himalayan linguistics

A free refereed web journal and archive devoted to the study of the languages of the Himalayas

Himalayan Linguistics

Terminological proposals for the Nuristani languages

Jakob Halfmann

University of Cologne

ABSTRACT

Recent years have seen an increase in the variety of language names used in linguistic works for the Nuristani languages (Indo-Iranian), which are spoken in Eastern Afghanistan and to some extent across the border in Northern Pakistan. This increase is driven by efforts to recognize local practices, but it has also created confusion and inconsistencies in many areas, making terminological justifications a necessary part of every publication regarding these languages. A unified terminology would not only be convenient for linguists and other researchers, in so far as scientific usage can influence colloquial practices it would also increase the visibility and recognizability of the languages of this little-known language family on the national and international level, thus also potentially benefiting the native speaking communities. The proposals given in this paper aim to create an internally consistent terminology that is scientifically precise, accurate, and recognizable for speakers, while also being congruent with the naming principles put forward for the Glottolog database.

KEYWORDS

Nuristani, Indo-Iranian, terminology, language names

This is a contribution from *Himalayan Linguistics, Vol. 20(1): 28-64.*
ISSN 1544-7502

Himalayan Linguistics, Vol. 20(1). © Himalayan Linguistics 2021
ISSN 1544-7502

*Terminological proposals for the Nuristani languages**

Jakob Halfmann
University of Cologne

1 Introduction

The names used by linguists for the Indo-Iranian languages of the Nuristani family[1] have changed frequently over the last two centuries, usually concurrently with increased availability of accurate information about the languages and their speakers. In recent decades, a greater awareness for the concerns and practices of the native speaking communities has led to several changes to the terminology, which are, however, often inconsistently applied, thus sometimes producing confusing and inaccurate results. The most recent set of changes mostly follows the example of American anthropologist Richard Strand, who spent several years living among the Kom people of southeastern Nuristan in the 1960s and thereby gained an inside perspective on the cultures and languages of Nuristan.

At the proposal of Strand (1973: 297), the name "Nuristani" for the language family as a whole replaced the earlier terms "Kafir" or "Kafiri" (Persian/Arabic for 'unbeliever, infidel'), which were originally adopted when Nuristan had not yet been converted to Islam[2] and was generally named "Kafiristan" ('land of the unbelievers') by neighboring Muslim peoples. This change is an obvious choice, considering both the clearly offensive nature of the word "Kafir" to the Muslim people of Nuristan and its inaccuracy, since the only remaining non-Muslims of the region, the Kalasha of Chitral, speak an Indo-Aryan language. Since the Nuristani languages are for the most part spoken within the borders of the Afghan province of Nuristan, the term "Nuristani" is a reasonable replacement, and it is by now fully established in the linguistic (Buddruss 1977; Mayrhofer 1984; Mayrhofer 1992; Degener 2002) and anthropological literature (Edelberg & Jones 1979; Cacopardo & Cacopardo 2001; Cacopardo 2016).[3]

---

* In this paper, unless otherwise indicated, the following non-IPA transcription conventions are employed: macron = vowel length; underdot = retroflex; y = [j]; č, ǰ = [t͡ɕ], [d͡ʑ]; ċ, ǰ =[t͡ʂ], [d͡ʐ]; ř, ň = [ɻ], [ɳ]; ü = [y]; ö = [ø]; ë = [ɘ]. Otherwise, transcriptions follow IPA usage.

[1] Glottocode: *nuri1243*. The most reliable published survey on the languages of Nuristan is still Strand (1973), which should be consulted in conjunction with the corrections added in Strand (2001). For various views of the position of Nuristani between the Iranian and Indo-Aryan branches, see Degener (2002), Morgenstierne (1973), Buddruss (1977), Mayrhofer (1984), and Lipp (2009).

[2] I.e., before the military conquest of Amir Abdurrahman Khan, then ruler of Afghanistan, in 1896.

[3] A rare recent exception is Irgens-Møller (2009). Grjunberg (1971b, 1980) uses both terms side by side, differentiating by historical era, while noting that it may generally be "more appropriate" to use "Nuristanis" (*Nuristancy*) instead of "Kafirs" (1971b: 266) when talking about the people, rather than their languages.

The attempt to establish scientific naming conventions for the individual languages has proved far more problematic. Lentz's (1937) description of the situation as it was encountered by the German Hindu Kush Expedition of 1935 gives an impression of the difficulties involved:

> The number of language names and especially of tribal names, which are used in Nuristan, is much larger than that of the language or even dialect areas. [...] a large part of the designations even for wide areas consists of village names. When we asked an informant for the name of tribe and language, we usually received the statement of his home village.
> In this way the remarkably high number and the significant divergences in the reports about Kafir tribes among scientific authors can be explained. And a certain vacillation in this point in the statements even of reliable informants, which at times could not even be resolved by the most detailed and exact questioning, is then not surprising any more but almost natural. (Lentz 1937: 266–267; own translation)

In order to make the reasoning behind my proposals in this area more understandable, I will try to establish a groundwork of useful criteria for scientific language naming in section 2 with reference to the standardizing principles set down by Haspelmath (2017) for the Glottolog database (Hammarström et al. 2020). After an overview of the Nuristani family in section 3, I will discuss naming issues separately for each of the four best-known languages in the fourth section, finally adding some preliminary remarks on the underdocumented varieties of "Tregami" and "Zemiaki".

It should be noted that my proposals are directed at a unification of scientific naming conventions within linguistics. It is of course not my intention to prescribe particular terms for everyday usage in communities of native speakers. The distinction between language names in the scientific domain and in colloquial use is usually not very large and may even be unnoticeable, but it is an important one to keep in mind, when faced with situations such as the ones discussed in this article.

2 Principles for Language Naming

Talking about a variety of different languages in a scientific way is the central task of linguists. For this purpose, naming individual languages and differentiating them from each other terminologically in a way that reflects the purely linguistic layer of variation is an absolute necessity. As Haspelmath (2017: 91) notes, determining the most appropriate name for a language is "not primarily a research question" in itself. Nevertheless, it is a question that must be solved in order for the results of research to be communicated in an effective and understandable way. We could content ourselves with exclusively using unique identifiers (like ISO 639-3 codes or Glottolog's "glottocodes") to refer to specific languages, but this practice is neither intuitive nor practical for human beings, who usually have a much easier time relating words and names with meaning than codes and are often already used to giving languages names. This has the inconvenient result that we must compromise between a purely scientific classification of phenomena, any number of names that may already be in colloquial use and a tradition of research which may have established certain terms in the existing literature. Additionally, we would like to be respectful of the wishes

of speaker communities (in so far as these can be clearly identified), especially when it comes to avoiding obviously offensive names.

As becomes clear from the examples cited by Haspelmath (2017), individual linguists have rather different ideas about what constitutes a good compromise for naming a language. Some try to stay as close as possible to the self-designation (or "autoglottonym"), while others (like Haspelmath) prioritize the possibility of integration into the metalanguage. It is therefore likely that not every linguist will agree with the 11 principles adopted for use in the Glottolog database, as is also evidenced by the Dryer's (2019) response, but they constitute a principled catalog of objective criteria that can serve as a point of orientation in the complicated terminological issues surrounding Nuristani languages.[4] Some of these principles are merely explicit statements of preferences that have long been commonly held in linguistics, others may be more controversial:

1) Language names (like city names) are loanwords, not code-switches.
2) Names of non-major languages are not treated differently from names of major languages.
3) Each language has a unique name.
4) New language names are not introduced unless none of the existing names is acceptable for some reason.
5) Language names that many speakers object to should not be used.
6) Language names in English are written with ordinary English letters plus some other well-known letters.
7) Highly unusual pronunciation values of English letters are not acceptable.
8) Language names must be pronounceable for English speakers.
9) Language names begin with a capital letter.
10) Language names may have a modifier-head structure.
11) The usage of prominent authors is given substantial weight.

In the following section I will examine to what degree the various terms used in the literature on Nuristani languages conform to these principles, and how they can serve as a guide for reaching a terminological compromise that all or most authors, as well as native speakers, who may or may not be involved in the scientific discourse, will be able to accept. Such a solution, beside improving conditions for linguists, could also have some positive real-world consequences by making it possible where necessary to refer to Nuristani varieties in unambiguous and clearly recognizable ways, thus improving the wider visibility of these minority languages (cf. Haspelmath 2017: 85).

[4] An anonymous reviewer points out that it may be problematic to rely exclusively on a set of principles proposed by a single author. It is true that these principles do not represent all possible views of language naming within linguistics, nor is it, to my knowledge, an immediate priority of the editors of Glottolog to adapt all language names in the database to conform to these principles. Nevertheless, as an explicit statement of preference by a standardizing institution within the field, they provide a helpful frame of reference. For the proposals given in this paper I do not treat them as inviolable rules but rather as a set of priorities to be conformed to or diverged from based on an evaluation of competing points of view.

3 Overview of the Nuristani Family

A map of the geographical distribution of the Nuristani languages in and around the Afghan province of Nuristan is given in figure 1. The language names used here follow from the discussion in section 4 below. ISO 639-3 codes are added in square brackets for clarity. Many of the high mountain areas on the map are sparsely populated, if at all. The assignment of such areas to a particular language is based on settlements at lower elevations in the same valley. Settlement dots give an impression of population density. Second language speakers and minority populations are not represented. Also not represented are nomadic or recently settled Gujar populations speaking the Central Indo-Aryan Gujari language. Some of the smaller languages may be overrepresented on the map in cases where only a part of the given village's population still speaks the ancestral language. Locations and boundaries may not be exact or up-to-date in all areas.

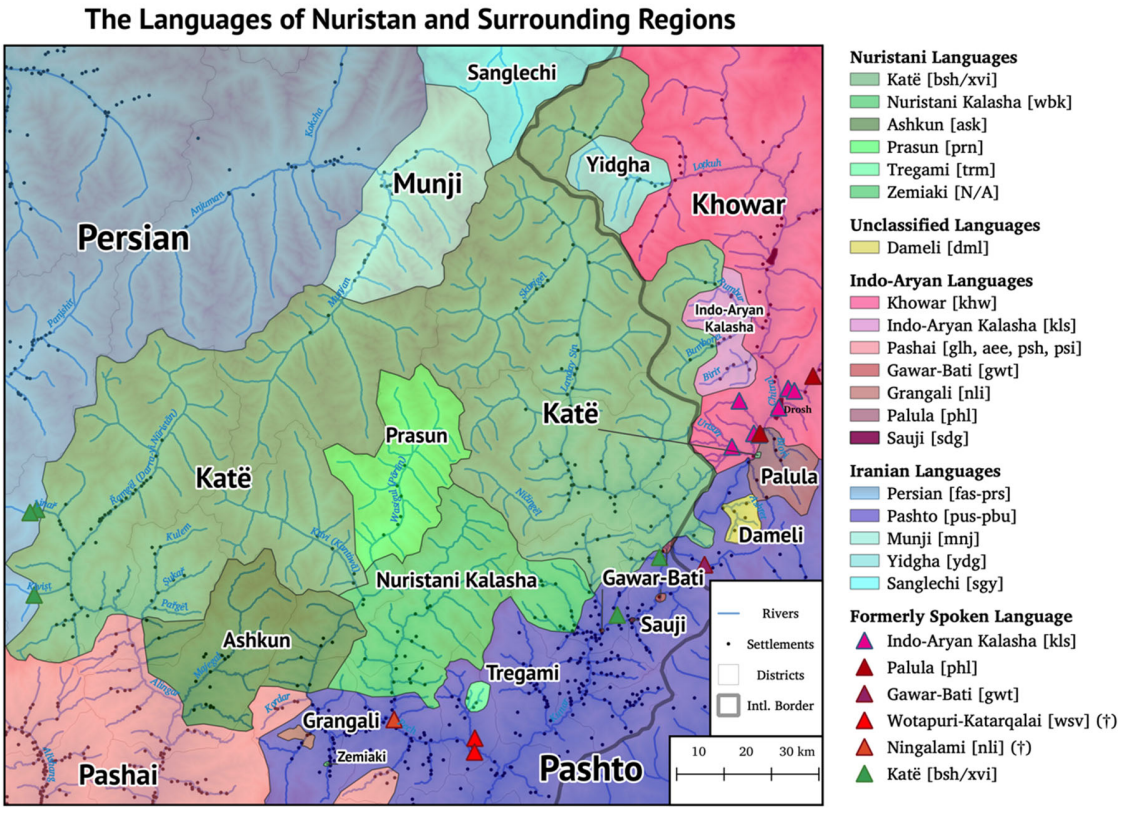

Figure 1. Map of the Nuristani languages[5]

[5] Made in QGIS 3.16 based on U.S. National Imagery and Mapping Agency (~1979) – Joint Operations Graphic 1:250.000 maps (NI42-3, NI42-4, NI42-8), supported by OpenStreetMap data and Google satellite images. Elevation data from Jarvis et al. (2008), settlements from AIMS. River names and language boundaries are based on Lentz (1937), Herrlich et al. (1937), Morgenstierne (1950, 1954), Buddruss (1960b), Edelberg & Morgenstierne (1974), Christensen (1980), Strand (1973, 1997a, 1998a, 1998b, 1998c), Grjunberg (1971a, 1980), Decker (1992), Cacopardo & Cacopardo (2001), Cacopardo (2007: 256), Beyer & Beck (2011), Beck (2012), Heegård (2015), Liljegren (2016), as well as information provided by Qiamuddin Khadim and Sviatoslav Kaverin.

Figure 2 contains a likely genealogical tree classification of the Nuristani languages. The general outline of the tree follows Strand (1973: 302). The later proposal of a Northern and Southern subgroup advocated by Strand (2001: 258) is not adopted here, as there are no compelling historical-comparative arguments for such a separation (cf. e.g., Nelson 1986: 64–66). Strand's subgroupings are based on an unclear concept of "speech postures" such as "prognathizing" (cf. Strand 2016: 72) as opposed to more reliable criteria like shared phonological and morphological innovations.[6] Justifiable subgroupings may yet emerge from a detailed examination of such criteria. The genealogical position of Dameli is still unclear, it may belong to either the Nuristani family or the Indo-Aryan family (cf. Morgenstierne 1942: 144–148; Halfmann forthc.). The status of Tregami as a separate languages or as a subvariety of Nuristani Kalasha as well as the status of Zemiaki as a separate subvariety of Nuristani Kalasha is also still an open question (cf. section 4.5 below).

