



DARDESTĀN II. LANGUAGES

ii. LANGUAGES

The Dardic (< OInd. *darád-* “the people who live next to Kashmir”; cf. *dārada-*, *darada-*, designating the population of northern India, and modern *dard*, *dārd*, self-denomination of the speakers of Gurezi, one of the Shina dialects) languages are a group of Indo-European languages spoken in part of Nūrestān and adjacent areas along the Kabul river and its tributaries in the mountain region that encompasses northeastern Afghanistan, northern Pakistan, and northwestern India (see i, above).

In the literature this group is sometimes also called Piśāca (the old term for the population of northwestern India), and Paiśāci (obsolete name of one Middle Indian language of this region).

Classification.

The main languages and dialects of the Dardic group are the following.

Eastern subgroup. The eastern subgroup includes Kashmiri in the Kashmir valley; Shina in the districts of Gilgit and Tangir, north of Kashmir; Phalura (or Palola) and the closely related Sawi; and a number of languages and dialects sometimes referred to generally as Kōhestānī (lit., “of the mountains”) in the Indus, Swat, and Panjkora basins: Maiyan (so called by native speakers but Kōhestānī by others) with the Kanywali dialect, Torwali, and Bashkarik (or Diri, known in another dialect variant as Garwi).



Central subgroup. The central subgroup is further subdivided into northern and southern groupings. The northern grouping includes Khowar (or Chitrāli, Chitrārī, Chatrorī, Arniya) and Kalasha in the [Chitral](#) region. The southern grouping includes Tirahi, Gawar (or Gawar-bati, lit., “language of the Gawar people”), Katarkalai (or Wotapuri, referring to another dialect), Shumashti, Glangali (closely related Ningalami, reported in the literature but apparently no longer extant), and Pashai, a large group of extremely divergent dialects or closely related languages, in the southern part of Nūrestān and adjacent areas.

Nūrestānī languages. The Nūrestānī languages (also known traditionally as Kafiri languages) are sometimes included as a western subgroup of the Dardic group of languages (see Shaw, 1876, pp. 146-47; Grierson; Morgenstierne, 1945; idem, 1974). They are the languages of Afghan Nūrestān, known as Kafiristan until the people adopted Islam on the eve of the 20th century. This subgroup includes Kati (including the eastern dialect Bashgali), Waigali (or Wai, Wai-alā), and related Tregami (or Gambiri) and Zemiaki; Ashkun and the closely related language or dialect Wamai; and Prasun (or Paruni, Wasin-veri, Veron). These languages have much in common with the Dardic languages and are spoken in close geographical proximity to them, but their origin is not the same (see below). The attribution of the Dameli language, which exhibits both Nūrestānī and Dardic features, is not clear. Some authors also include in the Dardic group the Ḍumaki language, spoken by a people scattered in groups in Hunza and Nagar. Genetically, however, it belongs to the Central Indo-Aryan languages (being close to Gypsy), rather than to the Dardic group.

Linguistic classification.

The Dardic languages are an offshoot of the Indo-Aryan languages of the post-Vedic period. The Nūrestānī languages belong neither to the Iranian nor to the Indo-Aryan group but represent instead an independent branch of the Aryan family of Indo-European languages. This relationship is represented schematically in [Table 1](#).

The main classifying features within this family of languages is the different phonetic evolution of certain Proto-Aryan consonants.

Development of the Proto-Aryan stops. The Proto-Aryan phonological system included four kinds of stop: voiced and unvoiced, aspirated and unaspirated (*dh, *d, *th, t; *bh, *b, etc.). In Proto-Indo-Aryan these stops remained separate phonemes, whereas in the Nūrestānī languages the aspiration was lost



(**dh* and **d* > **d*, **th* and **t* > **t*, etc.), leaving a system with two kinds of stop: voiced and unvoiced (*d*–*t*, etc.). In Proto-Iranian only the voiced aspirate stops lost their aspiration (**dh* and **d* > **d*, etc.), while the unvoiced aspirated stops, with a few exceptions, became fricatives (**th* > **θ*, **ph* > **f*, **kh* > **x*); the system of four stops was thus changed into a system of two stops, voiced and unvoiced, and a fricative (**d* – **t* – **θ*). Examples include Proto-Indo-Aryan (and Old Indian) **khara* “donkey,” Dardic (Kalasha) *khār*, Nūrestānī (Kati) *kur*, Proto-Iranian (and Avestan) **xara* (Pers. *kar*).

