from the more recent sources that “peace” is not the basic meaning in any of the languages. The available translations as well as examples from use, in combination with such derivate meanings as “mediator” and “unfair”, make it clear that the central meaning of the terms is something similar to “law” or “justice”. All other meanings can plausibly be accounted for as secondary developments from this basic meaning: “law, justice” > “court of law” = “mediation/settlement (of a dispute)” > “peace (achieved by mediation/a settlement)”.

Among later researchers in Nuristani historical linguistics, Morgenstierne’s proposal of a connection with OIA *labdhi-* has only been questioned by Strand (1999), who instead adduced as potential cognates Russian *lad* “agreement”, *lada* “spouse”, Tocharian B *laare* “dear” and Lycian *lada* “wife”, with a reference to Winter (1965: 191). These connections would presuppose a direct inheritance of the word from Proto-Indo-European, which is already a difficult assumption to make, given the development of PIE **l* in Indo-Iranian, but it becomes yet more implausible when we consider that earlier borrowings in Katë render **l* as *ř*, as discussed above. It is further impossible to derive the Nuristani forms from Winter’s (1965: 191) reconstruction **loHd^(h)os*, since intervocalic *-d^(h)-* would have disappeared in all Nuristani languages except Prasun, where it would have turned into *l*.

A more plausible idea comes from Tāza (2017: 1589), who compares the Avestan word *dāta-* “law”, based on its semantic and phonological similarity with the Nuristani Kalasha word *lād*. Mere superficial similarity is of course not enough to establish an etymology, and for Morgenstierne this connection would have been inadmissible on the grounds of historical phonology (**d > l* is not regular in Nuristani Kalasha and intervocalic *-t* would have been lost). What Morgenstierne could not have known is that a large number of new Bactrian documents from Afghanistan, which significantly expanded the available corpus of Bactrian language materials, would be discovered a few decades after his passing in 1978. In these documents a cognate of the Avestan and common Iranian word for “law” **dāta-* appears, which is spelled in the Graeco-Bactrian script as <λαδο> and meant approximately “law, lawsuit, trial, court, judgement” (Sims-Williams 2007: 225). This word attests to the typically Bactrian sound changes **d > l* and **t > d*, and while vowel length is not consistently marked in Bactrian orthography (though it was phonemic) it is safe to assume that this word inherited a long *ā* from Proto-Iranian. Its phonological shape would therefore have been /*lādə*/ or /*lād*/, depending on the interpretation of word-final <o> (cf. Morgenstierne 1970: 126; Sims-Williams 1988). This Bactrian word

³ Quite a few open questions remain with regard to this merger, but since Old Iranian languages show exclusively *r* and the oldest Old Indo-Aryan texts show predominantly *r* for both PIE **r* and **l*, whereas younger Old Indo-Aryan has both *r* and *l*, but in a distribution that cannot be reconciled with that in other Indo-European languages, it cannot be maintained that the PIE phonemic distinction survived in Proto-Indo-Iranian. The presence of words with *l* in New Iranian languages does not contradict these points, especially since the words in question can probably all be explained as secondary. See Mayrhofer (2002) for a detailed discussion of the issue.

would seem to be the ideal candidate as the source for the Nuristani word family presented above. The vowel development $*\bar{a} > \text{Kt. } o, \text{ NKal. } \bar{a}, \text{ A. } \bar{a}$ is regular in stressed monosyllables, as is final devoicing ($-d > -t$) in the Northeastern and Southeastern dialects of Katë. The free variation of $-d$ and $-t$ in word-final position is also as expected for Nuristani Kalasha (Buddruss 1975: 46). Since all of the relevant Nuristani words, including those in Katë, reflect the Bactrian l as l , the borrowing must be post-Proto-Nuristani. How the preservation of the plosive $-d$ is to be interpreted depends on whether or not the Bactrian word had a final vowel at the time of borrowing. If a final vowel was present and also survived the process of loanword adaptation,⁴ then the borrowing would have happened after the loss of single intervocalic plosives but before the loss of final vowels. If the attested chronology of sound changes in neighbouring Middle Indo-Aryan Gāndhārī can be taken as a model for the Nuristani developments – which is not unlikely given the close contact attested by loanwords and the striking similarity of the changes – then this would place the borrowing somewhere in the first century AD (Fussman 1989: 463–4). If the final vowel was not there, then it is difficult to draw any certain conclusions, since a word-final plosive might potentially have been preserved when intervocalic plosives were dropped. However, as I will argue below, it is most probable that the final vowel was in fact still present at the time of borrowing.