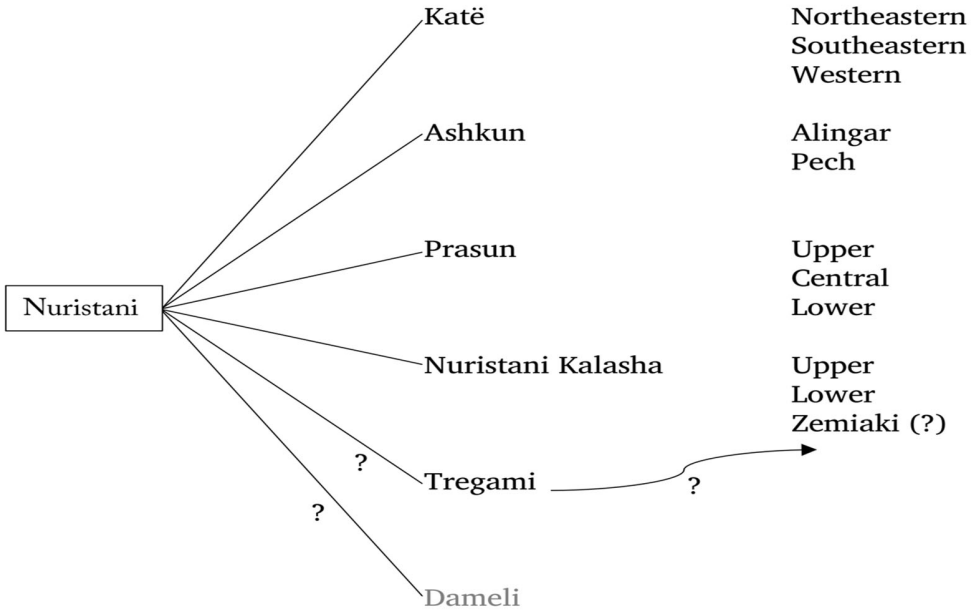

Figure 2. Genealogical classification of the Nuristani languages

[6] Cf. the explanation of "prognathizing" in Strand (2013): "It is widely recognized that the opposing modes of 'sweetness' versus 'harshness' are expressive of conciliation and belligerence, respectively. At the root of these contrasts are the extremely ancient metaphors of feminine, associated with high pitch, versus masculine, associated with lower pitch. Prognathizing, with its jutted-out jaw, is probably at its base also a signal of belligerence, employed by otherwise 'sweet'-talking Irânians and Nûristânîs to express insolence. This is not to suggest that today's speakers are necessarily belligerent or nice, only that such attitudes prevailed at the time when such postures were adopted, and the postures have persisted. Perhaps encoded in the ancient frontings and backings that define the linguistic branches of Indo-Irânian are waves of confrontation and conciliation that echo ancient ethnic encounters."

4 Proposals for the Individual Languages

4.1 Katë = "Kati" or "kâmk'ata-vari"

The largest language of Nuristan (Glottocode: *kati1270*) currently has two separate ISO 639-3 codes, *bsh* and *xvi*, one being identified with the name "Kati" (*bsh*) and the other with the name "Kamviri" (*xvi*). As we will see, these two terms are taken from two different terminological complexes, whereas the code *bsh* itself results from yet another term. This already gives a first impression of the complications that need to be solved.

The ethnic groups speaking the largest language of Nuristan call themselves (ordered approximately by numbers): Katë [kaˈt̪ə], Kom [kɔm], Mumo [muˈmɔ], Kṣto [kʂt̪ɔ], J̌amčo [d͡ʑamˈt͡ɕɔ], Binyo [binˈjɔ] and J̌aži (J̌aši) [d͡ʑaˈʑi] (cf. Strand 1997a). The separation into two languages by ISO 639-3 runs contrary to the widespread perception as a single mutually intelligible language among speakers. The coherence of the language and mutual intelligibility of its dialects is confirmed by Strand (2001: 253) and is also asserted by my native-speaking consultants.[7] The ISO-classification is probably based on a number of sentences by Grjunberg (1980: 16–17), who was most familiar with the western dialect of Kulem and tentatively separated "Kamviri" (a term he adopted from Strand's (1973: 298) list of self-designations) from what he calls "Kati proper", i.e. the dialects of the Katë[8] people. However, contrary to Grjunberg's (1980: 16) statement, the dialects of the ethnic Katë are linguistically not necessarily closer to each other than they are to the dialect of the other ethnic groups.[9] It would make more sense to separate the dialectological classification from ethnic terms and to distinguish instead a Western dialect (with common features but some internal diversity), a Northeastern dialect and a Southeastern dialect as is done in Figure 2 above.

A specific self-designation for the language as a whole, independent of ethnic boundaries, does not exist among native speakers (cf. Grjunberg 1980: 22–23). Nevertheless, as has been pointed out, they understand very well that their varieties are related to each other at the level of dialects, whereas the other Nuristani languages are understood to be qualitatively different.

[7] Grjunberg (1980: 24, 93) was successful in having his consultant Nuri Said, a speaker of the Western dialect of Kulem, interpret a recording of the Southeastern dialect. Parsons (1995: 7) writes that speakers from Řamgël (western dialect) claimed to understand recordings of the Northeastern dialect very well, but reported difficulties in understanding the Southeastern dialect. He did not test the actual level of comprehension, but relied on speakers' statements, the interference of ethnic or political factors can therefore not be excluded.

[8] <ë> here stands for a close-mid central vowel [ə]. Earlier sources transcribe this sound as [ɘ]. Before adopting his current transcription system (see fn. 16), Richard Strand occasionally transcribed it as [ɨ] (e.g. in Strand 1985) to emphasize that it is a more close vowel than a mid-central [ə].

[9] They are separated by a significant morphological isogloss: The Northeastern dialect, spoken by ethnic Katë, forms a progressive stem in *-t-*, whereas the Western dialect, also spoken by Katë people, uses *-n-*, just like the Southeastern dialect, which is spoken by other ethnic groups (cf. Strand 1973: 299). The loss of nasalization in Western dialects (cf. Grjunberg 1980: 167; cf. also Strand 1999a, 1999b) additionally sets these apart from the Northeastern and Southeastern group. The most notable feature of the Southeastern dialect is a number of sound changes resulting in the intervocalic lenition of consonants (cf. Strand 1973: 299). The dialect of the Mumo people, tentatively classed as a transitional dialect by Strand (1973: 299), judging from my limited field data linguistically belongs to the Southeastern group, as it uses the *-n-* progressive, has undergone characteristic southeastern sonorizations (e.g. 3rd person copula *azë* vs. *asë* in other dialects) and shares vowel developments in unstressed syllables with the Southeastern group.

Linguists therefore require an all-encompassing term for the language as a whole, but there is no immediately obvious pre-existing name for this purpose.

Historically, besides "Kafir", the first term used by foreign researchers to refer to the largest language of Nuristan was "Siah-Posh" (e.g., Raverty 1864; Capus 1889; Andreev[10]). In this era, so little was known in the outside world about the languages of the region, that this term was used very inconsistently and the articles claiming to be about "the Siah-Posh language" describe rather distinct languages.[11] The word itself originates from the Persian phrase *siyāh-pōš* 'black-clad', referring to the color of garments then widespread in certain areas of Kafiristan, the coordinate term being *safēd-pōš* 'white-clad'. As Cacopardo (2016: 74) notes, "Siah-Posh" and "Safed-Posh" are outsiders' terms which do not correspond to a local reality. They have at most been locally adopted to explain the real differences between ethnic and linguistic groups to outsiders in terms they were familiar with.

Following Robertson's (1896: 74) rejection of the division into "Siah-Posh" and "Safed-Posh" and especially of the latter term as "more convenient than scientifically correct", this naming convention was slowly abandoned in favor of the Khowar word *Bašgali*, spelled in English as "Bashgali", "Bashguli" or "Bashgeli" (used by Davidson 1902; Konow 1911; Grierson 1911; Konow 1913; Grierson 1919).[12] This term is derived from the Khowar name *Bašgal* for the valley where the eastern portion of the language's speakers lives. Like the language, this valley and its river were not historically referred to with a single name by the local residents. Instead, separate ethnically-based names designated segments of the valley, e.g. *Katḗgël* for the northern half inhabited by Katë people, *Mumgúl* for the area inhabited by Mumo people or *Kamḗston* for the area inhabited by Kom people. Today it is locally known by its Pashto name *Laṇḍáy Sin* ('short river') and indicated as such on the map in Figure 1.

The term "Bashgali" is the origin of the first of the two ISO codes: *bsh*. During the time of its use, outside authors, who were usually British colonial officers, still often failed to correctly distinguish the various languages of the region and confusion about the true depth of linguistic diversity in the region still prevailed. An earlier example of this is G. W. Leitner's *The Bashgeli Kafirs and their language* (1880), which actually describes the Indo-Aryan Kalasha language of Chitral, which is not spoken in the "*Bašgal* valley" at all.

The first professional linguist to work on the languages of Nuristan was the Norwegian Georg Morgenstierne. In his *Report on a linguistic mission to Afghanistan* (1926) he introduced for the first time the name "Kati" which would remain in use up to the present day. He gives the following justification for the name change:

[10] Andreev's materials were collected in Tashkent in 1902, but only published in the 1990s by A.L. Grjunberg (Grjunberg 1994, 1995a, 1995b).

[11] Capus (1889: 213) confidently proclaims about one of his informants: "Cet homme est bien Siahpouche." ('This man is surely Siah-Posh'), not noticing that this one spoke a different language than the other man he titled "Siahpouch(e)".

[12] Grierson (1919: 29), following Robertson's (1896: 74) observations, explains his view of the relationship between the two terms as follows: "It appears that the Siāh-pōsh Kāfirs, who, roughly speaking, people the northern half and the East of Kāfiristān, all speak various dialects of one language, of which Bashgalī, the speech of the people inhabiting the valley of the Bashgal River, may be taken as the type. All the tribes who wear the dark-coloured raiment seem at once to understand each other, and to be able to converse fluently and without hesitation".

> [...] the tribe [...] [is called] *Katī* and their language *Katī-verī* or *Katī*. As Bashgal only includes a small part of their country (it was even denied that the Bashgalis were real Katis), I prefer to call this most important language of Nuristan *Katī*. (Morgenstierne 1926: 40)

At this early stage, Morgenstierne had not yet fully understood the phonology of the language, and mixed up some of the vowels in the transcription: The name of the Katë tribe in IPA transcription is [kaˈʈə], the name of the language [kaˈʈəβə̣ɾi]. In addition to these two native terms, Morgenstierne was probably given a Persian adjective in *-ī* derived from the name Katë, resulting in *Katë-ī* [kaʈəˈi], which could be approximately rendered into English as "Katë-ese"[13]. Having heard the latter word as *Katī*, he apparently generalized this form to the other two terms.

The denial that "Bashgalis" were "real Katis" was probably actually a statement only about the people of lower *Bašgal/Laṇḍay Sin*, who ethnically do not belong to the Katë group. Morgenstierne had at this point not been to Nuristan in person and mainly gathered his information from the community of Katë people living in Kabul. He notes that "most of the Kafirs I met came from the western valleys of Ramgel, Kulum and Ktivi. I did not get hold of any individual from Kamdesh and the lower Bashgal [...] valley" (Morgenstierne 1926: 40). It is therefore understandable that he did not come into contact with the names of other ethnic groups speaking the same language, who live mostly in the southeastern part of Nuristan. The name "Kati" was not immediately accepted: Ivanow (1931: 154) returned to the Khowar-derived term "Bashgali", since his consultant, an ethnic Kom who had lived for some time among the Katë,[14] knew only this word and did not recognize Morgenstierne's term, at least not in the form in which Ivanow pronounced it before him. In 1967, after spending quite a bit more time studying the language and interacting with its speakers, Morgenstierne acknowledges in a footnote: "Kate'ī with Persian *yā-yi nisbatī* from the tribal name Kate, may be the more correct [form]. The native name is *Kate-weri* 'Kate-speech'. The dialect of the *Kām* tribe in the lower Bashgal valley is in reality a branch of the Kati language." (Morgenstierne 1967: 1378). He still does not clearly identify the central vowel in his spelling "Kate-weri", but the vowel contrasts inside the word are now correctly shown.

Shortly after this, the ethnic composition of Nuristan and the self-designations used by the speakers of Nuristani languages became better known outside of Nuristan through the work of Richard Strand. It became clear to outside linguists then, that in fact each of the relevant ethnic groups (Katë, Kom, Mumo, Kṣto, Jamčo, Binyo and Jaži/Jaši) uses their own ethnic name plus the word *vërí* [βə̣ˈɾi] (Western and Northeastern) / *virí* [viˈɾi] (Southeastern) meaning 'language, speech, word, issue' when they refer to their language. In the Southeastern forms regular vowel changes and stress shifts appear in the resulting compound: *Mumvíri*, *Kṣtéviri*, *Kamvíri*, *J̌ašvíri.* Additionally, there are a few Pashto terms of somewhat unclear origin (cf. Grjunberg 1980: 17), which may be used with an adjective ending *-ī* to refer to both the tribal groups and their languages:

[13] Strand (2016: 70) considers this a "nonce-word", but Parsons (1995: 7) reports the same term from Ř amgël, apparently independently. It is at least possible that this term is or was regularly used in the Persian-influenced north-western parts of Nuristan.

[14] The text recorded by Ivanow is, somewhat surprisingly, in the northeastern dialect (Grjunberg 1980: 22). It seems that the consultant had shifted from his native dialect to the dialect of *Břëgomaṭol* (prs. *Barg-i Matāl*), where he had lived for some time, or at least that he preferred to present this dialect to an outsider.

Kantoz(ī) = Katë, *Kamoz(ī)* = Kom and *Kushtoz(ī)* = Kṣto. These terms probably contain preserved earlier forms of the self-designations.

Despite the common use of ethnically-based terms to designate the language, the existing ethnic boundaries do not coincide with linguistic ones: The Katë people speak the Western and Northeastern dialect, all remaining groups together speak the Southeastern dialect. Additionally, Strand (1973: 299) notes that the section of the Kṣto people living in *Dũŋul* speaks the language of Kalashüm (see section 4.3)[15] whereas some ethnic Kom have completely shifted to Pashto and some Katë speak only Persian. Nevertheless, and in spite of the absence of a common name for the language, Strand, especially in more recent publications (e.g. Strand 2016: 70), advocates that only the ethnically-based endonymic terms or "autoglottonyms" should be used by scientific authors. He justifies this as follows (e-mail to participants of the Cologne Nuristani Workshop, 13th February 2020; quoted with permission):

> Ever since I began living in Nûristân in 1967, I have advocated for using native rather than foreign names for the languages and peoples of that region. My motivation for using native terms is based on the sensibilities of the native speakers, who generally disdain the terms used for them by outsiders. In my publications I have always striven to use native names, all of which have appeared on my website since 1999. By now you all should be familiar with these names.

Strand usually presents the relevant terms in his own personal transcription system, which is not very intuitive for readers without previous knowledge.[16] Thus we find "*kât'a-vari*" (= [kaˈʈɘβᵱɘɾi]) and "*kâmv'iri*" (= [kamˈviɾi]). Strand has coined an inclusive endonymic term himself by combining the names of the two largest groups Katë and Kom with the word *vëŕi* resulting in "*kâmk'ata-vari*" (= [kamˈkɘʈɘβᵱɘɾi]).[17] He notes that this term was coined "in the Kâmviri dialect, but with the Kâta-vari form *var'i* 'language'" (Strand 1997a).[18] On his website he also offers the more inclusive *kâmk'ata-mumkṣt'a-vari* [kamˌkɘʈɘmumˈkʂʈɘβᵱɘɾi], which, however, still does not include the names of the smaller groups Jamčo, Binyo and Jaži/Jaši.

Strand's new conventions were adopted by many recent authors, but his transcription system is usually not understood very well, resulting in a number of confusing and contradictory forms. Some, like Bashir (2006, 2010: 6), fully adopted his spellings with only removal of accent marks and added capitalization. Others, more problematically, removed the distinguishing diacritic from *â*, creating "Kam-Kataviri" (Di Carlo 2016: 107), "Kamkata-viri (including Kati)" (Heegård 2015: 21) or "Kataviri" (Perder 2013: 8; Beyer & Beck 2011: 42; Liljegren & Haider

[15] Their Kṣto identity is not confirmed by Tāza (2017: 1059), who states that the inhabitants of *Dūrĕgal* "are *Kalaṣa* and speak the *Kalaṣa* language" (own translation).