Development of the Indo-European palatal velars. Indo-European **ǵ*(*h*), **ǵ* became Proto-Aryan palatal affricates **j*′(*h*), **ć* (i.e., *dź*(*h*), *tś*). They remained in Proto-Nūrestānī and Proto-Iranian as separate phonemes but merged with the palatalized velars in Proto-Indo-Aryan (see below).

Development of the Proto-Aryan palatalized velars and palatal affricates. In Proto-Aryan the Indo-European velars were palatalized before front vowels: **g*(*h*), **k* before *e* and *i* > **j*′(*h*), **č*. When Proto-Aryan **e* changed into **a*, these palatalized velars became separate phonemes, contrasting with the Proto-Aryan descendants of the Indo-European voiced and unvoiced palatals: **j*′(*h*), **ć*. As with the stops, the aspiration remained in Proto-Indo-Aryan but was lost in Nūrestānī and Proto-Iranian, leaving only **ǰ*, **č* and **j*′, *ć*, which remained as separate phonemes. In Proto-Indo-Aryan **j*′*h*, **j*′*h* and **ǰ*, **j*′ then merged into single phonemes *h* and *ǰ*, *č* remained, and **ć* changed to palatal *ś*. In Nūrestānī **ǰ* remained, **j*′ changed to **ž* (> *z*), and **ć* became the dental affricate *c* (*ts*). In the Iranian languages **ǰ* and **č* remained, while **j*′ became *z* or *d* and **ć* became *s* or *θ*. Examples include Indo-European **dek̑m* “ten” (Latin *decem*, etc.) > (late) Proto-Aryan **daća* > Nūrestānī (Kati) *duc*, Avestan *dasa*, Old Persian **daθa* (Pers. *dah*), Old Indian *daśa*, Dardic (Dameli) *daš*; Indo-European **ǵenu*–/**ǵonu*– “knee” > Proto-Aryan **j*′*ānu*– “knee” > Nūrestānī (Kati) *jō* (i.e., *dzō*), Old Indian *jānu*–, Avestan *zānu*– (Pers. *zānū*); and Proto-Aryan **aj*′*ham* “I” (cf. Latin *egō*, etc.) > Nūrestānī (Kati) *vūze*, Avestan *azəm*, Old Persian *adam*, Old Indian *aham*, Dardic (Pashai) *ā*.

Development of Indo-European s. Indo-European *s* remained in Proto-Aryan except in a few positions. Proto-Aryan *s* remained in Nūrestānī and Proto-Indo-Aryan but became *h* in Proto-Iranian. Indo-European **s* became **š* after Indo-European **i*, **ə*, **u*, **r*, **k*, **ǵ* in most of Proto-Aryan, but not consistently in Nūrestānī, where in a few words the *s* has remained after *u*, for example, in Kati *mūsə* “mouse,” as opposed to Old Indian *mūṣ-*, Iranian (Persian) *mūš* (for



more details, see Morgenstierne, 1945, pp. 225-36; idem, 1973, pp. 227-347; idem, 1974, pp. 6-8; Buddruss).

The Nūrestānī and Dardic languages are therefore genetically quite distinct, and the similarities between them that can be observed today and that have obscured their origins must be ascribed to various causes: their common Proto-Aryan origin, effects of a common substrate, parallel developments resulting from their geographic proximity, converging phonetic developments, and mutual influences and borrowings. As a result of all these trends, Dameli, for instance, now occupies what may be termed an “intermediate” position between the two groups. The separate nature of the Dardic languages is still clear, however, from their close relationship with other Indo-Aryan languages, especially Punjabi, Lahnda, Sindhi, and the dialects of Western Pahari.

The assumption of a connection between the Dardic languages and the Middle Indian Paiśācī language is based partly on their common geographical location and partly on such phonetic features as the devoicing of postvocalic (occasionally also initial) voiced consonants. This feature is not shared by all Dardic languages, however, and may be a regional, rather than a genetic, phenomenon. A genetic relationship between the two groups is therefore doubtful.

The Dardic languages.

The Dardic language group can be further subdivided according to genetic and structural criteria.

1. The four-way opposition among the stops (*dh-d-th-t*, etc.) remains in most of the Kōhestānī languages but is reduced through the merger of the voiced aspirated and unaspirated stops (*dh* and *d > d*, etc.) to a three-way opposition elsewhere, except perhaps in Kalasha, Phalura, and some of the dialects of Kashmiri. In some of these languages the loss of the voiced aspirated stops gave rise to a tonal opposition, similar to that found in some neighboring Indo-Aryan languages.

2. In Proto-Khowar-Kalasha intervocalic and final **t* become **δ*, which then became *l* in Northern Kalasha but *r* in Southern Kalasha and Khowar. (There is a similar distribution of the change from **t* in the Gypsy dialects, to *r* in Syria, to *l* in Europe.)

