Hindu Kush, or the Greater Hindu Kush, in this context is really a shorthand for the remote region where the ranges of the Hindu Kush, the Karakoram, the Pamirs and the westernmost extension of the Himalayas meet (Liljegren 2014: 134–138; Bashir 2016: 264). These northwestern outskirts of the subcontinent are inhabited by at least 50 distinct ethnolinguistic communities (Hammarström, Forkel & Haspelmath 2017; Lewis, Simons & Fennig 2016). The Hindu Kush, in this sense, is part of the territories of several countries -- primarily Afghanistan, Pakistan and India. The geographically most salient feature is its mountainous environment, especially vis-à-vis the Indo-Gangetic plains situated south of it. While being a transit zone of sorts between the cultural spheres of South Asia, Central Asia, West Asia and the Himalayas, this is simultaneously the easternmost extension of Iranian languages, the northernmost extension of Indo-Aryan languages as well as the westernmost extension of Sino-Tibetan. Apart from those three phylogenetic components, the region is also home to Nuristani, at least two Turkic language enclaves and the language isolate Burushaski. These six major phylogenies of the linguistic landscape of the Hindu Kush will be introduced briefly.

**Indo-Aryan** (which along with Iranian and Nuristani belongs to the larger Indo-Iranian branch of Indo-European) is the largest phylogenetic component, making up at least half of the languages in the Hindu Kush region, relatively evenly distributed in a southern belt stretching from east to west. Those can be grouped into at least nine relatedness clusters or groups (Pashai, Kunar, Chitral, Kohistani, Shina, Kashmiri, Western Punjabi, Rajasthani, and Central), although the exact placement of a few of them remains uncertain (Strand 1973: 207–208; 2001: 251; Bashir 2003). In the past, the label “Dardic” was collectively applied to languages belonging to the six first-mentioned groups, all of them Northwestern Indo-Aryan languages, with a longstanding presence in the region. That label is, however, no longer relevant as a classificatory entity (Morgenstierne 1961). The region’s Western Punjabi varieties (such as Pahari-Pothwari and Hindko) are really part of a larger Punjabi continuum with an extension far south of the region, and as such probably have more in common with the closest main Indo-Aryan languages of the Indo-Pakistani plains than with most of the other Indo-Aryan varieties in the region. Gojri (Rajasthani) is the language of nomadic or semi-nomadic Gujurs, today a language spoken in pockets throughout the region and beyond, with a significant concentration in Kashmir, whereas its closest Rajasthani relatives are found at a considerable distance, deep into the main belt of Indo-Aryan (Sharma 2002: 44–45). Domaaki (Central) is a relative newcomer to the region. As the language of a small enclave of musicians and blacksmiths, it has during its 200-300 years in the area acquired a number of features typical of neighbouring and locally dominant languages (Weinreich 2011: 165–166). Like Gojri, its closest relatives are found in the plains of North India.

**Iranian** is the second largest component, if counted in number of distinct language varieties. Most of those languages are spoken in the western half of the region, primarily in Afghanistan, but also in Pakistan and in adjacent areas of Tajikistan and China. A number of them are spoken next to Indo-Aryan languages, and patterns of bilingualism involving Indo-Aryan and Iranian alike have most likely existed for a prolonged period. While the various varieties and sub-varieties of Dari (Afghan Persian) or Tajik that are spoken in the region are considered Western Iranian, all the other Iranian language varieties represented here are Eastern (Skjaervø 1989: 370). Pashto, the numerically dominant of those and also the geographically most wide-spread, is distantly related to the so-called “Pamiri” languages, a relatively loose cluster of Iranian languages primarily spoken in the Pamirs and surrounding areas (Payne 1989: 417; Èdel’man & Dodykhudoeva 2009: 773). While Parachi, a small language community in the southwestern part of the Hindu Kush region, in general terms also belong to East Iranian, and more specifically to a south-eastern subgroup, the classification of this language poses problems of a particular kind, as there are features revealing deep historical affinities with Northwestern Iranian (Skjaervø 1989: 370; Kieffer 2009: 693; Morgenstierne 1926: 27–28). Parachi is only closely related to one other Iranian language, namely Ormuri, a small linguistic island surrounded by Pashto in Pakistan’s Waziristan area.

**Nuristani** is spoken in a system of remote valleys in a confined area of northeastern Afghanistan, with some minor spill-over into adjacent areas of Pakistan. This group, with its 5-6 languages, forms a third, albeit numerically minimal, branch of Indo-Iranian (Strand 1973: 297–298; Morgenstierne 1961: 139). While it is neither Iranian nor Indo-Aryan, the historical and interactional relationship between the neighbouring Indo-Aryan communities and the Nuristani communities is complex and far from straight-forward and the exact point at which Nuristani branched off from the rest of Indo-Iranian is a far from settled issue (Degener 2002; Zoller 2005: 10–15). Nuristani can be further sub-classified into a northern and a southern group (Strand 2001: 258–259; Buddruss & Degener 2017: 13).