Most of the Nuristani derivatives listed above can easily be explained as later formations, but a notable exception is that of Kt. *NE/SE* *ladér*, *w* *ladír*; *A.* *ladír* “mediator”. This word would most plausibly be connected to Bactrian $\lambda\alpha\delta\omicron\beta\alpha\rho$ /*lād(ə)var(ə)*/ “judge” (Sims-Williams 2007: 225),⁵ but the phonological development is not completely clear. The preservation of intervocalic $-d-$ points either to borrowing after dropping of single intervocalic plosives or to a consonant cluster ($-dv-$) in the source form. The vowel correspondence Kt. *w* $i \sim$ Kt. *NE/SE* e in the stressed second syllable of a disyllabic word, is otherwise only known from Kt. *w* *apší*, *NE/SE* *apšě* “water mill” (f).⁶ These forms can most plausibly be explained as the outcome of an earlier compound $*\bar{a}p$ -*pešáni* “water mill”, consisting of a descendant of PIIr. $*ap-$ \sim $*\bar{a}p-$ “water” (Kt. \bar{o} , NKal. $\bar{a}w$ etc. “water”) and $*pešáni$ a derivative of PIIr. $*paiš-$ “to grind”, paralleled in its formation by OIA *peṣaṇī-* “grindstone” (cf. semantically New Persian *āsyāb* “water mill” $<$ $\bar{a}s$ “grindstone” $+ \bar{a}b$ “water” (Ḥasan-Düst 2014: 66)). The root shows the inherited Nuristani development of PIIr. $*š > \bar{s}$ instead of IA \bar{s} .

If this parallel is admissible, Bact. $\lambda\alpha\delta\omicron\beta\alpha\rho$ /*lād(ə)var(ə)*/ may have been borrowed as $*\bar{l}ādivári$ or $*\bar{l}ādvári$ with substitution of i for \bar{o} ,⁷ as well as the common Nuristani penultimate stress. The preservation of a final vowel in this borrowed form would point to a rather early period in the history of the Bactrian language, which would be compatible with a borrowing somewhere around the first century AD, as suggested above. The two possible borrowed forms open two separate routes for further sound development, but both of these routes would in fact lead to the same result: first, if the originally borrowed

⁴ A central vowel might conceivably have been dropped during borrowing into Nuristani varieties since it probably did not exist in their phonemic inventories at the time.

⁵ Sims-Williams (2007: 225) also gives the cognates Parthian *d'dbr* /*dādbar*/ and Manichaean Middle Persian *d'ywr* /*dāywar*/. As an anonymous reviewer points out, Róna-Tas (2016: 70) sees a borrowing of the same Bactrian word also in the Old Turkic title $\langle \text{H} \rangle \text{elteber}$, which would then have been re-borrowed into Bactrian as $\omega\lambda\tau\omicron\beta\eta\rho$.

⁶ A similar correspondence is found in K. *NE/SE* *ačě*, *w* (Kulem, Ārangēl) *ačī*, *w* (Ktivi) *ačē* “eye(s)”, which is easiest to derive from a PNur. form $*ačāni$ “eyes” corresponding to OIA *akṣ-ān-i* “eyes (pl.)”. The divergent outcome in Ktivi – reported both by Morgenstierne (1978: 3) and Strand (2011) – may be attributable to the difference in original vowel quantity.