[16] Strand has given up his earlier transcription system found in Strand (1973) and used for citing his data, e.g. in Degener (1998a). In the current system <a> is used for [ɘ] and <â> for [a], which expresses the conviction that [ɘ] as the "basic vowel" should receive the more basic symbol. The grapheme <a> is also used as a marker for vowel length (e.g., <âa> = [aː]; <ua> = [uː]), as well as for [i] after palatal consonants, both of which are analysed as allophones of the same phoneme /a/. Additionally, voiced intervocalic sounds are represented as voiceless in the transcription with the argument that they are the result of predictable sonorization processes, though the processes are not in fact fully predictable (Strand 1997b, 1999b, 2007).

[17] To my knowledge, this name is not in use among any native speakers. I only discussed the term with one native speaker, who happened to reject it.

[18] He diverges from this convention with the spelling "Kâmkata-viri" in Strand (2001: 253).

2011: xvi), all in fact with the Southeastern ("Kamviri") form *virí*, which is not actually used by the Katë people.

That this spelling is usually interpreted as [kataviri] by unfamiliar readers, is shown by the Urdu transliteration (کاتاویری *kātāvīrī*) which the Pakistani NGO "Forum for Language Initiatives" uses in a published video.[19] The same organization, which advocates for better visibility and language rights for the minority languages of Pakistan, has put out the following statement (bolding added):

> **Shekhani** language; the 20th language FLI takes up
> FLI reaches another milestone by taking up the **Kataviri** language, the 20th language the organization has targeted to preserve and promote. [...] The **Kataviri** language was found vigorous in orality with strong position on identity scale but lagging behind in literacy. [...] The **Kataviri** language is spoken in hilly peripheries of Chitral valley, sharing the borders with Afghanistan from where the speakers immigrated to the area some 130 years back. [...] The language is called **Shekhani** by locals which is a group name of Nooristani languages spoken in Chitral: **Kataviri** and **Kamviri**, however the native speakers like to be identified as **Kati** speakers. [...] **Shekhani** is the 20th language FLI has started to develop by enabling its speakers in language documentation[20]

The negative real-world consequences of the terminological confusion are undeniable, when even an organization dedicated to promoting minority languages, which turns to scientific works for clarity, must mention three different terms to refer to the language, and finally resorts to the potentially offensive[21] colloquial term "Shekhani" used by the Chitrali majority.

Confusion with the older term "Kati" has created hybrid forms like "Kativiri" (cf. Glottolog; Bashir 1988: 34; Decker 1992: 129; Perder 2013: 62). Parsons (1995: 6–7), apparently integrating Robertson's (1896) colonial spelling "Katir", writes "Katerviri". The confusion of terminological complexes is also apparent in the combination "Kati-Kamviri", or "Kati and Kamviri" (cf. ISO 639-3; Di Carlo 2016: 115; Grjunberg 1980: 16–17; Fussman 1983), an inconsistent mix of Morgenstierne's and Strand's conventions, which disregards Morgenstierne's inclusion of the southeastern dialect ("Kamviri") within the meaning of the larger term "Kati" and Strand's rejection of "Kati". Use of "Kati" is continued in more recent times by Degener (2002: 104), Hegedűs (2002: 189, 2012: 147), Blažek & Hegedűs (2012: 42), Tikkanen (2008: 261), Koptjevskaja-Tamm & Liljegren (2017: 217–218) and Liljegren (2016: 24), who uses "Shekhani" as the main term and also mentions "Kamviri" and "Kamkata-viri", the latter again diverging from Strand's original "*kâmk'ata-vari*".

---

[19] Forum for Language Initiatives (FLI) - "Covid-19 explained in the Shiekhani language" [sic], available at https://www.youtube.com/watch?v=DBgMq9AgVEY.

[20] Forum for Language Initiatives (FLI) - "Shekhani language; the 20th language FLI takes up", available at https://fli-online.org/site/shekhani-language-the-20th-language-fli-takes-up/.

[21] The term Shekhani (*Šexānī*), which contains the Arabic title *Šayḫ* as a (derisive) epithet for relatively recent converts to Islam, is also sometimes used as a self-designation by speakers living in Chitral (see e.g. Decker 1992: 131), but it is strongly rejected by Strand (e-mail, February 13th 2020; quoted with permission): "many of you still persist in using incorrect and out-of-date names, some of which are highly disparaging to *Nûristânîs* (names like *Šexânî* for any Nûristânî language...)".

It is therefore no overstatement to say that a clarification of the naming issue is necessary. The most obvious solutions would be to either return to the Morgenstierne convention "Kati", or to consistently adopt Strand's proposals, perhaps with a more accessible spelling. When judged according to the standards of Haspelmath (2017, cf. section 2), we find that Morgenstierne's introduction of the term "Kati" only (arguably) violated one principle:

> 4) New language names are not introduced unless none of the existing names is acceptable for some reason.

Acceptability is of course a matter of taste and degree. To Morgenstierne the name "Bashgali" was not acceptable since the speakers he first interacted with were not from "Bashgal" and called themselves by a different name. But this would not be a valid argument in the view of Haspelmath (2017: 90), who explicitly disqualifies closeness to the self-designation as a criterion for language names, arguing that

> many major languages have c onventional English names that are quite remote from the autoglottonym (*German, Greek, Finnish, Chinese*, etc.). By the principle of 2), names of non-major languages should not be treated differently from names of major languages, so an English name that is different from the autoglottonym is not per se problematic. (fn.: On the contrary, one might even say that a language name that sounds more like an older English word is preferable because it makes the language sound more like bigger, more prestigious languages which generally have more English-sounding names, e.g. names including an English suffix.)

In theory linguists should thus have held onto "Bashgali", but as the name change has already happened, there is another principle that can be invoked:

> 11) The usage of prominent authors is given substantial weight.

Georg Morgenstierne is without a doubt the single most influential figure in the history of linguistic research in the entire region. Additionally, his usage of "Kati" was adopted by most subsequent authors including A.L. Grjunberg, who wrote the first reliable grammatical description of a variety of this language (Grjunberg 1980). Having accepted Morgenstierne's changes we should therefore have held onto the term "Kati". But we must also accept the reality that this is not what happened. When judging Strand's conventions for the entire language ("*kâmk'ata-vari*") as well as for its subvarieties ("western *kât'a-vari*", "eastern *kât'a-vari*" and "*kâmv'iri*"), we find that 7, possibly 8 out of 11 principles are violated. I will address each of these principles individually:

> 1) Language names (like city names) are loanwords, not code-switches.

Strand's terms are clearly intended as code-switches - the native pronunciation is encoded in exact detail, the native word for 'language' (in its corresponding dialectal form) is appended at the end, when the word for 'language' in the metalanguage could do just as well and it is intended that the words be produced as closely as possible to the native pronunciation. Often they are even printed in lower case and italics, giving them a distinctly unintegrated appearance.

2) Names of non-major languages are not treated differently from names of major languages.

As mentioned above, the names of major languages in English (and other languages) were not chosen according to their closeness to the self-designation but became established through historical contacts, with some diverging rather significantly from the self-designation. As this is accepted and unproblematic for major languages, Strand's introduction of new names to avoid it with minor languages violates this principle.

4) New language names are not introduced unless none of the existing names is acceptable for some reason.

The previously used term "Kati" was not acceptable to Strand as it is "not a recognizable name in any Nûristânî language" (Strand 2016: 70). One native speaker with whom I discussed the issue did find the word "Kati" recognizable in the sense that he recognized it as sounding similar to the ethnic name Katë, but he did indeed not use it himself. It is by all appearances not an offensive name but merely a natively non-existent one. One could even argue that this characteristic might have especially qualified it to serve in the function of a comprehensive language name for scientific works, as such a term does not (yet) exist among native speakers.

But we can also again appeal to the principle that conventions of major scholars should be followed, with the justification that Richard Strand has also been a prominent figure in Nuristani language studies. In recent decades he has become the single most prominent authority who is consulted by other writers whenever expertise in this area is required. Moreover, Strand himself explicitly argues against the principle of retaining established terms (e-mail, February 13th 2020; quoted with permission):

> Some have argued that the names Kati, Waigali, Ashkun, and Prasun are by now established names that we linguists must retain. To this I say "nonsense"; these names are not so venerable. They were essentially made up by our esteemed mentor Georg Morgenstierne back in the 1920s, because he was unsure of the proper native names. [...] Especially as linguists we should strive to inform our audiences of the proper native names, just as we would any other native lexical item. What we should not do is sit in our armchairs in some distant land pontificating on what to call other peoples' languages, rather than just calling them what the native-speakers want us to.

It is certainly correct that Morgenstierne's "Kati" actually does not have a particularly long tradition. Before 1926 "Bashgali" was used and it was only a few decades later in the 1970s that it began to be contested again by Strand. The principle of least possible name changes has thus already been violated so frequently for this particular language, that it has essentially lost its usefulness. Returning to "Kati" or "Bashgali" may not significantly reduce the confusion, also considering the fact that the language has never really become widely known under any particular name. On the other hand, it is of course not impossible to use a conventional name while also informing audiences about the practices of native speakers, just as a book titled *Grammar of German* may also contain the information that the language is called *Deutsch* by its speakers.

As to the wishes of native speakers, so far no complaints or demands regarding the naming of the largest language of Nuristan have been publically directed to linguists. When examining the practices of native speakers who have written about their language to audiences outside the region (e.g., Ġulāmullāh 1966; Jano 1991; Sun-Aro 2016), we find that almost none of them follow the usages of linguists and most simply use the name "Nuristani" (Persian *Nūristānī*), which in official Afghan government use effectively also often refers to the largest of the Nuristani languages. A native speaker with whom I discussed the issue also argued for "Nuristani" as the most appropriate designation, since in his opinion the smaller languages of Nuristan were not important enough to receive this label. It seems that, when a meaning inclusive of all dialects is required, "Nuristani" (also in the form *Nuristoní-vëri/viri*) is the word of choice among native speakers.

It is, however, with good reason that this name was adopted - in scientific use - for the language family as a whole and not for one of the individual languages. As Cacopardo (2016: 72) explains, the label "Nuristani" "has been vested with high status in Afghanistan, it has been to a certain extent internalized and has become [...] a 'performative' identity that has a role on the socio-political arena of present-day Afghanistan, while the old patterns of identity [...] still play their role at the local level behind the façade of this recent construction." It is therefore not surprising that speakers use this imprecise label to refer to their languages when explaining themselves to wider Afghan society or even to international audiences: It is a recognizable term that affords a certain status and a basis for advocacy while obscuring older internal differences. But as linguistic researchers who are necessarily concerned with even those details of linguistic variation that are of little concern to most "outsiders", we cannot use this label as a name for any of the individual languages because this would come at the cost of glossing over the existing local diversity.[22] Nor could a linguist adopt the point of view that the smaller languages of Nuristan are too unimportant to deserve this name. Instead it serves its best purpose as a label for the language family, a function in which it is securely established in linguistic terminology. A variation with a modifier such as "Northern Nuristani" could be considered, but this would lead to further problems, since this term is already in linguistic use for a putative subgroup of the language family (cf. section 3 above).

Among other terms, the name "Kati" was accepted by Jan Mohammad, a native speaking linguist who wrote a master's thesis about his own language (Mohammad 1991). Ġulāmullāh gave his extensive hand-written Persian-language grammar of the Southeastern dialect (Ġulāmullāh 1966) the title *Girāmir-i zabān-i* <*ktyḥ wyry*> (to be read *katë-viri*?). This form is unusual both in its spelling and in its application to the southeastern dialect of Kamdesh (*Kombëgřom/Kombřom*). It may have arisen from his personal contact with Morgenstierne who was using the word "Kati". In the text of the grammar Ġulāmullāh only refers to the language as *lisān-i nūristānī* 'the Nuristani tongue'.

Returning to the Glottolog principles, the following points can also be mentioned:

6) Language names in English are written with ordinary English letters plus some other well-known letters.

[22] An anonymous reviewer argues that there would be no issue with using the term *Nuristani* as "the standard term for a 'single underlying language'", elaborating that "English has such broadly/widely distinct varieties, and yet one primary name". The meaning of this comment is not quite clear to me, as this would essentially amount to using the term for the language family as a whole, which is its established usage besides being exactly what I am suggesting.

The circumflexed *â* is accepted under this principle as one of the "other well-known letters" (Haspelmath 2017: 85), but the prime symbols (′) used by Strand to indicate stress, probably in approximation of the IPA stress sign (ˈ), directly violate this principle. Replacing them with apostrophes would also contradict the standards set by Haspelmath (2017: 86–87).

7) Highly unusual pronunciation values of English letters are not acceptable.

This principle leaves a bit of space for interpretation, but the pronunciation of <a> as [ɘ], not in cases of vowel reduction but as a normal value in stressed syllables, and contrasted to <â> [a] probably qualifies as "highly unusual", as it is not used in this way in English nor, it seems, in any other of the world's languages.

8) Language names must be pronounceable for English speakers.

This principle is not necessarily violated, because the spellings are more or less readable for English speakers, but it is unlikely that even trained linguists will intuitively produce the intended pronunciation.

9) Language names begin with a capital letter.

This principle is sometimes violated by Strand, especially on his website, but could easily be conformed to without causing inconvenience. Use of lower-case letters after a hyphen is only expressly discouraged for language names made up of two "sub-names", but it is also potentially problematic in the component of the name meaning 'language' (cf. Haspelmath 2017: 88–89).

11) The usage of prominent authors is given substantial weight.

This principle was originally violated by Strand when he diverged from Morgenstierne's conventions but now stands in favor of his newly adopted conventions, as he has become a prominent author himself.

It becomes clear in this way, that the older convention of Morgenstierne fares far better by the standards of Haspelmath (2017) than Strand's proposals, but it has also been noted that the terminological confusion has already been caused and cannot now be reversed by simply returning to the older term. In any case, it may be desirable to reconcile the practical Glottolog principles with Richard Strand's concern for accuracy and the practices of native speakers. This also addresses the concerns voiced by Dryer (2019: 583), that Haspelmath (2017) underemphasizes "the role that nonlinguists play in choice of language names". I therefore propose to set up a consistent terminology that remains mostly recognizable for those used to the established scientific labels, but more closely approximates actual usages by native speakers and at the same time allows a better integration into the metalanguage than Strand's terms, in accordance with the Glottolog principles.

The introduction of a new term violates once again the severely abused principle of least possible name changes, but hopefully contributes to a terminology that will remain stable over a longer period of time, being acceptable to a variety of audiences.