3. The augment was retained only in Khowar and Kalasha (e.g., *akāren* “they



did”).

It is possible to distinguish still further subdivisions, for instance, Phalura-Sawi-Shina, and Shumasti-Glangali-Ningalami, but the mutual relationships among the Dardic languages are still unclear at many points, primarily because of insufficient research, so that, for instance, the phonology of several Dardic languages is very scantily known. Furthermore, determination of which similarities should be ascribed to common ancestry and which have been caused by linguistic interaction still remains to be sorted out

History of study of the Dardic languages.

The only Dardic language with a long literary tradition is Kashmiri, which was formerly written in the Śārada and Nagari alphabets and more recently in the Urdu modification of the Arabic script. Written versions of Khowar, Shina, and Maiyan, also based on the Arabic alphabet, have appeared only recently.

Although Dardic languages were occasionally mentioned in travelers' notes and Christian missionaries actually published parts of the Bible in Kashmiri (in the Nagari alphabet) in the 1820s, the study of Dardic languages actually began in the 1830s, when the first professional linguistic data, in the form of compact vocabularies and glossaries, were published by M. P. Edgeworth, R. Leech, G. T. Vigne, and others. Between the 1860s and 1880s more comprehensive information appeared, along with studies of the folklore and ethnography of the region. The most important works of this period were those by J. Campbell, G. W. Leitner, J. H. Knowles, and K. F. Burkhard. At about the same time such scholars as R. B. Shaw and Wilhelm Tomaschek incorporated material on the Dardic and Nūrestānī languages into their studies of comparative Indo-Iranian linguistics.

By the 1890s the study of the Dardic languages had assumed a systematic character. In 1896 G. A. Grierson initiated and supervised the multivolume *Linguistic Survey of India*, an attempt to gather data on all known languages of India and neighboring countries in conformity with general guidelines and a standard questionnaire. The *Survey* included linguistic outlines and lexical and textual material for the Dardic languages and a consolidated dictionary of Dardic vocabulary. Grierson also wrote several works on various Dardic languages, notably a grammar and an extensive dictionary of Kashmiri. A major contribution to the description of Shina, Kashmiri, and other languages was made by T. Grahame Bailey, beginning in



the first decade of the twentieth century. Georg Morgenstierne devised the accepted genetic classification of the Dardic and Nūrestānī languages. In the 1920s he began to publish descriptions of many Dardic, Nūrestānī, and Iranian languages, as well as texts, vocabularies, and historical essays. In the 1950s Georg Buddruss began publication of texts, vocabularies, and descriptions of languages, supported by profound historical analysis. He drew on a wide range of material to revise the historical-genetic classification of the Nūrestānī languages, thus largely completing the classification of the Aryan language group.

In recent decades research has been focused on the contemporary state of languages (e.g., in the works of B. B. Kachru, B. A. Zakharyin, and A. L. Gryunberg) and on typological and regional analysis of the Dardic languages (e.g., by W. N. Toporov, Gérard Fussman, and others). The material of Dardic languages has also been used for historical analysis of the Indo-Aryan and other languages by R. L. Turner, Manfred Mayrhofer, I. M. Steblin-Kamenskii, and others.

Exchange of loanwords with Iranian languages.

Contact between the Dardic and Iranian languages and the resulting mutual exchange of loanwords varied with several factors. In folklore, particularly poetry, and religious texts there are a large number of Persian words and elements, as well as some from Arabic borrowed through Persian. Such borrowings have often been indirect, through Urdu in the modern Dardic languages of India and Pakistan and through Darī or Pashto (see [AFGHANISTAN v, vi](#)) in those of Afghanistan, especially in some of the Pashai dialects. In Kashmiri, the traditional Dardic literary language, they came directly from literary Persian.

Persian and Arab-Persian loanwords fall into large lexical groups. One such group is the concrete nouns (names of animate and inanimate objects and substances): Kashmiri *pādšāh*, Khowar *bāčha*, Kalasha *bādša*, Kanyawali *bādšā* “padishah”; Kashmiri *wəz’ir* “vizier,” *yīnsān* “human being,” *hamdard* “courtier,” *k’itāb* “book”; Shina *bādām* “almond”; Kalasha *biaban* “desert”; Kashmiri *šah(i)r* “city,” *kākaz* “paper.” Another is names of abstract nouns: Kashmiri *wādi*, Khowar *wada* “promise”; Kashmiri *mōjizi* “wonders,” *yād* “memory,” *nūr* “light”; Kalasha *bo* “smell,” *khiyal* “thought,” *kuwat* “strength”; Khowar *wafā’*, Kashmiri *waphā* “fidelity”; Khowar *itifaq* “unity,” *žān* “life, soul.” Adjectives



constitute another group: Kashmiri *phān'ī* “temporal, perishable,” *mušk'ul* “difficult,” *waphādār* “devoted,” *g°anāgār* “sinful”; Kalasha *bedarkar* “ill, sick”; Khowar *kam* “little, less, few.”