**Sino-Tibetan** is present in the eastern-most part of the Hindu Kush region (as defined above), in Pakistan’s Baltistan and in Ladakh on the Indian side of the so-called Line-Of-Control. The 3-4 varieties represented are all part of a Western Tibetan grouping and forms the geographic westernmost fringe of a phylogenetic entity comprising hundreds of languages spoken throughout the highlands of South East Asia and beyond. The two with one another closely related language varieties Balti and Purik are in comparison with Tibetan in general considered particularly archaic (Bielmeier 1985: 14–15), but give at the same time evidence of developments resulting from intensive contact with neighbouring Indo-Aryan languages (Zeisler 2005: 57).

**Turkic** is represented by two Turkic language enclaves with a relatively stable presence in the northernmost part of the region: Kyrgyz, in the far-flung highlands of the Wakhan corridor, where it is spoken by herdsmen; and Uzbek in Afghan Badakhshan. Those languages belong to two separate branches of Turkic, Kipchak (Kirchner 1998: 344), and Uyhgur-Karluk (Johanson 2006: 287–288; Boeschoten 1998: 357), respectively. The southern sub-variety of Uzbek that is spoken in the region displays a great deal of interference from Iranian languages (Reichl 1983: 31).

**Burushaski**, finally, is for all practical purposes still best considered a language isolate, although having been the subject of numerous attempts at tying it to various families of the world, such as Indo-European (Čašule 2009), Yeniseian in Siberia (Toporov 1971) and some of the language families in the Caucasus (Holst 2017). The presence of the language in this particular region is without doubt of a very old age, possibly surpassing most of the languages or phyla mentioned above, and there is reason to believe that it once had a considerably larger geographical distribution (Tikkanen 1988: 305). Many of the surrounding languages have left traces, not the least in its lexicon (Berger 1998: 3), and there is considerable convergence between Burushaski and quite a few of the neighbouring languages, whether Indo-Aryan, Iranian or Tibetan (Tikkanen 1988: 310; Hock 2015: 128). Today, it is spoken by less than 100 000 people in the extreme North of Pakistan, close to the border of China.

The linguistic profile of this region and its significance as a contact zone or linguistic area has been the topic of a discussion going on for several decades (Toporov 1970; Èdel’man 1980; 1983: 16; Fussman 1972: 389–399; Bashir 1996a; 1996b; 2003: 823; 2016; Baart 2014; Tikkanen 1999; 2008; Liljegren 2017; Koptjevskaja-Tamm & Liljegren 2017: 215–223; Emeneau 1965; Skalmowski 1985; Masica 1991: 43; 2001: 259), but the tendency has been to focus on individual features and phenomena, sometimes based on relatively sparse data, and more seldom have there been attempts at applying a higher degree of feature aggregation with tight sampling.

In the present study, we collected and analyzed comparable primary data from a very tight sample of 59 Hindu Kush-Karakoram language varieties (for practical reasons limited to Afghanistan, Pakistan and India). All the six phylogenies referred to above were represented, as were most of the relevant subgroupings. For some of the languages spoken across international borders we aimed at representing such a language as used in these different countries. Unfortunate circumstances and a tight schedule made it impossible to collect data from a few languages that were initially targeted, such as Domaaki (Indo-Aryan), Chilisso (Indo-Aryan), Pothwari-Pahari as spoken in Kashmir (Indo-Aryan), Poguli Kashmiri (Indo-Aryan), and Tregami (Nuristani). A few other languages previously listed for the region are most likely moribund or spoken by only a few individuals, such as Tirahi (Indo-Aryan), Grangali (Indo-Aryan), Shumashti (Indo-Aryan), and Gowro (Indo-Aryan)

The data collected allowed for setting up a basic word list of comparable meanings (representing kinship, numerals, basic actions, substances and objects) as well as for classifying each variety according to a number of binary structural features (reflecting phonological, morphological, syntactic and lexico-semantic properties). As is expanded on in publications yet to appear, a comparison of the basic lexicon across theses languages lines up closely with established phylogenetic classification, whereas clustering of structural similarity is clearly related to geographical proximity within the region and often cuts across phylogenetic boundaries. The strongest evidence of areality tied to the region itself (vis-à-vis South Asia in general on the one hand and Central/West Asia on the other) relates to phonology and lexical structure, whereas word order and alignment features mostly place the region’s languages within larger areal or macro-areal distributions. Some morphological properties related to grammatical categorization (e.g. gender) display a high degree of genetic stability.