For the largest language of Nuristan, I propose the term **Katë**. This term has the following advantages:

- It is recognizable for audiences used to Morgenstierne's "Kati".
- It accurately represents the phonological form of the native word [kaˈt̪ɘ̈] and is recognizable for native speakers and compatible with their habits.
- It is not a code-switch, the word for 'language' is not included but rather expressed in the metalanguage.[23]
- It can be interpreted and pronounced by native speakers of English and other languages. The character <ë> is an earlier IPA convention for [ə] before the 1993 revision (IPA 1993) and [ë] can still be used in the IPA to represent a centralized close-mid front vowel. The IPA symbol [ə] should not be used in the language name, as it is a special character (cf. Glottolog principle 6)). The grapheme <ë> is known from French and Dutch, though it is used here rather as in Albanian and Luxembourgish (i.e., for a central vowel). Since the letter appears in English names of Albanian cities (Durrës, Shkodër), its use in language names is also acceptable in accordance with Glottolog principle 1) (language names are like city names). In the same way <ë> could be used for central vowels in place names and ethnic names of Nuristan when they are loaned into English or other Latin-script languages.
- Removal of diacritics in contexts where they are unwanted or cause technical difficulties is unproblematic, as the vowel contrast will still be preserved unlike in Strand's transcription.
- It is internally consistent with most of the other language names proposed below in being a term loaned from (northeastern) Katë, historically the "lingua franca" of Nuristan.

I hold that the advantages outweigh the following negative points:

- Another new term is added to the already confusing variety of names.
- The usage of prominent authors (Morgenstierne, Strand) is not followed completely, but rather combined into a compromise.
- It is in origin an ethnically-based name and privileges the name of the largest ethnic group over the others. This is done for continuity with Morgenstierne's term and to avoid an overly long hyphenated name (Katë-Kom-Mumo-Kshto-Binyo-Jamcho-Jashi). Strand's term is only marginally more inclusive by also including the second largest group. Strand (p.c., 2020) voices concerns that such a term "implies a false hegemony of the Kâta tribe over the Kom tribe". While I disagree that this implication necessarily follows from the term, this point could indeed become problematic, if the term were to cross over from scientific into colloquial usage. The current term "Kamviri" for the Southeastern dialect has similar issues, as it is also used by linguists for the language of the Kṣto people (e.g. in Fussman & Ġulāmullāh 1983), who are engaged in a tribal feud with the Kom that has led to the Kom burning down their main village in 1998 and turning its inhabitants into refugees (cf. Strand 1997a). It is unlikely under these circumstances that they would like their speech variety to be named after the Kom. Such complications of tribal politics may be difficult to reconcile with precise terminology. On the other hand, colloquial usage

[23] Many speakers I was in contact with used the contact language's (English, Arabic, Persian) word for 'language'. It should be noted however that language names which include a word for 'language' are established for three languages of the same region: Gawar-Bati [gwt], Khowar [khw], and Brokskat [bkk]. But cf. also Lentz (1937: 273): "Gawar".

> tolerates greater ambiguity than scientific labels and can make use of any labels appropriate to the situation at hand ("Nuristani language", "Kom language", "Katë language", "Kṣto language", etc.).

4.2 Prasun = "Prasuni" or "Vâs'i-vari"

The language denoted by the Glottocode *pras1239* and ISO 639-3 code *prn* is the smallest independent Nuristani language by number of speakers. The entire speaker population belongs to one ethnic group, the self-designation of which is *Wasi* [βaˈsi] and lives in the central valley of Parun (Persian *Pārūn*, native name *Wasigul*) The language is natively called <*Wasi-weri*> (Buddruss & Degener (2017: 5); Strand (1998-2000): <*vâs'i vari*> = [βaˈsiβəri]) 'Wasi language'.

The speakers are first mentioned by Robertson (1896) who described them as

> [...] the Presun tribe, also called Viron by their Musalmán neighbours [...] They inhabit the Presungul, and are entirely different from the Siáh-Posh tribes on the one hand, and from the Wai and the Ashkun people on the other. (Robertson 1896: 78)

"Presun" was probably Robertson's rendering of the Katë name for this group, which is *Přasyu* (Western; Strand 1999a), *Přazĩ̃* (Southeastern; Strand 1999b) or *Přasyũ* (northeastern). The first description of the language appeared in the "Linguistic Survey of India" (Grierson 1919), where it was named "Wasĩ-weri or Veron" and described as being spoken by the "Prēsun". The reason for Grierson's changes to Robertson's "Viron" and "Presun" is not clear, as the "corrected" forms are actually less accurate. The name "Viron"/ "Veron" is taken from the Yidgha name *Wīrōn* (cf. Morgenstierne 1931: 443). "Wasĩ-weri" is clearly based on the native name, the nasalization in *Wasi* also appears in Morgenstierne's notations (Morgenstierne 1949: 190; there presented as dialectal), but could not be confirmed by more recent investigators (cf. Buddruss & Degener 2016: 857). Georg Morgenstierne, in his *Report on a linguistic mission to Afghanistan* (1926), chose the name "Prasun", giving the confusing addition "(Vasi-Veron)". Regarding this choice he writes: "My Kati and Ashkun informants called the valley and the language *Pr'asũ*, *Pr'asun*, *Pr'asĩ̃*, and said that the Parsi [= Persian] name was Parun" (Morgenstierne 1926: 46).

It is possible that the native name cited by the LSI did not appear among the initial data he elicited from his first informant and that he therefore relied on "Kati and Ashkun informants" for an appellation. But the decision could also be interpreted as a conscious return to Robertson's terminology ("Presun"), while correcting the first vowel according to his newer findings. In any case it was the name "Prasun", which would become established in the linguistic literature, particularly following the publication of a more detailed description in Morgenstierne (1949).

In 1956 and again in 1970, Georg Buddruss was able to conduct linguistic fieldwork in the Parun valley, collecting an extensive amount of data, which long remained essentially unpublished. However, excerpts from Buddruss's fieldwork data appeared, e.g., in Buddruss (1960a, 2005), Snoy (1962), Jettmar (1975) and Degener (2016). Other articles relied on Morgenstierne (1949), e.g. Hamp (1969, 1977). All of these publications use the name "Prasun". In 2016 and 2017 the bulk of the fieldwork data was published by Buddruss and Degener under the title *Materialien zur Prasun-Sprache des Afghanischen Hindukusch* (*Materials on the Prasun language of the Afghan Hindu Kush*), Volume I: Texts and glossary (2016) and Volume II: Grammar (2017). These two volumes

are by far the most extensive and authoritative linguistic publications on the language available today. Three recent articles by Hegedűs (2017, 2018, 2020) also use "Prasun".

We therefore have a relatively clear continuity of terminology since the time of Robertson (1896) ("Presun") and a clear majority of the most relevant authors using the name "Prasun". The standard databases ISO 639-3 and Glottolog on the other hand puzzlingly present "Prasuni" as the official name. This is probably an (unintentional?) blend of Persian *Pārūnī* and "Prasun". It has no justification in the history of research or in the linguistic reality, but it has found entrance into some publications (e.g. Koptjevskaja-Tamm & Liljegren 2017: 218).[24]

Another author who diverged from the common usage of "Prasun" is Richard Strand. He revived the use of the native term in the scientific discourse, which had first been introduced by Grierson (1919). Strand (1973: 299) contains the spelling "Wasi-weri", but in his later articles (Strand 2001, 2016) the spelling is changed to fit his personal transcription system ("Vâsi-vari" = [βaˈsiβɽi]) Strand was followed, e.g., by Bashir (2010: 10) (in brackets) and by Heegård (2015: 21) who removed the distinguishing diacritics ("Vasi-Vari") and thereby caused the same problems already observed in the case of Katë.

Again the various terms may be weighed against each other with consideration of the principles set forth by Haspelmath (2017). The principle of least possible name changes is violated about equally as often by "Prasun" (1 spelling change < "Presun" and changing back from "Wasī-Weri") as it is by "Vâsi-vari" (2 spelling changes < "Wasi-Weri" < "Wasī-Weri" and change from "Presun"). Otherwise "Prasun" is consistent with all other principles, whereas "Vâsi-vari" violates 5 further principles:

1) Language names (like city names) are loanwords, not code-switches.
2) Names of non-major languages are not treated differently from names of major languages.

These principles are violated for the same reasons already explained in the section on Katë.

3) Each language has a unique name.

The name cannot be confused with any other language's name but the varying spellings ("Wasi-weri", "Vâsi-vari", "vâsi-vari", "Vasi-Vari") contradict this principle.

7) Highly unusual pronunciation values of English letters are not acceptable.

As mentioned in the section on Katë, <a> for [ɘ] is an unusual pronunciation value, especially when contrasted with <â> for [a].

9) Language names begin with a capital letter.

[24] An anonymous reviewer finds no issue with the form "Prasuni", arguing that "It is pretty common to have a language name ending in -i". However, an adjective suffix *-ī*, which exists in Arabic, Persian and many Indo-Aryan languages, is not common in the morphological systems of the Nuristani languages. It would therefore be unusual to append it to a Nuristani name. If one wishes to use the suffix for some reason, the Persian word *Pārūnī* would be more appropriate.

This is only violated by the variant "vâsi-vari" (see Strand 2001: 258).

11) The usage of prominent authors is given substantial weight.

This principle is violated by Strand's usages in a more significant way for this language than was the case for Katë, since the most extensive and up-to-date publications all use the name "Prasun". In this case the usage of Strand has less of a particular weight, since he has not published any works about the language.

I therefore propose that **Prasun** can remain the conventional name for scientific publications. The inconsistent dialect names given on Glottolog ("Central Prasun", "Lower Prasun", "Upper Wasi-Weri") are better represented simply as Upper, Central and Lower Prasun. This proposal has the following advantages:

- It retains complete continuity with Morgenstierne (1926, 1949) and to some extent with Robertson (1896). It is consistent with the usage of the most extensive and up-to-date publications on the language by Buddruss & Degener (2016, 2017).
- It roughly represents the phonological form of an existing term used in the region (/přasyũ/) and is somewhat recognizable for native speakers who will be familiar with terms used by neighboring peoples.
- It is not a code-switch, the word for 'language' is not included but rather expressed in the metalanguage.
- It can be interpreted and pronounced by native speakers of English and other languages.
- No special characters or diacritics are used.
- It is mostly internally consistent with the other names proposed in being a term loaned from (northeastern) Katë, historically the "lingua franca" of Nuristan.

I hold that these outweigh the following negative points:

- The usage of one author (Strand) is not followed.
- The name is not borrowed from the language itself but rather from the former regional lingua franca, this is done to retain continuity with earlier usage. The Katë term is likely etymologically the same word as the self-designation (cf. Morgenstierne 1931: 443).

4.3 Nuristani Kalasha = "Kalaṣə-alā" or "kalaṣa-alâ"

The language spoken in the southern part of Nuristan known as Kalashüm (*Kalaṣüm* [kalaˈʂym]), which has been given the Glottocode *waig1243* and the ISO 639-3 code *wbk*, has also been known under various designations. Native speakers, depending on the dialect, refer to it as *Kalaṣa-alā* [kʌlʌˌʂʌ.ʌˈlɑ(ː)] (cf. Degener 1998a: 2) or *Kalaṣə-alā* [kʌlʌˌʂə.ʌˈlɑː][25] (cf. Tāza 2017) both of which mean 'Kalaṣa language'. *Kalaṣa* [kʌlʌˈʂʌ] / *Kalaṣə* [kʌlʌˈʂə][26] is a wider ethnic self-

[25] IPA interpretation follows statements by Degener (1998a: 29–30) and Strand (1999c: 239).

[26] According to Strand (2007), the upper dialect's final <ə>, described by him as [ɨ], is a word-final allophone of the phoneme /a/ ("basic vowel"). Under this analysis the transcription <kalaṣa> can be used for both dialectal forms. It

designation in southern Nuristan. Lentz (1937) first reported it in the German spelling "Kalasche" and gave a short description of its use:

> A more comprehensive linguistic and tribal designation seems to be *Kalasche* [...]. The interpretations of the designation given to us are not uniform. It was used both for the tribe and language of the Waigal valley, as well as in a much wider sense, with the inclusion of more or fewer neighboring tribes. [...] It has at any rate no relation to the name of the *Kalasch* people in Chitral (Lentz 1937: 266; own translation)

Samīʿullāh Tāza, himself a native speaker and member of the ethnic group, defines the term as follows:

> It is one of the large tribes (*qabāyil*) of the Nuristani nation (*qawm*), which in a wider sense includes the people of the *Aṣkũ-gəl* and *Wāmā* areas and the Waigal valley district including the three villages of *Tregām* in Nuristan. From a historical perspective the Kalash people of Chitral are also part of it. (Tāza 2017: 1434; own translation)

It follows from this definition that the ethnic term is not just applied to the speakers of the language *wbk*, since the inhabitants of the areas identified here as *Aṣkũ-gəl* and *Wāmā* speak a different language (ISO 639-3: *ask*). While speakers of that language apparently do identify as ethnic *Kalaṣa*, it seems that they do not call their language "*Kalaṣa* language" (see section 4.4).

The status of the language of Tregam, which also has its own ISO 639-3 code (*trm*) is not clear, it may be either a dialect of *wbk* or an independent language (see section 4.5). The "autoglottonym" of this variety is not known.

The "Kalash" people of Chitral across the border in Pakistan, who are today most commonly known as "Kalasha" in English, speak an Indo-Aryan language (ISO 639-3: *kls*), which means that their linguistic relationship with Nuristani is relatively distant. Their language is also most commonly called "Kalasha" in English. In their own language they call themselves <*Kal'as'a*> [kɐˈɻʸɐʂɐ] and their language <*Kal'as'amon*> or <*Kal'as'amandr*> 'Kalasha language'. Their equivalent to Nuristan's "Kalashüm" is <*Kal'as'um*> or <*Kal'as'gum*> 'the Kalasha community'.[27] Pronunciation differences between Chitral's *Kal'as'a* and the Nuristani term *Kalaṣa/Kalaṣə* are slightly different vowel qualities, a velarized [ɻʸ] and stress on the second syllable rather than the third.

The speakers of the language *wbk* are concentrated in the district of Wānt-Waigal, in and around the Kalashüm valley, which has often been called the "Waigal valley" by foreign authors. *Waigal*[28] (Strand: *väigal*) is the largest village of this valley and takes its name from the *Wai* (Strand: *väi*) ethnic subgroup, which according to Strand (1998b) inhabits the upper part of the valley

seems, however, that <a> and <ə> are not in complementary distribution in the dialect of Zhönchigal, cf. *drəgələ* 'long', *draŋə* 'leg' (Tāza 2017).

[27] Words cited from Trail & Cooper (1999); orthographic and IPA representation follows Heegård (2015: 31–33) and Cooper (2005).

[28] Phonemic transcriptions here and in the following taken from Tāza (2017).

comprising the villages of Waigal, Ameshdesh (*Amēšdēš*), Jamich (*J̌amıč,* Strand: *ǰâmač*) and Zhönchigal (*Ẓönčigal/Ẓõčigal*). Tāza (2017: 1847), a native of Zhönchigal, does not quite confirm this,[29] as he attributes the name *Wai* only to the inhabitants of Waigal:

> Historically speaking, the *Wai* are a national group (*grūp-i qawmī*) of the *Kalaṣa* tribe of the Nuristani nation, who are resident in the areas connected to Waigal district, province of Nuristan. Even though they refer to themselves as *Kalaṣə*, the other Nuristani tribes know them under the name *Wai*, which is not correct. Today, in Waigal district the *Wai* are a separate ethnic group which lives in the two villages of the name Waigal.[30]

He instead uses the terms *war-dēšī* and *war-ǰən* in the meaning indicated by Strand for *Wai* (the latter is also mentioned by Strand as a synonym in the form *varǰan*):

> *war-dēšī* - It is said of the people of the upper villages of the Waigal valley, such as Waigal, Zhönchigal, Ameshdesh, Jamich and other small related villages. This word is only used for geographical identification. The people of these villages are sometimes also referred to by the name *war-ǰən*.