Some loanwords occur as the nominal components of compound verbs, as in Persian and other Iranian languages: Khowar *jawab d-* “to answer,” *wada k-* “to promise”; Kashmiri *taft'iš kar-* “to check,” *kar- arz* “to complain,” *kar-khošt* “to make merry, to rejoice.” Much less often a simple verb has been borrowed, for example, Khowar *neweš-* “to write.” Occasionally borrowed conjunctions, particles, and other auxiliaries are found: Kashmiri *agar*, Shina *āgār* “if”; Kalasha *albat* “probably”; Kashmiri *hargāh k'ī* “if only”; Kashmiri *magar*, Khowar *magam* “however.”

In regions of direct contact between various Dardic languages and the adjacent Iranian languages, that is, in the Hindu Kush and part of the Pamirs, there are mutual lexical borrowings between Iranian and Dardic languages, as well as common borrowings from other sources. The Dardic languages have borrowed from live and extinct Iranian languages throughout the entire area of contact. In Khowar several Iranian sources from different periods can be distinguished (Mor-genstierne, 1973, pp. 241-55).

In turn, the Iranian languages of this area have borrowed a vast vocabulary from Dardic languages or through Dardic languages from other Indo-Aryan sources. These words are mostly related to local life, traditional household items, plants, and the like, for example, terms for birch bark: Shughni *birūj*; Rushani, Bartangi, Roshorvi *birū/ūj*; Yazghulami *bəruž, bəruj*; Ishkashmi *biriz', bəriž, bruž*; Sanglechi *bərež, bəriž*; Zebaki *bruž, bərež*; Tajik dialect *bəruj, buruj*; Pashto *barj*; Khotanese Saka *bruṃja-* “bark.” Of special interest is the term for cotton: Shughni *čipōs* (in the collocation *čipōs rūyan* “cotton oil”), Rushani *čipūās*, Yazghulami *k'əbes*, Wakhi *kəbas*. These terms are all descended from **kap(p)āsa-* (cf. Pali *kappāsa-* < OInd. *karpāsa-*, Khowar *kəbos*, Marathi *kāpūs*, Nepali and Hindi *kapās*, as the word spread eastward following the cultural reality: Burushaski *gupas*, Werchikwar *γupas*, Uighur *kāpāz, kepāz*, and so on; Turner, no. 2877). Despite the obscure origin of this word in Old Indian (see Mayrhofer, *Dictionary* fasc. 3, pp. 174-75), it certainly came into these Iranian languages from Indo-Aryan, presumably Dardic, languages at a relatively early period, as can be deduced from phonetic transformations: the transition **k > č* before **a* in Shughni and Rushani, also typical of indigenous words, and the transition **-p- > b* in Yazghulami. Wakhi *kəbas* could have been borrowed directly from Khowar or



from Indo-Aryan before the change of *p* to *b* in Wakhi. The term for a particular kind of cotton cloth, which continues the same Old Indian prototype *karpāsa-* and has been borrowed by other Iranian languages through other channels, exhibits a more archaic form in Iranian than the one observed in the Dardic languages, for instance, Persian *karbās*, Tajik *karbos*, but Kalasha *kravas*, Khowar *karvas*.

In the Iranian languages of the Pamirs the term for the local style of kerchief is probably a similar borrowing with relatively early phonetic transformations: Shughni-Bajuwi *cēl*, Rushani-Khufi, Roshorvi *cīl*, Yazghulami *cil* (possibly from Rushani), Wakhi *čil*. Among Dardic equivalents are Kalasha *čēl-*, Dameli *čel*, Pashai *čilā-* (Turner, no. 4910). Another Iranian borrowing is Shughni *lāq* “old, worn trousers, tatters,” *lēq* “old, torn quilt”; Rushani *loq*, Bartangi *löq* “clothes”; Rushani, Khufi, Bartangi *lēq* “old clothes, rags”; Yazghulami *luq* “clothes, fabric, rags” (pl. *laqáθ* “quilts”); Wakhi *luq* “rags.” This word is probably related to Indo-Aryan **lakka-* “defective” (Turner, no. 10877) or possibly to Ossetic *lyg/lux*, the onomatopoeic origin of which cannot be excluded.