References

Baart, Joan L. G. 2014. Tone and stress in North-West Indo-Aryan: A survey. In Johanneke Caspers, Yiya Chen, Willemijn Heeren, Jos Pacilly, Niels O. Schiller & Ellen van Zanten (eds.), *Above and Beyond the Segments*, 1–13. Amsterdam: John Benjamins Publishing Company.

Bashir, Elena. 1996a. Mosaic of tongues: Quotatives and complementizers in Northwest Indo-Aryan, Burushaski, and Balti. In William L Hanaway & Wilma Heston (eds.), *Studies in Pakistani popular culture*, 187–286. Lahore: Lok Virsa Pub. House and Sang-e-Meel Publications.

Bashir, Elena. 1996b. The areal position of Khowar: South Asian and other affinities. In Elena Bashir & Israr-ud-Din (eds.), *Proceedings of the Second International Hindukush Cultural Conference* (Hindukush and Karakoram Studies v. 1), 167–179. Karachi: Oxford University Press.

Bashir, Elena. 2003. Dardic. In George Cardona & Danesh Jain (eds.), *The Indo-Aryan Languages*, 818–894. London: Routledge.

Bashir, Elena. 2016. Pre-1947 convergences. In Hock Hans Henrich & Bashir Elena (eds.), *The Languages and Linguistics of South Asia, A Comprehensive Guide*, 264–284. Berlin, Boston: De Gruyter Mouton. doi:10.1515/9783110423303. https://www.degruyter.com/view/product/429184 (22 November, 2018).

Berger, Hermann. 1998. *Die Burushaski-Sprache von Hunza und Nager 1. Grammatik* (Neuindische Studien 13). Wiesbaden: Harrassowitz.

Bielmeier, Roland. 1985. *Das Märchen vom Prinzen Čobzaṅ: eine tibetische Erzählung aus Baltistan: Text, Übersetzung, Grammatik und westtibetisch vergleichendes Glossar* (Beiträge Zur Tibetischen Erzählforschung Bd. 6). Sankt Augustin: VGH Wissenschaftsverlag.

Boeschoten, Hendrik. 1998. Uzbek. In Lars Johanson & Éva Á Csató (eds.), *The Turkic Languages*, 357–378. Reprint edition. London: Routledge.

Buddruss, Georg & Almuth Degener. 2017. *Materialien zur Prasun-Sprache des Afghanischen Hindukusch 2. Grammatik* (Harvard Oriental Series 84). Cambridge, MA: Harvard University Press.

Čašule, Ilija. 2009. Burushaski Numerals of Indo-European Origin. *Central Asiatic Journal* 53(2). 163–182.

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

Èdel’man, Džoi Iosifovna. 1980. K substratnomu naslediju central’no-aziatskogo jazykovogo sojuza. *Voprosy jazykoznanija* 5. 21–32.

Èdel’man, Džoi Iosifovna. 1983. *The Dardic and Nuristani languages* (Languages of Asia and Africa). Moscow: Nauka.

Èdel’man, Džoi Iosifovna & Leila R. Dodykhudoeva. 2009. The Pamir languages. In Gernot Windfuhr (ed.), *The Iranian Languages*, 773–786. 1st edn. London; New York: Routledge.

Emeneau, Murray Barnson. 1965. *India and historical grammar*. Annamalainagar.

Fussman, Gérard. 1972. *Atlas linguistique des parlers dardes et kafirs*. Paris: École française d’Extrême-Orient; Dépositaire: Adrien-Maisónneuve.

Hammarström, Harald, Robert Forkel & Martin Haspelmath (eds.). 2017. *Glottolog 3.0*. Jena: Max Planck Institute for the Science of Human History. http://glottolog.org/ (29 August, 2017).

Hock, Hans Henrich. 2015. The Northwest of South Asia and beyond: The issue of Indo-Aryan retroflexion yet again. *Journal of South Asian Languages and Linguistics* 2(1). 111–135. doi:10.1515/jsall-2015-0005.

Holst, Jan Henrik. 2017. *Die Herkunft des Buruschaski* (Kaukasien-Reihe). Aachen: Shaker Verlag.

Johanson, Lars. 2006. Uzbek. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 287–290. Amsterdam [etc.]: Elsevier.

Kieffer, Charles M. 2009. Parachi. In Gernot Windfuhr (ed.), *The Iranian Languages*, 693–720. London: Routledge.

Kirchner, Mark. 1998. Kirghiz. In Lars Johanson & Éva Á Csató (eds.), *The Turkic Languages*, 344–356. Reprint edition. London: Routledge.