⁷ The Bactrian reduced vowel might have been phonetically similar to the central vowels common in the region today (phonetically [ə~i]).

form was **lādivári* (with a medial vowel *-i-*), then the word would have gone through the same umlaut process as **āp-pešáni* leading to **lādivér~*lādivír*. This form would then have been affected by the syncope of pre-stress vowels which affected all of **e*, **i* and **u* in varying contexts, producing something similar to **lādvér~*lādvír*. If the medial vowel was not present in the borrowed form, the umlaut process would lead directly to **lādvér~*lādvír*. This, together with cluster simplification *dv > d* (cf. Kt. *du* “door” < **dvāra-*) and the regular outcome of pre-stress **ā > K. w/NE a ~ SE a* (cf. Kt. *w/SE daré* “door-jams” < **dvārāka-*), would explain the attested forms.

The decision between the two options is made easier by the Ashkun form *ladír*, which can be explained from **lādvári* but not from **lādivári*, since Ashkun did not generally go through the same pre-stress syncope as Katë (compare A. *pisās* with Kt. *pšaš* “cat”). The reconstruction without a medial vowel *-i-* is therefore preferable, since it does not require the assumption of a borrowing from Western Katë to explain the Ashkun form. This fact may have some implications for the reconstruction of the Bactrian sound system and the interpretation of its orthographic representation, pointing to the possible lack of any sound value of <o> between the elements of a compound, already at a time when word-final <o> was still pronounced.

While the arguments for the identification of Kt. *ladír ~ ladér* as an outcome of the borrowed Bactrian word *λαδοβαρο* are not beyond doubt, it seems quite clear that at least the Bactrian word *λαδο* “law” was borrowed into the predecessors of the Nuristani languages, and that this would have happened after they had already become separate languages. It is not unlikely that such a word may have been borrowed when the speakers of ancestral Nuristani varieties came into contact with the judicial institutions of one of the bactrophone empires ruling on both sides of the Hindu Kush mountain range. If the dating of the borrowing to the first century AD is correct, the Kushan empire is the most likely culprit. Whether the speakers of earlier Nuristani languages lived under the dominion of the Kushans or whether they only had a passing familiarity with the judiciary of a neighbouring empire, is not retrievable from the linguistic facts, but the presence of the loanword points at least to some degree of contact.

Abbreviations

Katë = Kt.:

w = Western dialect

NE = Northeastern dialect

SE = Southeastern dialect

Nuristani Kalasha = NKal.

z = Žōnčigal dialect

n = Nišeygrām dialect

Ashkun = A.

Prasun = Pr.

z = dialect of Zumu

i = dialect of Ištewi.

OIA = Old Indo-Aryan

Av. = Avestan

PIE = Proto-Indo-European

Competing interests. None

References

- Buddruss, Georg. 1975. “Gāndhārī-Prakrit *chada* ‘Ton’”, *Studien zur Indologie und Iranistik* 1, 37–48.
- Buddruss, Georg. 1977. “Nochmals zur Stellung der Nūristān-Sprachen des afghanischen Hindukusch”, *Münchener Studien zur Sprachwissenschaft* 36, 19–38.