Differentiated from this subgroup within the *Kalaṣa/Kalaṣə* grouping are the *Čimiā* and *Nišēi* (Strand: *čima-nišei*). The *Nišēi* live in the village of Nisheygram (*Nišēigrām*), whereas the *Čimiā* inhabit *Čimī*, which includes the villages of Kegal (*Kēgal*), Müldesh (*Mūldēš*) and Akun (*Akūṇ*). An additional, smaller group is formed by the people of *Wānt* (Strand: *vântä*) (cf. Strand 1998b). The terms *war-ǰən* 'upper people' and *bēr-ǰən* 'lower people' are used by Tāza (2017: Introd. 17) to designate the two main dialects. Strand instead calls them *varǰan* and *čima-nišei.*

The first specific name (apart from "Kafir" and both "Safed-Posh" and "Siah-Posh") that was used for the language of Kalashüm by foreign researchers was "Wai-alā", i.e. "Wai language", recorded in the "Linguistic Survey of India" (Grierson 1919: 45), which is the name that was given by the interviewed native speakers in Waigal village. Grierson also incorrectly states "The Wai Kāfirs call the country in which they live 'Waigal'" (Grierson 1919: 45), which certainly contributed to the later confusion of the village with the entire valley.

The term "Wai-alā" was replaced by Morgenstierne in his *Report on a linguistic mission to Afghanistan* (1926) with a Persian adjective formed from the village name, first spelled "Waigeli" and later modified to "Waigali" for his sketch of the language's grammar published 28 years later (Morgenstierne 1954). This replacement of "Wai-alā" was perhaps also inspired by Robertson's (1896: 75) earlier use of "Waigulis" as an ethnic designation. "Waigali" is the form which became widespread in the literature and which was used up until recent times, in particular also by Georg Buddruss who conducted linguistic fieldwork in Nisheygram in 1969, publishing a number of shorter articles as a result (Buddruss 1983, 1987, 1992). Eventually however, Richard Strand, beginning with Strand (1973), brought the self-designation (in its southern dialectal form) back

[29] Discrepancies between the accounts given by Strand (1998b) and Tāza (2017) are in many cases probably the result of microregional differences: Strand gathered most of his information in Nisheygram in the lower part of the valley whereas Tāza mostly refers to practices from the upper part of the valley.
[30] I.e., upper and lower Waigal.

into the scholarly focus, currently usually given in his personal transcription system as "*kalaṣa-alâ*". Analysis of Buddruss's field data was carried on by Almuth Degener (Degener 1994, 1998b, 2001), who did not adopt the use of "Waigali" for the language as a whole. One of her articles (Degener 1998b) is titled "Waigali-Lieder zur Islamisierung Kafiristans (Waigali Songs on the Islamization of Kafiristan)" but the subject of the article is specifically the dialect of the village of Waigal, which she distinguishes as "Waigali" from the "Kalaṣa-alā" of Nisheygram. In her most extensive work on the language of Nisheygram, a grammatical description published in 1998, she explains:

> Morgenstierne (1954) and Buddruss (1987) and (1992) use the language name "Waigali" for a group of dialects of the Waigal valley, among them the dialect of Nisheygram. This term is an expedient choice, since it is established in the scientific literature and is most well-known. But it has the disadvantage that it does not distinguish between the various languages and dialects and that it can also be confused with "Waigali" as a designation for the language of Waigal. [...] The only unambiguous language name is Nišey alā 'language of the inhabitants of Nisheygram' cited by Katz (1982: 51). This does not occur in the texts collected by Buddruss.
>
> In the present work the abbreviation N. therefore stands for "Nišey-alā, i.e. Kalaṣa-alā or Waigali as it is spoken in Nisheygram". By "Waigali" we refer to the dialect of the village of Waigal. (Degener 1998a: 2–3; own translation).

In the title of her work Degener (1998a) did not commit to a particular name, calling the subject of her book simply *Die Sprache von Nisheygram im afghanischen Hindukusch (The language of Nisheygram in the Afghan Hindu Kush)*. Outside of the initial discussion quoted above she almost exclusively uses the abbreviation "N." to refer to the language. This diplomatic solution is a respectable choice and it is true that readers are duly informed about the relevant names in the book, but it is not a sustainable solution for the long-term. Focussing on dialect names in order to evade the naming controversy prevents us from capturing the wider context of the dialects and takes us farther from the goals originally aimed for with the change to self-designations: visibility and recognition of the language communities.

With Degener's grammar, the most detailed linguistic work on the language of Kalashüm published until then, "Kalaṣa-alā" also came into use in Degener's conventional transcription, which is also essentially the same as the one used by Tāza ("Kalaṣə-alā"),[31] allowing for the dialectal vowel difference. On the other hand, the name "Waigali" was also carried on by more recent authors, sometimes with "Kalaṣa-alā" in brackets (e.g., Bashir 2010: 10), sometimes without mention of the native name (e.g., Hegedűs 2002: 189; Hegedűs 2012: 147; Blažek & Hegedűs 2012: 42; Tikkanen 2008: 261; Koptjevskaja-Tamm & Liljegren 2017: 218; Liljegren 2019: 10). "Waigali" is also still the only name recognized by Glottolog and ISO 639-3.

It was in response to this continued practice, that Samīʿullāh Tāza eventually decided to address a formal appeal to international linguists working on Nuristani languages, asking that the name "Waigali" be replaced in linguistic usage by "Kalaṣa-alā", as it was originally based on a misconception (email, February 13th 2020).

[31] In the titles of his own works Tāza uses both the wider label "Nuristani" and (in brackets) the more specific "Kalaṣə-alā".

As with the Katë language, linguists are faced with the conflicting priorities of terminological continuity and respect for local usages. Nevertheless, the facts lie somewhat differently for this language. Judged by the principles of Haspelmath, we find that Morgenstierne's "Waigali" potentially violates 3 of them, 2 more than "Kati".

The introduction of "Waigeli" in the place of "Wai-alā" and then again the change to "Waigali" violated the principle of least possible name changes. The use of "Waigali" against the explicit objections of native speakers violates another principle. As Haspelmath points out "[l]anguage documenters sometimes operate with a vague concept of 'speaker community', but unless the community is organized in a formal way that is widely recognized, reference to the community's views must be interpreted with caution". In this case the community is not formally organized and I cannot objectively evaluate how many speakers object to the term "Waigali", but the opinion of a leading intellectual from the community like Samīʿullāh Tāza, who participates in the scientific discourse and has published several books about the language in Afghanistan, including a comprehensive trilingual dictionary of around 2000 pages, is certainly a decisive factor. In the person of Samīʿullāh Tāza a prominent author comes together with a prominent representative of the community, allowing us to invoke two principles in favor of his proposals. In general, the principle of following the usage of prominent authors is violated by either of the terms in common use, since both have prominent users (Morgenstierne and Buddruss vs. Tāza, Strand and Degener).

As for the term "Kalaṣa-alā" (Degener 1998a), "Kalaṣə-alā" (Tāza 2017), "KalaSa-alâ" (on the same page alongside "Kalasha-alâ") (Strand 2001: 253), or "kalaṣa-alâ" (Strand 1998b), first of all a unification of spelling conventions would be desirable. The variant spellings possibly contradict the principle that each language should have a unique name. Macrons and underdots are explicitly discouraged for language names by Haspelmath (2017: 85–86), as are special characters like <ə> and names that begin with a lower-case letter (Haspelmath 2017: 88). We thus have the options of "Kalashë-ala" and "Kalasha-alâ". If we want to avoid code-switches and treating minor languages differently from major languages, it would be better to remove the word for 'language' from the name and instead to express it in the metalanguage.[32] This creates the problem of distinction from Indo-Aryan "Kalasha" [kls]: Since every language should have a unique name, we can either use the provision of Glottolog principle 10) ("Language names may have a modifier-head structure") and thus distinguish "Nuristani Kalasha" from "Kalasha" or "Indo-Aryan Kalasha", or rely on the vowel difference between "Kalashë" and "Kalasha".[33] Another possible strategy would be again to use a borrowing from Katë, which has the term *Kalṣë́* [kalˈʂə] (cf. Strand 1999b), in which the unstressed second syllable is reduced due to the sound changes of this language. A resulting name form "Kalshë" would allow for maximal contrast with the name of Indo-Aryan Kalasha, which is natively stressed on the second syllable, and would increase the internal consistency of the terminology. However, the introduction of a new exonymic term would not be advisable in response to objections against the exonym "Waigali".

[32] Some have attempted to use the word for 'language' to distinguish "Kalasha-mon" [kls] from "Kalasha-ala" [wbk] (Mørch & Heegård 1997, 2008). This convention has not caught on however (cf. the title of Heegård 2015) and is rather unintuitive. It also creates issues with dialectally different forms of the word "language" (*mon* vs. *mandr* in *kls*).
[33] Tāza (p.c., 2020) states that he has no disagreements with these solutions, though he adds that the term "Kalaṣə-alā" will nevertheless likely continue to be used by writers within Afghanistan.

While acknowledging the acceptability of "Kalashë", I propose that **Nuristani Kalasha** is preferable as the more unambiguous term, which already contains the necessary information required to distinguish it from Indo-Aryan Kalasha. Additionally, speakers of both languages consider themselves to belong to the same ethnic group, a distinction according to differences of pronunciation would therefore somewhat misrepresent the situation.[34] In contexts where no precise distinction is necessary the modifier "Nuristani" may also be abbreviated after the first mention. I propose that the two main dialects could be distinguished as "Upper Nuristani Kalasha" and "Lower Nuristani Kalasha".

In summary, the term "Nuristani Kalasha" holds the following advantages:

- It accurately represents the phonological form of the native term (/kalaṣa/ or /kalaṣə/) and is recognizable for native speakers and compatible with their habits, in this way it improves upon the rejected term "Waigali".
- It retains some continuity with usages like "Kalasha-alā" or "Kalaṣə-alā".
- The modifier allows the immediate identification of the relevant distinguishing feature of the two languages named "Kalasha", thus being easily interpretable also for non-specialists.
- It is not a code-switch, the word for 'language' is not included but rather expressed in the metalanguage.
- It can be interpreted and pronounced by native speakers of English and other languages.
- No special characters or diacritics are used.

I hold that these outweigh the following negative points:

- Another name is introduced to the already confusing array of terms.
- There is no continuity to the earlier "Waigali".
- The usage of prominent authors (Tāza, Degener, Strand, Buddruss, Morgenstierne) is not followed exactly but rather combined into a compromise.

[34] A parallel to this situation can be encountered in the case of Western and Eastern Yugur, two languages spoken in China. Western Yugur is a Turkic language and Eastern Yugur is a Mongolic language, but speakers of both languages together form one ethnic group (cf. Roos 2000: 1–5). "Nuristani Kalasha" and "Indo-Aryan Kalasha" are arguably even better terms than these, since they distinguish the languages by their relevant feature and, in contrast to other methods of distinction, do not require additional explanations such as "Western Kalasha, which is a Nuristani language, and Eastern Kalasha, which is an Indo-Aryan language" or "Kalasha-ala, which is a Nuristani language, and Kalasha-mon, which is an Indo-Aryan language".
An anonymous reviewer argues that this solution "gives the impression that it were a regional dialect of Kalasha [...] Just because the speakers of the variety under consideration identify themselves as (ethnic) 'Kalasha' is by itself not a justification to call what they speak as 'Kalashi' (e.g., if ethnic Kashmiri people shift to speaking Urdu as their first language, we cannot change the name of Urdu language to 'Kashmiri')". The reviewer's criticism takes the established name of Indo-Aryan Kalasha as a natural given, but both peoples, as far as historical records exist, have always called themselves by variants of the "Kalasha" name and we are not in a position to decide whether one of the two groups adopted the name earlier than the other. The hypothetical Kashmiri comparison is therefore not fitting. It is true that the modifier "Nuristani" could also be interpreted as referring to Nuristan in a geographical sense instead of the Nuristani language family. It would not be inappropriate if taken this way, but this interpretation should be precluded by the modifier Indo-Aryan before the language spoken in Chitral. The name "Kalashi" is not proposed in this paper.

- The internal consistency of the terminology is somewhat compromised as this name is not borrowed from Katë. This is done for greater continuity with previous usages and to correspond more closely to the demands of speakers.
- The distinctiveness against Indo-Aryan Kalasha is only somewhat awkwardly achieved by a modifier. For maximal distinctiveness, a modifier "Indo-Aryan" also has to be added to the established "Kalasha". It could be argued that the similarity of the names is quite appropriate, in that it acknowledges the historical and cultural relationship between the peoples speaking these languages, though they are linguistically quite distinct.
- The ethnic grouping *Kalaṣa*, even just within Nuristan, is more inclusive than the group speaking the language given the name "Nuristani Kalasha", as it includes the speakers of the language discussed in section 4.4. This is not a significant issue, since these apparently do not refer to their language by a term meaning 'Kalasha language' (this is confirmed by Tāza (p.c., 2020), cf. also section 4.4.).

4.4 Ashkun or "Ashkun & Wama/Wāmāī" or "ÂSkuNu-saNu-vîri/âṣkuňu-saňu viːri"

This language has been given the Glottocode *ashk1246* and the ISO 639-3 code *ask*. Ethnically the speakers of this language are part of the wider *Kalaṣa* grouping (cf. Strand (1973: 299), Tāza (2017: 1434) quoted above), but are additionally divided into three distinct groups, which Strand identifies as *Aṣkuṇu* (<*âṣkuňu*> = [aʂkuˈʈu]), *Sëṇu* (<*saňu*> = [səˈʈu]) and *Gřamsëṇa* (<*gřâmsaňâ*> = [gɻamsəˈʈa]).[35] The main villages of the *Gřamsëṇa* and *Sëṇu* are often mentioned under their Pashto names, respectively *Ačanu* (natively *Gřamsëṇa-Gřam*) and *Wāmā* (natively *Samë*[36]).

As in the case of Katë, there apparently exists no native term that includes the varieties of all ethnic groups, even though the linguistic relationship is clearly recognized by speakers (cf. Grjunberg 1980: 23). Strand (1998c) argues that his term "*âṣkuňu-saňu vːiri* [sic]" "comes close" to being all-encompassing. In his classification tree he uses only "ÂSkuNu" as the superordinate term (Strand 2001: 258). Locally the terms *Aṣkuṇu-vēri* ('*Aṣkuṇu* language'), *Sëṇu-vīri*[37] ('*Sëṇu* language') and *Gřamsëṇa-vīri* ('*Gřamsëṇa* language') are used, corresponding to the ethnic divisions. Linguistically it appears that the latter two are the same dialect (cf. Grjunberg 1981: 228), though Strand (1998c) separates them from each other due to the ethnic distinction. This points to a dialectal division by watershed into an Alingar dialect and a Pech dialect, though the position of the dialect of Titin is not fully clear.