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

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2016. *Ethnologue: Languages of the World, Nineteenth edition*. Online version. Dallas, Texas: SIL International. http://www.ethnologue.com (21 February, 2017).

Liljegren, Henrik. 2014. A survey of alignment features in the Greater Hindukush with special references to Indo-Aryan. In Pirkko Suihkonen & Lindsay J. Whaley (eds.), *On diversity and complexity of languages spoken in Europe and North and Central Asia* (Studies in Language Companion Series 164), 133–174. Amsterdam: John Benjamins Publishing Company.

Liljegren, Henrik. 2017. Profiling Indo-Aryan in the Hindukush-Karakoram: A preliminary study of micro-typological patterns. *Journal of South Asian Languages and Linguistics* 4(1). 107–156. doi:10.1515/jsall-2017-0004.

Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.

Masica, Colin P. 2001. The definition and significance of linguistic areas: Methods, pitfalls, and possibilities (with special reference to the validity of South Asia as a linguistic area). In Peri Bhaskararao (ed.), *The yearbook of South Asian languages and linguistics 2001*, 205–267. London: SAGE.

Morgenstierne, Georg. 1926. *Report on a Linguistic Mission to Afghanistan* (Serie C I - 2). Oslo: Instituttet for sammenlignende kulturforskning.

Morgenstierne, Georg. 1961. Dardic and Kafir Languages. *Encyclopedia of Islam*, vol. 2, Fasc. 25, 138–139. New Edition. Leiden: E.J. Brill.

Payne, John R. 1989. Pamir languages. In Rüdiger Schmitt (ed.), *Compendium linguarum Iranicarum*, 417–444. Wiesbaden: Dr. Ludwig Reichert Verlag.

Reichl, K. 1983. Syntactic Interference in Afghan Uzbek. *Anthropos* 78(3/4). 481–500.

Sharma, J. C. 2002. Gojri and its relationship with Rajasthani. *International Journal of Dravidian Linguistics* 31. 27–46.

Skalmowski, Wojciech. 1985. The linguistic importance of the Dardic languages. *Journal of Central Asia* 8(1). 5–15.

Skjaervø, Prods Oktor. 1989. Modern East Iranian languages. In Rüdiger Schmitt (ed.), *Compendium linguarum Iranicarum*, 370–383. Wiesbaden: Dr. Ludwig Reichert Verlag.

Strand, Richard F. 1973. Notes on the Nuristani and Dardic Languages. *Journal of the American Oriental Society* 93(3). 297–305. doi:10.2307/599462.

Strand, Richard F. 2001. The tongues of Peristân. Appendix 1. In Alberto Cacopardo & Augusto Cacopardo (eds.), *Gates of Peristan: History, Religion and Society in the Hindu Kush* (Reports and Memoirs 5), 251–257. Rome: Istituto Italiano per l’Africa e l’Oriente (IsIAO).

Tikkanen, Bertil. 1988. On Burushaski and other ancient substrata in northwestern South Asia. *Studia Orientalia* 64. 3030–325.

Tikkanen, Bertil. 1999. Archaeological-linguistic correlations in the formation of retroflex typologies and correlating areal features in South Asia. In Roger Blench & Matthew Spriggs (eds.), *Archaeology and language IV: Language change and cultural transformation*, 138–148. London: Routledge.

Tikkanen, Bertil. 2008. Some areal 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. Karachi: Oxford University Press.

Toporov, Vladimir Nikolaevǐc. 1971. *Burushaski and Yeniseian languages: some parallels*. Academia.

Toporov, Vladimir Nikolayevich. 1970. About the phonological typology of Burushaski. In Roman Jakobson & Shigeo Kawamoto (eds.), *Studies in General and Oriental Linguistics Presented to Shiro Hattori on the occasion of his sixtieth birthday*, 632–647. Tokyo: TEC Corporation for Language and Educational Research.

Weinreich, Matthias. 2011. Domaaki jazyk. In Tatiana I Oranskaia, Yulia V Mazurova, Andrej A Kibrik, Leonid I Kulikov & Alexandr Y Rusakov (eds.), *Jazyki mira: Novye indoariiskie jazyki*, 165–194. Moscow: Academia.

Zeisler, Bettina. 2005. On the position of Ladakhi and Balti. In John Bray (ed.), *Ladakhi histories: Local and regional perspectives* (Brill’s Tibetan Studies Library v. 9), 41–64. Leiden ; Boston: Brill.

Zoller, Claus Peter. 2005. *A Grammar and Dictionary of Indus Kohistani: Volume 1, Dictionary* (Trends in Linguistics 21–1). Berlin: Mouton de Gruyter.