- Buddruss, Georg and Almuth Degener. 2015. *Materialien zur Prasun-Sprache des afghanischen Hindukusch. Teil I: Texte und Glossar.* (Harvard Oriental Series 80.) Cambridge, MA: Harvard University Press.
- Degener, Almuth. 1998. *Die Sprache von Nisheygram im afghanischen Hindukusch.* (Neuindische Studien 14.) Wiesbaden: Harrassowitz.
- Fussman, Gérard. 1989. "Gāndhārī écrite, Gāndhārī parlée", in Colette Caillat (ed.), *Dialectes dans les Littératures Indo-Aryennes*, 433–502. (Publications de l'Institut de Civilisation Indienne 55.) Paris: De Boccard.
- Grjunberg, Aleksandr L. 1980. *Jazyk kati. Teksty, grammatičeskij očerk.* (Jazyki vostočnogo Gindukuša.) Moscow: Izdatel'stvo "Nauka".
- Halfmann, Jakob. 2021. "Terminological proposals for the Nuristani languages" *Himalayan Linguistics* 20/1, 28–64. doi:10.5070/H920150079
- Ḥasan-Dūst, Moḥammad. 2014. *Farhang-e Rīs-ē-s-ēnāḥtī-ye Zabān-e Fārsī.* Vol. 1. Tehran: Farhangestān-e Zabān-e va Adab-e Fārsī.
- Klimburg, Maximilian. 1999. *The Kafirs of the Hindu Kush. Art and Society of the Waigal and Ashkun Kafirs.* Vol. 1. Stuttgart: Franz Steiner Verlag.
- Mayrhofer, Manfred. 2002. "Zur Vertretung der indogermanischen Liquiden in den indo-iranischen Sprachen", *Indologica Taurinensia, the Journal of the International Association of Sanskrit Studies* 28, 149–61.
- Morgenstierne, Georg. 1929. "The language of the Ashkun Kafirs", *Norsk Tidsskrift for Sprogvidenskap* 2, 192–289.
- Morgenstierne, Georg. 1970. "Notes on Bactrian phonology", *Bulletin of the School of Oriental and African Studies, University of London* 33/1, 125–31.
- Morgenstierne, Georg. 1973. "Die Stellung der Kafirsprachen", in Georg Morgenstierne, *Irano-Dardica*, 327–43. (Beiträge zur Iranistik 5.) Wiesbaden: Ludwig Reichert Verlag.
- Morgenstierne, Georg. 1978. *Kati Vocabulary* (Unpublished Manuscript digitized by the Norwegian National Library, Oslo – Ms.fol.4201, D:23, Kati XIII). https://www.nb.no/items/URN:NBN:no-nb_digimanus_298571.
- Róna-Tas, András. 2016. "Bayan and Asparux. Nine notes on Turks and Iranians in East Europe", In András Róna-Tas, Éva Csató, Bo Utas and Lars Johanson (eds), *Turks and Iranians. Interactions in Language and History*, 65–78. (Turcologica 105.) Wiesbaden: Harrassowitz.
- Sims-Williams, Nicholas. 1988. "Bactrian language" [updated 2011]. *Encyclopædia Iranica, online version.* <https://www.iranicaonline.org/articles/bactrian-language>.
- Sims-Williams, Nicholas. 2007. *Bactrian Documents from Northern Afghanistan, Vol. II: Letters and Buddhist Texts.* London: The Nour Foundation and Azimuth Editions.
- Strand, Richard. 1999. *Kāmv'iri Lexicon (Last edited 2019).* <http://nuristan.info/IngFrameL.html>.
- Strand, Richard. 2008. *Saṅu v'iri Lexicon (Last edited 2011).* <http://nuristan.info/IngFrameL.html>.
- Strand, Richard. 2011. *Kāt'a vari (kt'i vřā-i vari) Lexicon (Last edited 2011).* <http://nuristan.info/IngFrameL.html>.
- Sun-Aro, Sulaymān. 2016. *Farhang-i Nūristānī.* Kabul: Samar.
- Tāza, Samī'ullāh. 2017. *Farhang-i Zabān-i Nuristānī (Kalaṣā-alā): Ma'nā wa Tas'riḥ-i Luḡāt, bā Zabānhā-yi Paḫtō wa Dari.* Kabul: Summer Institute of Linguistics.
- Winter, Werner. 1965. "Tocharian evidence", in Werner Winter (ed.), *Evidence for Laryngeals*, 190–211. (Janua Linguarum, Series Maior XI.) The Hague: Mouton & Co.