In foreign publications the earliest records of this language are found in Trumpp (1862, 1866), the later article being a slightly modified German version of the first. The German article also includes a translation of Clark (1865). This, in turn, is a translated summary of the diaries of the Pashtun Christian missionaries Fazl Ḥaqq and Nūrullāh who entered what was then Kafiristan to spread Christianity. In these diaries, originally written in "invisible ink" made from lime juice, some individual words and a short song in the same language were also recorded, which were

[35] Authors disagree over the phonological status of the /ɳ/ in these words. Buddruss (2006: 192) gives the allophones as [ɽ~ɳ], transcribing phonetically as <ṽṛ> or <ṇ>, Strand gives the allophones as [ɻ~ʈ~ɳ] transcribing all as <ň> (Strand 2011).

[36] Following Strand (1998c). Grjunberg (1981: 229) instead gives *Šamə́*, Buddruss (2006: 189): *šamə́*, *šamə̄*.

[37] Following Strand (1998c). Grjunberg (1981: 229) gives *Seṇu virí*, Buddruss (2006: 189): *sə̄ṛú*, Morgenstierne (1952: 117): *säṇ'ū*.

reproduced by Clark (1865: 204). All of these early works refer to the language only with the undistinguished "Kafir" label. Trumpp (1866) gives at least the following qualification:

> That the name Kāfir is meaningless coming from the mouth of the Muhammadans, already results from its meaning. I therefore immediately asked the Kāfirs, whom we shall name this way for the time being until we have learned their proper name, how they call their own country, and received the prompt response: Wāmasthān. (Trumpp 1866: 392; own translation)

He goes on to draw untenable connections between "Wāmasthān" and other place names of the region. In reality this seems to be an extended form of the Pashto village name *Wāmā* meaning 'land of Wama'.[38] A native ethnic designation is first mentioned by Tanner (1881: 290–293) who names the "Sanu Kafirs" (fn.: "The *n* of Sanu is very nasal") and gives some sentences of their language. An additional wordlist is found in Capus (1889: 214–216) who causes some confusion by calling his vocabulary collected in Mashhad (identified as "essentially Ashk[un]" by Morgenstierne (1929: 193)) "Kafir-Siahpouche" and his Katë vocabulary "Kafir-Siahpouch [sic] Bachgali-Loudhé"[39]. All of these scattered publications were not yet understood to represent the same distinct language. A few years later Robertson (1896: 75) still wrote: "There is another important tribe called the Ashkun, of whom, however, it was most difficult to get any information".

This was the first mention of the name "Ashkun", which would establish itself in the literature both for the language and for the people. The "Linguistic Survey of India" (Grierson 1919) contained close to no information on the language, but modified Robertson's designation to "A<u>sh</u>kund", probably based on an ad-hoc folk etymology given by a speaker of a different language.[40] It was not recognized at the time that some of this language had earlier been recorded by the above-mentioned authors. Grierson (1919: 31–32) instead associates Trumpp, Tanner and Capus's data with "Ba<u>sh</u>galī".

Morgenstierne, in his *Report on a linguistic mission to Afghanistan* (1926), returned to Robertson's "Ashkun". This was also the first Nuristani language he described in some detail (Morgenstierne 1929), working with the consultants Abdul Karim from *Majegal* and Abdur Razák, whose native village could not be determined with certainty. Abdur Razák claimed to be from *Titin* but Abdul Karim denied this, claiming that he was in fact from *Wāmā* (cf. Morgenstierne 1929: 192). Based on the linguistic data it appears that Abdur Razák indeed spoke the dialect of *Wāmā/Samɘ*[41] whereas Abdul Karim's speech represents the dialect of the *Aṣkuṇu*.

---

[38] The language in his articles does not show the characteristics of the dialect of *Wāmā/Samë*. The word *glām* rather than *grām* for "village" (Trumpp 1866: 397) is typical of the dialect of the *Aṣkuṇu*. Village names extended with the Persian suffix *-stān* are common in pre-Islamic poetry and songs from the region, cf. Degener (1998b: 387, 529) and Strand (1997c).

[39] "Loudhé" = Khowar *Loṭdeh* = the Katë village of *Brëgomaṭol* (*Barg-i Matāl*).

[40] Grierson (1919: 68): "We know nothing whatever about this dialect except that the word *A<u>sh</u>-kund* means 'bare mountain'." The reported meaning 'bare mountain' and the *d* in the name were later found to be without justification (cf. Morgenstierne 1926: 44).

[41] This does not invalidate his statement that he was from *Titin*. As we have seen in the case of Ivanow's (1931) informant from Kamdesh, people in Nuristan apparently easily adapt to other dialects when they change their places of residence. In fact, Morgenstierne (1929: 192) somewhat strangely writes that Abdul Karim claimed Abdur Razák was "from *Wāmaī*" (not "from *Wāmā*"). The actual claim might therefore just have been that the sentences from Abdur Razák shown to him were in *Wāmāī* (Persian for "*Wāmā*-ese").

Later Morgenstierne published supplements to his foundational article under the titles "Additional Notes on Ashkun" (1934) and "The Wama dialect of Ashkun", part of his "Linguistic Gleanings from Nuristan" (1952).

Lentz relates the experience of the German Hindu Kush Expedition of 1935 in trying to establish the name of the language:

> The village of Wama by the central Pech river according to the common Nuristani view apparently stands on its own. In *Kulam* [Kulem] and in the village of *Waigel* [Waigal] the language had a separate name [...]. In the place itself the independence of the tribe was stressed but a separate name for the language was rejected. This [the language] was said to be identical with *Aschkun* [Ashkun]. (Lentz 1937: 267; own translation)

In 1981, Grjunberg published a number of sentences collected in the 1960s, referring to the language descriptively as *The Language of Wama village* (*Jazyk selenija Vama*) (Grjunberg 1981). More material in this dialect, collected in 1956 and 1970, was published by Buddruss in 2006 together with a short grammar sketch. The name of the language used there is *The Wama language* (*Die Wama-Sprache*). In 2008, Strand published a wordlist of the dialect of *Wāmā/Samɔ* assembled in 1968 on his website under the names "*saňu vi:ri*" and "*saňu-vîri*" (Strand 2008). Recent survey articles by Strand use the combined name "ÂSkuNu-saNu-vîri" (Strand 2001: 253) or "Âṣkuňu-Saňu-vîri" (Strand 2016: 66), though judging by the text on his website he also seems to accept "Ashkun" as a name.[42] Some of the most recent authors have included Strand's usage in brackets (Bashir 2006, 2010: 7) or adopted it while removing diacritics ("Ashkunu" - Heegård 2015: 21). Others have held onto "Ashkun", e.g. Hegedűs (2002: 189, 2012: 147), Blažek & Hegedűs (2012: 42), Tikkanen (2008: 261), Lehr (2014: 175) and Liljegren (2019: 10). The name recognized by Glottolog and ISO 639-3 remains "Ashkun". Additionally, Glottolog indicates as dialect names the puzzling terms "Ashuruveri", "Gramsukraviri" and "Suruviri", which can only possibly be explained as the result of a failed OCR scan of Strand's dialect names.

The possible language names we are therefore faced with today are "Ashkun", "Wama" or "Wamai", "Âṣkuňu", "ÂSkuNu-saNu-vîri" and "Âṣkuňu-Saňu-vîri".

The name "Ashkun", which has been transmitted since the time of Robertson (1896), who probably transcribed it from the speech of Katë speakers, does not violate any of Haspelmath's (2017) principles. The use of "Wama" by Grjunberg and Buddruss was intended to be more precise, since the people of that village do not consider themselves *Aṣkuṇu*, but it violates the principle that new names should not be introduced without absolute necessity. Since *Wāmā* is the Pashto name of the village, it also goes against Strand's preference for native terms. The introduction of "Âṣkuňu", "ÂSkuNu-saNu-vîri" and "Âṣkuňu-Saňu-vîri" by Strand also violates the introduction principle and additionally is incompatible with 7-8 further principles:

1) Language names (like city names) are loanwords, not code-switches.

[42] Quotes: "Ashkun forms with Kalaṣa-alâ and Tregâmi the Southern Group of Nuristâni languages"; "Morgenstierne's field data remain the major source on Ashkun"; "No account of the history of the Ashkun-speaking peoples has been recorded"; "it would appear that Ashkun speakers have spread 'over the top'", all from Strand (1998c).

2) Names of non-major languages are not treated differently from names of major languages.

These principles are violated for the same reasons already explained in the section on Katë.

3) Each language has a unique name.

The name is not confusable with any other language's name but the varying spellings and the concurrent use of "Âṣkuňu", "Âṣkuňu-Saňu-vîri" and "Ashkun" in addition to the dialect names, which are also variable in their spelling, go against this principle.

6) Language names in English are written with ordinary English letters plus some other well-known letters.

The letters *Â* and *î* are permitted by Haspelmath (2017: 85), but <ň> (used in Czech) and <ṣ> are likely not "well-known". The <:> for vowel length in the variant "*saňu vi:ri*" is not permitted by the Glottolog principles.

7) Highly unusual pronunciation values of English letters are not acceptable.

As explained in the section on Katë, <a> for is [ɘ] is quite unusual. The standard interpretation of <ň> would probably be [n], possibly [nʲ] for speakers familiar with Czech. The value [ɳ] is not a usual association. The convention of upper-case letters for retroflexes is also highly unusual in normal English writing.

8) Language names must be pronounceable for English speakers.

The sound [ɳ] is very difficult to pronounce for average English speakers. "Âṣkuňu-Saňu-vîri" is difficult to interpret because of the amount of diacritics. In the case of "ÂSkuNu-saNu-vîri" the visual impression alone will scare off most English speakers from trying to pronounce the term.

9) Language names begin with a capital letter.

A sub-principle of this is violated by the spelling "ÂSkuNu-saNu-vîri", where the second "sub-name" <saNu> starts with a lower-case letter. This cannot be easily remedied when the capital form of the letter is already given the value of the corresponding retroflex. The form "Âṣkuňu-Saňu-vîri" is therefore much more suitable.

11) The usage of prominent authors is given substantial weight.

This principle is violated but also supports the proposals in the same way as already explained for the case of Katë.

It seems therefore that "Ashkun" fares much better than Strand's proposals, and in light of its concurrent adoption by Strand himself and the absence of an all-encompassing self-designation, I propose that "Ashkun" should remain the scientific term for this language. This is consistent

with the system also applied in the case of Prasun in being a name borrowed from Katë, where the word is *Aṣkũ* [aʂˈkũ] (cf. Strand 1999b), which can intuitively be rendered into English spelling conventions as **Ashkun**.

This name has the following advantages:

- It retains complete continuity with Morgenstierne (1926) and even with Robertson (1896).
- It accurately represents the phonological form of an existing term used in the region (/aṣkũ/) and is recognizable for native speakers and mostly compatible with their habits.
- It is not a code-switch, the word for 'language' is not included but rather expressed in the metalanguage.
- It can be interpreted and pronounced by native speakers of English and other languages.
- No special characters or diacritics are used.
- It is internally consistent with the other names proposed in being a term loaned from (northeastern) Katë, historically the "lingua franca" of Nuristan.

I hold that these outweigh the following negative points:

- The usage of one prominent author (Strand) is not followed.
- The name is not borrowed from the language itself but rather from the former regional lingua franca, this is done for continuity with earlier usage and internal consistency of the terminology. It should be noted that Northeastern Katë *Přasyũ* (Southeastern *Přazũ*) and *Aṣkũ* (= southeastern) are not as phonologically parallel to each other as the conventional names "Ashkun" and "Prasun" might imply. Nevertheless, a change to "Prasyun" under violation of continuity principles would not have been worth the increased internal consistency and therefore was not proposed in section 4.2.
- It is in origin an ethnically-based name and privileges the name of the largest ethnic group over the others, this is done for continuity with Morgenstierne's term. Strand's term is only marginally more inclusive by also including the second largest group but not the third. The quoted reports of Lentz (1937) additionally support the view that the *Sëṇu* people do not necessarily object to their language being named after the *Aṣkuṇu*.

4.5 Zemiaki and Tregami or "Gambiri"

These two varieties are severely underdocumented. "Tregami" is currently recognized by ISO 639-3 as an independent language (trm), whereas Glottolog places it together with Nuristani Kalasha in a "Waigali-Tregami" language grouping (Glottocode: *waig1246*). Both databases classify "Zemiaki" as a subvariety of Indo-Aryan Grangali (ISO 639-3: nli; Glottolog: *zemi1238* in *gran1245*). This placement of "Zemiaki" in the Indo-Aryan family is based on outdated information reported in Morgenstierne (1932: 24) ("probably also a dialect of Pashai") and Morgenstierne (1945: 241): "The inhabitants are said to be immigrants from Waigel. But the *Žimakē* or *Žumiaki* dialect may yet belong to the same group as Shum[ashti], and not to the Kafir languages."

Later the following, also outdated, information is added in Morgenstierne (1950: 58):

> [...] in 1949 my Waigeli guide told me that the *Zəmyakō žəba* was identical with *Kuṛaŋgalī*, probably a mistake for *Gṛaŋgalī*, since Kuṛaŋgal is Pashai-speaking. According to Lentz [(1937: 273)], *J̌umiaki* is the name of a village near Gelangel inhabited by emigrants from *Nišēi* in Waigel. It is, however, more likely that a dialect akin to Ningalami is spoken in Gṛaŋgal as well as in "*J̌umiaki*". Only inquiries on the spot can settle the question.

Such inquiries were made in 1963 by A.L. Grjunberg, who could establish that the language, unlike *Gṛangali*, indeed belongs to the Nuristani family and is very close to Nuristani Kalasha, though not identical with any of the known dialects (Grjunberg 1971a: 15–16). The inhabitants were said to have migrated to their present location from the Kalashüm (Waigal) valley after a quarrel some centuries ago (Grjunberg 1971a: 15). He also introduced the language name "Zemiaki" while mentioning various other terms that are in use:

> The area of the upper reaches of the *Čapa-dara* river including its tributary *Sulejmanša* and the village of Zemiaki carries the name *Lindalam* among the local population. The name of the village itself is *Zemiaki* or *Zame* among the Afghans [Pashtuns] and *Džamlam* among its inhabitants. (Grjunberg 1971a: 15)

"Zemiaki" thus appears to be the Pashto name for the village where the language is spoken. On the other hand, Strand (1998b) gives the name of the village as <zamyâ>. More information was given in Grjunberg (1999: 123) where the names of the language are quoted as Pashto *Zemyaki žə́ba* '*Zemyaki* language' and "autoglottonym" *J̌amlám-am bašá* 'language of *J̌amlam*'.

The first foreign mention of "Tregami" seems to be in Lumsden (1860: 124–167), who published a short vocabulary of "the Kaffir Language as spoken in Traieguma and Waigul", though as Morgenstierne points out, it was not possible to separate out the "Tregami" words from this list with the knowledge available at the time. Lentz (1937: 274) found varying pronounciations of the name of "Tregam", the area where "Tregami" is spoken, which was said to mean 'three villages'.

The only description of the language available today is the short sketch by Morgenstierne (1952: 120–135), which appears under the heading "Dialects of Tregām" also including the treatment of the Indo-Aryan language of neighboring *Woṭapūr* and *Kaṭārqalā*. Here he names the language "Gambiri" after the village *Gambīr*, the dialect of which he recorded. Despite the extremely short amount of time Morgenstierne was able to spend with native speakers, his findings allowed him to establish that "Gambiri" is especially closely related to Nuristani Kalasha, so much so that he chose to include the "Gambiri" materials in the vocabulary of his 1954 account of Nuristani Kalasha ("Waigali"). There he notes:

> I have included words from Gambiri in the Voc[abulary] because they, on the whole, agree closely with Waig[ali], and because it is, in most cases difficult to distinguish G from Waig. in the older sources. But, judging from the scraps of information which I was able to gather about the morphology of G, it differs markedly from that of Waig., and on that account G must be reckoned as a separate language. (Morgenstierne 1954: 158–159)

The name "Tregami" came back into use with Buddruss's (1960b: 8) demarcation of that language from the Indo-Aryan language of *Woṭapūr* and *Kaṭārqalā*[43] and with Strand's survey article (1973: 298, 300).

At the present stage of research it is still not clear whether the Upper and Lower dialects of Nuristani Kalasha are more closely related to each other than either is to "Tregami" or "Zemiaki". It may be the case that one or both of the latter two are part of the same level of Nuristani Kalasha dialect variation. More data and further comparative investigation is required to settle this question. As to the naming issue, potential future publications based on new primary data may set a precedent for the conventional naming of these two languages. Until then, I propose that **Tregami** and **Zemiaki** are adequate preliminary designations.

5 Conclusion

The situation surrounding language names in Nuristan is clearly quite complex, not just because of the variety of names in colloquial use, but also because of the inconsistent and disagreeing practices of scientific authors. It is my hope that a conscious reflection of the terminology and its various domains of use (use in scientific publications, native speaker use in the native language, colloquial use in various languages), which I have attempted in this paper, will lead to greater clarity about which practices are appropriate for linguists to follow and which are not. If my suggestions do not meet with approval, I hope that the summary of practices and priorities encountered in the literature will nevertheless provide a frame of reference that can be appealed to when making terminological decisions. In this way terminological justifications will not need to feature quite as prominently in future discussions of these languages.

References

Bashir, Elena. 1988. *Topics in Kalasha syntax: An areal and typological perspective*. Ph.D. diss, University of Michigan.

Bashir, Elena. 2006. "Evidentiality in South Asian languages". In: Butt, Miriam; and Holloway King, Tracy (eds.), Proceedings of the LFG06 conference, workshop on South Asian languages, University of Konstanz, 30–50. Stanford, CA: CSLI Publications.

Bashir, Elena. 2010. "Traces of mirativity in Shina". *Himalayan Linguistics* 9.2: 1–55. doi:10.5070/h99223478.

Beyer, Daniela; and Beck, Simone. 2011. "A linguistic assessment of the Munji language in Afghanistan". *Language Documentation & Conservation* 6: 38–103. https://scholarspace.manoa.hawaii.edu/bitstream/10125/4506/beyer_beck.pdf

Beck, Simone. 2012. "The effect of accessibility on language vitality: The Ishkashimi and the Sanglechi speech varieties in Afghanistan". *Linguistic Discovery* 10.2: 157-233. doi:10.1349/PS1.1537-0852.A.399.

[43] Buddruss (1960b: 8): "[...] Here lie the three villages of *Kaṭar*, *Gambīr* and *Dewōz*, where a distinct Kafir language is spoken, about which I intend to report in a publication soon. It is advisable to use the name *Tregām languages* only for the dialects of these three villages, which stand alone separately and have no relation with the Dardic language of *Kaṭārqalā* described here." The announced publication never appeared.

Blažek, Václav; and Hegedűs, Irén. 2012. "On the position of Nuristani within Indo-Iranian". In: Sukač, Roman; and Šefčík, Ondřej (eds.), *The sound of Indo-European 2: Papers on Indo-European phonetics, phonemics and morphophonemics*, 40–66. Munich: Lincom Europa [LINCOM Studies in Indo-European Linguistics 41].

Buddruss, Georg. 1960a. "Zur Mythologie der Prasun-Kafiren". *Paideuma: Mitteilungen zur Kulturkunde* 7.4/6: 200–209.

Buddruss, Georg. 1960b. *Die Sprache von Woṭapūr und Kaṭārqalā: Linguistische Studien im afghanischen Hindukusch*. Bonn: Selbstverlag des Orientalischen Seminars der Universität Bonn.

Buddruss, Georg. 1977. "Nochmals zur Stellung der Nūristān-Sprachen des afghanischen Hindukusch". *Münchener Studien zur Sprachwissenschaft* 36: 19–38.

Buddruss, Georg. 1983. "Spiegelungen der Islamisierung Kafiristans in der mündlichen Überlieferung". In: Snoy, Peter (ed.), *Ethnologie und Geschichte: Festschrift für Karl Jettmar,* 73–88. Wiesbaden: Franz Steiner Verlag [Beiträge zur Südasienforschung 86].

Buddruss, Georg. 1987. "Ein Ordal der Waigal-Kafiren des Hindukusch". *Cahiers Ferdinand de Saussure* 41: 31–43.

Buddruss, Georg. 1992. "Waigali-Sprichwörter". *Studien zur Indologie und Iranistik* 16: 65–80.

Buddruss, Georg. 2005. "Māra's Mühlenbau. Analyse eines Prasun-Textes aus dem afghanischen Hindukusch". In: Kazanskij, Nikolaj Nikolaevič; Krjučkova, Evgenija Rodionovna; Nikolaev, Aleksandr Sergeevič; and Šackov, Andrej Vladimirovič (eds.), *Hṛdā́ Mánasā. Sbornik statej k 70-letiju so dnja roždenija Professora Leonarda Georgieviča Gercenberga*, 446–469. St. Petersburg: Nauka.

Buddruss, Georg. 2006. "Drei Texte in der Wama-Sprache des afghanischen Hindukusch". In: Bogoljubov, Mixail Nikolaevič (ed.), *Indoiranskoe jazykoznanie i tipologija jazykovyx situacij: sbornik statej k 75-letiju professora A.L. Grjunberga (1930-1995)*, 177–200. St. Petersburg: Nauka.

Buddruss, Georg; and Degener, Almuth. 2016. *Materialien zur Prasun-Sprache des afghanischen Hindukusch. Teil I: Texte und Glossar* Vol. 1. Cambridge, MA: Harvard University Press [Harvard Oriental Series 80].

Buddruss, Georg; and Degener, Almuth. 2017. *Materialien zur Prasun-Sprache des afghanischen Hindukusch, Teil II: Grammatik* Vol. 2. Cambridge, MA: Harvard University Press [Harvard Oriental Series 84].

Cacopardo, Alberto. 2007. *Chi ha inventato la democrazia? Antropologia politica fra Afghanistan e Pakistan*. Ph.D. diss, Università Degli Studi Di Sassari.

Cacopardo, Alberto. 2016. "Fence of Peristan: The islamization of the 'Kafirs' and their domestication". *Archivio per l'Antropologia e la Etnologia* CXLVI: 69–101.

Cacopardo, Alberto; and Cacopardo, Augusto. 2001. *Gates of Peristan: History, religion and society in the Hindu Kush*. Rome: Istituto Italiano per l'Africa e l'Oriente.

Capus, Guillaume. 1889. "Vocabulaire de langues pré-pamiriennes". *Bulletins de la Société d'Anthropologie de Paris* 12: 203–216.

Christensen, Asger. 1980. "The Pashtuns of Kunar: Tribe, class and community organization". *Afghanistan Journal* 7.3: 79-92.

Clark, Robert. 1865. "The Afghan missionaries in Kafiristan". *The Church Missionary Intelligencer* 1: 197–206.

Cooper, Gregory. 2005. *Issues in the development of a writing system for the Kalasha language*. Ph.D. thesis, Macquarie University.

Davidson, John. 1902. *Notes on the Baṣhgalī (Kāfir) language*. Calcutta: Asiatic Society.

Decker, Kendall. 1992. *Languages of Chitral*. Islamabad: National Institute of Pakistani Studies, Quaid-i-Azam University; Summer Institute of Linguistics (SIL) [Sociolinguistic Survey of Northern Pakistan 5].

Degener, Almuth. 1994. "Unauffälliges aus Nisheygram". *Studien zur Indologie und Iranistik* 19 (Festschrift Buddruss): 61–69.

Degener, Almuth. 1998a. *Die Sprache von Nisheygram im afghanischen Hindukusch*. Wiesbaden: Harrassowitz [Neuindische Studien 14].

Degener, Almuth. 1998b. "Waigali-Lieder zur Islamisierung Kafiristans". In: Kušev, Vladimir Vasil'evič; Luzeckaja, Nina Leonidovna; Rzehak, Lutz; and Steblin-Kamenskij, Ivan Mixailovič (eds.), *Central'naja Aziya, Vostočnyj Gindukuš*, 51–61. St. Petersburg: Peterburgskoe Vostokovedenie [Strany i Narody Vostoka XXX].

Degener, Almuth. 2001. "Hunters' lore in Nuristan". *Asian Folklore Studies* 60(2): 329–344.

Degener, Almuth. 2002. "The Nuristani languages". In: Sims-Williams, Nicholas (ed.), *Indo-Iranian languages and peoples*, 103–117. Oxford: Oxford University Press [Proceedings of the British Academy 116].

Degener, Almuth. 2016. "The Prasun verbal system and its terminology". *Historische Sprachforschung* 129: 184–195.

Di Carlo, Pierpaolo. 2016. "Retroflex vowels? Phonetics, phonology, and history of unusual sounds in Kalasha and other languages of the Hindu Kush region". *Archivio per l'Antropologia e la Etnologia* CXLVI: 103–120.

Dryer, Matthew. 2019. "Language names and nonlinguists: A response to Haspelmath". *Language Documentation & Conservation* 13: 580–585. http://hdl.handle.net/10125/24900

Edelberg, Lennart; and Jones, Schuyler. 1979. *Nuristan*. Graz: Akademische Druck- und Verlagsanstalt.

Edelberg, Lennart; and Morgenstierne, Georg. 1974. "Languages in Nuristan and adjacent regions (map)". In: Jettmar, Karl; and Edelberg, Lennart (eds.), *Cultures of the Hindukush: Selected papers from the Hindu-Kush Cultural Conference Held at Moesgård 1970*. Heidelberg: Franz Steiner Verlag [Beiträge zur Südasienforschung 1].

Fussman, Gérard; and Ġulāmullāh. 1983. "Un récit Kam-viri sur les coutumes pré-islamiques de Kamdesh". In: Snoy, Peter (ed.), *Ethnologie und Geschichte: Festschrift für Karl Jettmar*, 180–204. Wiesbaden: Franz Steiner Verlag [Beiträge zur Südasienforschung 86].

Grierson, George. 1911. "Note on Dr. Sten Konow's article on Bashgali". *The Journal of the Royal Asiatic Society of Great Britain and Ireland* Jan.: 195–197.

Grierson, George. 1919. *Specimens of the Dardic or Piśācha languages (including Kāshmīrī)*. Calcutta: Superindendent Government Printing, India [Linguistic Survey of India 8.2].

Grjunberg, Aleksandr Leonovič. 1971a. "K dialektologii dardskix jazykov (Glangali i Zemiaki)". In: Akademija Nauk SSSR - Institut Vostokovedenija (ed.), *Indijskaja i Iranskaja filologija. Voprosy dialektologii*. Moscow: Izdatel'stvo "Nauka", Glavnaja Redakcija Vostočnoj Literatury.

Grjunberg, Aleksandr Leonovič. 1971b. "Nuristan. Etnografičeskie i lingvističeskie zametki". In: Ol'derogge, Dmitrij Alexeevič (ed.), *Srednjaja i Central'naja Azija: Geografija, ėtnografija, istorija*, 264–287. Moscow: Izdatel'stvo "Nauka", Glavnaja Redakcija Vostočnoj Literatury [Strany i Narody Vostoka 10].

Grjunberg, Aleksandr Leonovič. 1980. *Jazyk Kati. Teksty, grammatičeskij očerk*. Moscow: Izdatel'stvo "Nauka" [Jazyki Vostočnogo Gindukuša].

Grjunberg, Aleksandr Leonovič. 1981. "K izučeniju kafirskix jazykov. Jazyk selenija Vama". *Iranskie jazykoznanie, ežegodnik* 1980: 228–231.

Grjunberg, Aleksandr Leonovič. 1994. "Materialy po ėtnografii i mifologii Nuristana iz arxiva M.S. Andreeva". *Studien zur Indologie und Iranistik* 19 (Festschrift Buddruss): 83–107.

Grjunberg, Aleksandr Leonovič. 1995a. "Kafirskij gimn bogu Monu (iz arxiva M.S. Andreeva)". *Peterburgskoe Vostokovedenie* 7: 605–620.

Grjunberg, Aleksandr Leonovič. 1995b. "Nuristanskie teksty iz arxiva M.S. Andreeva". In: Vasil'kov, Jaroslav Vladimirovič; and Gurov, Nikita Vladimirovič (eds.), *Sthāpakaśrāddham. Sbornik statej pamjati G.A. Zografa*, 144–161. St. Petersburg: Peterburgskoe Vostokovedenie.

Grjunberg, Aleksandr Leonovič. 1999. "Zemiaki jazyk/dialekt". In: Ėdel'man, Džoj Josifovna (ed.), *Dardskie i Nuristanskie Jazyki*, 123–124. Indrik [Jazyki Mira 7].

Ġulāmullāh Kāmdešī. 1966. "Girāmir-i zabān-i <ktyh wyry>". Unpublished manuscript.

Halfmann, Jakob. To appear. "Advances in the historical phonology of the Nuristani languages". In: Degener, Almuth; and Hill, Eugen (eds.), Proceedings of the International Workshop in the Languages of Nuristan - The Nuristani Languages in Synchrony and Diachrony, 1–2 November 2019, University of Cologne. *International Journal of Diachronic Linguistics and Linguistic Reconstruction*.

Hammarström, Harald; Forkel, Robert; Haspelmath, Martin; and Bank, Sebastian. 2020. *Glottolog* 4.2.1. Jena: Max Planck Institute for the Science of Human History. doi:10.5281/zenodo.3754591.

Hamp, Eric. 1969. "Two Prasun notes". *Indo-Iranian Journal* 12.1: 24–26.

Hamp, Eric. 1977. "The quality of nasal clusters in Prasun". *Acta Orientalia* 38: 41–42.

Haspelmath, Martin. 2017. "Some principles for language names". *Language Documentation & Conservation* 11: 81–93. http://hdl.handle.net/10125/24725.

Heegård, Jan. 2015. "Kalasha texts - with introductory grammar". *Acta Linguistica Hafniensia* 47. S1: 1–275. doi:10.1080/03740463.2015.1069049.

Hegedűs, Irén. 2002. "Proto-Nuristani *kurka 'young of an animal'". In: Cavoto, Fabrice (ed.), *The linguist's linguist. A collection of papers in honour of Alexis Manaster Ramer*, 189–195. Munich: Lincom Europa.

Hegedűs, Irén. 2012. "The RUKI-rule in Nuristani". In: Nielsen Whitehead, Benedicte; Olander, Thomas; Olsen, Birgit Anette; and Elmegård Rasmussen, Jens (eds.), *The sound of Indo-European: Phonetics, phonemics, and morphophonemics*, 145–168. Copenhagen: Museum Tusculanum Press, University of Copenhagen [Copenhagen Studies in Indo-European 4].

Hegedűs, Irén. 2017. "The etymology of Prasun atˈəg 'one; once, a (little)'". In: Simmelkjær Sandgaard Hansen, Bjarne; Hyllested Adam; Jørgensen, Anders Richardt; Kroonen, Guus; Larsson, Jenny Helena; Nielsen Whitehead, Benedicte; Olander, Thomas; and Mosbæk Søborg, Tobias (eds.),

Usque ad radices: Indo-European studies in honor of Birgit Anette Olsen, 249–260. Copenhagen: Museum Tusculanum Press [Copenhagen Studies in Indo-European 8].

Hegedűs, Irén. 2018. "Progress in Prasun linguistic studies". *Archiv Orientální* 86.1: 177–188.

Hegedűs, Irén. 2020. "The etymology of Prasun üpün 'one'". In: Bichlmeier, Harald; Šefčík, Ondřej; and Sukač, Roman (eds.), *Etymologus. Festschrift for Václav Blažek*, 201–214. Hamburg: Baar [Studien zur Historisch-Vergleichenden Sprachwissenschaft 14].

Herrlich, Albert; Lentz Wolfgang; Roemer, Werner; and von Rosenstiel, Klaus. 1937. "Geographische Bemerkungen". In: Scheibe, Arnold (ed.), *Deutsche im Hindukusch. Bericht der deutschen Hindukusch-Expedition 1935 der Deutschen Forschungsgemeinschaft*, 295–351. Berlin: Karl Siegismund Verlag [Deutsche Forschung: Schriften der Deutschen Forschungsgemeinschaft (Neue Folge) 1].

International Phonetic Association. 1993. "Council actions on revisions of the IPA". *Journal of the International Phonetic Association* 23.1: 32–34. doi:10.1017/S002510030000476X.

Irgens-Møller, Christer. 2009. *Music in Nuristan. Traditional music from Afghanistan. An investigation of the field recordings of Lennart Edelberg and Klaus Ferdinand 1953-54, 1964, 1970*. Moesgaard: Moesgaard Museum, Jutland Archeological Society [Jutland Archeological Society Publications 68].

Ivanow, Wladimir. 1931. "A specimen of Bashgali from Kamdesh". *Acta Orientalia* 10: 576–595.

Jano, Yāsamīn Akmal. 1991. *Afsānahā-yi Nūristānī*. Kabul: Center for Languages and Literature, Afghan Academy of Sciences.

Jarvis, Andrew; Reuter, Hannes Isaak; Nelson, Andrew; and Guevara, Edith. 2008. *Hole-filled seamless SRTM data V4*. International Centre for Tropical Agriculture (CIAT). http://srtm.csi.cgiar.org.

Jettmar, Karl; Jones, Schuyler; and Klimburg, Maximilian. 1975. *Die Religionen des Hindukusch*. Stuttgart: Kohlhammer [Die Religionen der Menschheit 4/1].

Katz, David. 1982. *Kafir to Afghan: Religious conversion, political incorporation, and ethnicity in the Väygal valley, Nuristan*. Ph.D. diss, University of California, Los Angeles.

Konow, Sten. 1911. "Notes on the classification of Bashgali". *The Journal of the Royal Asiatic Society of Great Britain and Ireland* Jan.: 1–47.

Konow, Sten. 1913. *Bashgali dictionary. An analysis of Colonel J. Davidson's notes on the Bashgali language*. New Delhi: Gian Publishing House.

Koptjevskaja-Tamm, Maria; and Liljegren, Henrik. 2017. "Semantic patterns from an areal perspective". In: Hickey, Raymond (ed.), *The Cambridge Handbook of Areal Linguistics*, 204–236. Cambridge: Cambridge University Press. doi:10.1017/9781107279872.009.

Lehr, Rachel. 2014. *A descriptive grammar of Pashai: The language and speech community of Darrai Nur*. Ph.D. diss, University of Chicago.

Leitner, Gottlieb Wilhelm. 1880. *Kafiristan, section 1. The Bashgeli Kafirs and their language*. Lahore: Dilbagroy.

Lentz, Wolfgang. 1937. "Sprachwissenschaftliche und völkerkundliche Studien in Nuristan". In: Scheibe, Arnold (ed.), *Deutsche im Hindukusch. Bericht der deutschen Hindukusch-Expedition 1935 der Deutschen Forschungsgemeinschaft*, 247–284. Karl Siegismund Verlag [Deutsche Forschung: Schriften der Deutschen Forschungsgemeinschaft (Neue Folge) 1].

Liljegren, Henrik. 2016. *A grammar of Palula*. Berlin: Language Science Press [Studies in Diversity Linguistics 8]. doi:10.26530/oapen_611690.

Liljegren, Henrik. 2019. "Gender typology and gender (in)stability in Hindu Kush Indo-Aryan languages". In: Di Garbo, Francesca; Olsson, Bruno; and Wälchli, Bernhard (eds.), *Grammatical gender and linguistic complexity* Volume I: General issues and specific studies, 279–328. Berlin: Language Science Press. doi:10.5281/zenodo.3462772.

Liljegren, Henrik; and Naseem Haider. 2011. *Palula vocabulary*. Islamabad: FLI (Forum for Language Initiatives) [FLI Language and Culture Series 7].

Lipp, Reiner. 2009. *Die indogermanischen und einzelsprachlichen Palatale im Indoiranischen. Neurekonstruktion, Nuristan-Sprachen, Genese der indoarischen Retroflexe, Indoarisch von Mitanni* Volume I. Heidelberg: Universitätsverlag Winter.

Lumsden, Harry Burnett. 1860. *The mission to Kandahar*. Calcutta: C.B. Lewis, Baptist Mission Press.

Mayrhofer, Manfred. 1984. *Lassen sich Vorstufen des Uriranischen nachweisen?*. Vienna: Österreichische Akademie der Wissenschaften [Veröffentlichungen der Iranischen Kommission 13].

Mayrhofer, Manfred. 1992. *Etymologisches Wörterbuch des Altindoarischen* Volume I. Heidelberg: Universitätsverlag Winter [Indogermanische Bibliothek, 2. Reihe: Wörterbücher].

Mohammad, Jan. 1991. *Causative constructions in Kati, a Nuristani language of Afghanistan*. Master's thesis, Ohio State University.

Mørch, Ida Elisabeth; and Heegård, Jan. 1997. *Retroflekse vokalers oprindelse i kalashamon i historisk og areallingvistisk perspektiv; Variation i sprogbeskrivelsen: vokallængde i kalashamon*. Master's thesis, University of Copenhagen.

Mørch, Ida Elisabeth; and Heegård, Jan. 2008. "Some observations on the variation in the pronunciation of Kalashamon as spoken inside and outside present-day Kalasha society". In: Israr-ud-Din (ed.), Proceedings of the Third International Hindu Kush Cultural Conference, 288–295. Oxford: Oxford University Press.

Morgenstierne, Georg. 1926. *Report on a linguistic mission to Afghanistan*. Instituttet for sammenlignende kulturforskning [Publikasjoner, Serie C 1.2].

Morgenstierne, Georg. 1929. "The language of the Ashkun Kafirs". *Norsk Tidsskrift for Sprogvidenskap* 2: 192–289.

Morgenstierne, Georg. 1931. "The name 'Munjān' and some other names of places and peoples in the Hindu Kush". *Bulletin of the School of Oriental Studies, University of London* 6.2: 439-444.

Morgenstierne, Georg. 1932. *Report on a linguistic mission to north-western India*. Instituttet for sammenlignende kulturforskning [Publikasjoner, Serie C 3.1].

Morgenstierne, Georg. 1934. "Additional notes on Ashkun". *Norsk Tidsskrift for Sprogvidenskap* 7: 56–115.

Morgenstierne, Georg. 1942. "Notes on Dameli, a Kafir-Dardic language of Chitral". *Norsk Tidsskrift for Sprogvidenskap* 12: 113–198.

Morgenstierne, Georg. 1945. "Notes on Shumashti". *Norsk Tidsskrift for Sprogvidenskap* 13: 239–281.

Morgenstierne, Georg. 1949. "The language of the Prasun Kafirs". *Norsk Tidsskrift for Sprogvidenskap* 15: 187–334.

Morgenstierne, Georg. 1950. *Notes on Gawar-Bati*. Oslo: Jacob Dybwad [Skrifter utgitt av det Norske Videnskabs-Akademi i Oslo. II: Hist.-Filos. Klasse 1].

Morgenstierne, Georg. 1952. "Linguistic gleanings from Nuristan". *Norsk Tidsskrift for Sprogvidenskap* 16: 117–135.

Morgenstierne, Georg. 1954. "The Waigali language". *Norsk Tidsskrift for Sprogvidenskap* 17: 146–324.

Morgenstierne, Georg. 1967. "Some folk-songs from Nuristan". In: *To honor Roman Jakobson: Essays on the occasion of his 70. birthday, 11. October 1966* Volume 2, 1378–1392. Berlin: De Gruyter [Janua Linguarum, Series Maior 32].

Morgenstierne, Georg. 1973. "Die Stellung der Kafirsprachen". In: Morgenstierne, Georg (ed.), *Irano-Dardica*, 327–343. Wiesbaden: Ludwig Reichert Verlag [Beiträge Zur Iranistik 5].

Nelson, David. 1986. *The historical development of the Nuristani languages*. Ph.D. thesis, University of Minnesota.

Parsons, Martin. 1995. *Ramgal valley, western Nuristan: Regional development report*. doi:10.2458/azu_acku_pamphlet_hn670_6_z9_p37_1995.

Perder, Emil. 2013. *A grammatical description of Dameli*. Ph.D. thesis, University of Stockholm.

Raverty, Henry George. 1864. "On the language of the Siah Posh Kafirs". *Journal of the Asiatic Society of Bengal* 33.3: 267–278.

Robertson, George Scott. 1896. *The Káfirs of the Hindu-Kush*. London: Lawrence & Bullet.

Roos, Martina Erica. 2000. *The western Yugur (Yellow Uygur) language. Grammar, texts, vocabulary*. Ph.D. diss, Leiden University.

Snoy, Peter. 1962. *Die Kafiren. Formen der Wirtschaft und der geistigen Kultur*. Ph.D. diss, Goethe-Universität Frankfurt.

Strand, Richard. 1973. "Notes on the Nūristānī and Dardic languages". *Journal of the American Oriental Society* 93.3: 297–305.

Strand, Richard. 1985. "Locality and nominal relationships in Kamviri". In: Zide, Arlene; Magier, David; and Schiller, Eric (eds.), Proceedings of the Conference on Participant Roles: South Asia and Adjacent Areas, 48–57. Bloomington, IN: Indiana University Linguistics Club.

Strand, Richard. 1997a. *The kât'a, k'om, mum'o, kṣt'o, bini'o, ǰâmč'o, and ǰâš'a* (Last edited 2011), at http://nuristan.info/Nuristani/Kamkata/kamkata.html Accessed Mar 29, 2021.

Strand, Richard. 1997b. *The sound system of kâmv'iri* (Last edited 2011), at http://nuristan.info/Nuristani/Kamkata/Kom/KomLanguage/Lexicon/phon.html Accessed Mar 29, 2021.

Strand, Richard. 1997c. *The Koms' History related by Qâzi Ghulâmullâh and Vakil Muhammad Kabir* (Last edited 2019), at http://nuristan.info/Nuristani/Kamkata/Kom/KomTexts/KomHist.html Accessed Mar 29, 2021.

Strand, Richard. 1998a. *The vâsi* (Last edited 2000), at http://nuristan.info/Nuristani/Vasi/vasi.html Accessed Mar 29, 2021.

Strand, Richard. 1998b. *The kalaṣa of kalaṣüm* (Last edited 2011), at http://nuristan.info/Nuristani/Kalasha/kalasha.html Accessed Mar 29, 2021.

Strand, Richard. 1998c. *The âṣkuňu, saňu, and gřâmsaňâ* (Last edited 2011), at http://nuristan.info/Nuristani/AshkunEtc/ashkun.html Accessed Mar 29, 2021.

Strand, Richard. 1999a. *kât'a vari (kt'i vřâ·i vari) Lexicon* (Last edited 2011), through http://nuristan.info/lngFrameL.html (no direct link) Accessed Mar 29, 2021.

Strand, Richard. 1999b. *kâmv'iri Lexicon* (Last edited 2019), through http://nuristan.info/lngFrameL.html (no direct link) Accessed Mar 29, 2021.

Strand, Richard. 1999c. "Review of Degener (1998): Die Sprache von Nisheygram im afghanischen Hindukusch". *Acta Orientalia* 60: 236–243.

Strand, Richard. 2001. "The tongues of Peristan". In: Cacopardo, Alberto; and Cacopardo, Augusto (eds.), *Gates of Peristan: History, religion and society in the Hindu Kush*, 251–257. Rome: Istituto Italiano per l'Africa e l'Oriente.

Strand, Richard. 2007. *Transcription and pronunciation of the Nuristâni languages* (Last edited 2011), at http://nuristan.info/Nuristani/phon.html Accessed Mar 29, 2021.

Strand, Richard. 2008. *saňu vi:ri Lexicon (Last edited 2011)*, through http://nuristan.info/lngFrameL.html (no direct link) Accessed Sept 30, 2020.

Strand, Richard. 2011. *The sound system of saňu-vi:ri*, at http://nuristan.info/Nuristani/AshkunEtc/SaNu/SaNuLanguage/Lexicon/phon.html Accessed Mar 29, 2021.

Strand, Richard. 2013. *Basic processes in the evolution of the Nûristânî languages*, at http://nuristan.info/Nuristani/BasicEvolutionaryProcesses.html Accessed Mar 29, 2021.

Strand, Richard. 2016. "Nûristânî". In: Hock, Hans Henrich; and Bashir, Elena (eds.), *The languages and linguistics of South Asia: A comprehensive guide*, 66–73. Berlin: De Gruyter Mouton [The World of Linguistics 7].

Sun-Aro, Sulaymān. 2016. *Farhang-i Nūristānī*. Kabul: Samar (SIL).

Tanner, Henry Charles. 1881. "Notes on the Chugáni and neighbouring tribes of Kafiristan". Proceedings of the Royal Geographical Society and Monthly Record of Geography 3.5: 278–301.

Tāza, Samīʿullāh. 2017. *Farhang-i Zabān-i Nuristānī (Kalaṣə-alā): Maʿnā wa tašrīḥ-i luġāt, bā zabānhā-yi Paχto wa Darī*. Kabul: Summer Institute of Linguistics (SIL).

Tikkanen, Bertil. 2008. "Some areal and phonological isoglosses in the transit zone between South and Central Asia". In: Israr-ud-Din (ed.), Proceedings of the Third International Hindu Kush Cultural Conference, 250–262. Oxford: Oxford University Press.

Trail, Ronald; and Cooper, Gregory. 1999. *Kalasha dictionary - with English and Urdu*. Islamabad: National Institute of Pakistani Studies, Quaid-i-Azam University [Studies in Languages of Northern Pakistan].

Trumpp, Ernst. 1862. "On the language of the so-called Kāfirs of the Indian Caucasus". *The Journal of the Royal Asiatic Society of Great Britain and Ireland* 19: 1–29.

Trumpp, Ernst. 1866. "Über die Sprache der sogenannten Kāfirs im indischen Caucasus (Hindū Kūsch)". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 20.2/3: 377–418.

Jakob Halfmann
jhalfma2@smail.uni-koeln.de