A Dardic source may also be attributed to Shughni *kappur*, Rushani *kapor*, Bartangi *kapör* “gourd, calabash”; Roshorvi *kapir* “gourd snuffbox”; Yazghulami *kapur*, *kapuâr* “elongated calabash”; Tajik dialect *kapar* “calabash,” all probably related to Old Indian *karpāra* “bowl.” Other plant names display similar patterns of borrowing. In the Iranian languages of the Pamirs a large number of terms for useful domestic plants were borrowed from the Dardic languages (and through them from other Indo-Aryan sources): Shughni, Rushani, Khufi, Bartangi, Roshorvi *pīnĵ*, Sariqoli *penĵ* (and Wanji *punĵev*) “millet”; Wakhi *šax*, Shughni (etc.) *χāš* “peas, beans”; Wakhi *kroš*, Ishkashmi *karoš*, *karāš* “cinna”; and so on (for more details, see Steblin-Kamenskii, pp. 24, 30, 33, 42, 46, 49 ff.).

Certain Dardic borrowings in Iranian languages can be linked with traditional taboos on words, which were usually replaced by either descriptive expressions or loanwords. For example, in certain Iranian languages of the area the original terms for the wolf that continue proto-Iranian **urka-* (cf. Av. *vahrka-*, OInd. *vṛka-*, Pers. *gorg*, Yazghulami *warǵ*, etc.) have been replaced by *χīθp* in the Barwozi dialect of Shughni and *χīθp* in Sariqoli, both borrowed from Wakhi; in Wakhi itself, however, *šapt* “wolf” is borrowed from **šapita-* in the precursor of Khowar (continued in modern Khowar as *šapīr*), which was a continuation of Old Indian **šapita-* “cursed,” from the verbal stem **šap-*



“to curse” (Mayrhofer, *Dictionary*, fasc. 21, p. 396; Turner, no. 12293). The term for wolf is also taboo in other languages of the area, but the substitution is made in a different way; cf. Ormuri *lēwū*, Pashto *lewə* “wolf” < **daiyā*- “of the dēvs” < **daiya*- “dēv, evil spirit.” In the non-Indo-European Burushaski language *ūr*k and *ur*k in the Werchikwar dialect are borrowed from Iranian Ishkashmi.

The Iranian, Dardic, and other languages of the area also have a common “local vocabulary,” which includes words, usually names of local things, household terms, and the like, the source and transmission of which are unknown or in doubt. These words include, for instance, Shughni *tāk* “gown string,” Rushani, Khufi *tāk* “collar, lapel,” Sariqoli *tok* “button,” Yazghulami *tak*^o “string at the collar of a woman’s gown,” Wakhi *tak* “button”; cf. Shina *ṭak*, Khowar *ṭak* “button, string of a woman’s gown,” Burushaski *ṭak*, Werchikwar *ṭak* “collar string, button.” Evidently this term initially denoted the common reality in the area: the string fastening the collar of the old local dress (cf. Tajik *yak-tak* “old-fashioned shirt with the string on one side”; on the other hand, Kurdish *tog* “string” is derived from Armenian). When the button appeared in this area the old term for “fastening” was retained in some languages and in others a new word was borrowed from Turkic: Shughni *tukmā*, Yazghulami *təkmá* “button.” Another example of the common regional vocabulary is the term for a large flat cake: Shughni *ḫipik* “pancake” and Sariqoli *ḫ(i)pik* “big flat cake” are borrowed from Wakhi, in which *ḫapik* “kind of cake” is itself derived from Khowar *šapik* “cake” (cf. Burushaski *šapik*).

In addition, the Dardic and Iranian languages share a number of structural features, also ascribable to a variety of causes, including substrate influence and interaction. In noun morphology the use of nouns like *gal(a)*, originally meaning “flock” (also found throughout Iran; e.g., Pashai *-kuli*, *-ēlā*, Gawar *-gila*, Kashmiri masc. *kyəl*, fem. *kyaj*; cf. Shughni *-galā*, *-xēl*, etc.) as plural markers may be noted. In pronominal inflection the remarkable convergence in the forms of the first-person plural pronouns Khowar *ispa* (direct and oblique cases), Wakhi *spo* (genitive) and the influence of the second-person singular pronoun on the second-person plural (e.g., Phalura sing. *tu*, pl. *tus*), is noteworthy; it may be compared with the “prefix” *ta-* used in the second-person plural in the Shughni group and Ishkashmi (Schmitt, p. 432). In both the Iranian languages of the area and Dardic (except Kashmiri) the numeral system is vigesimal; for example, fifty-five is expressed as “two score + ten +



five” or “two score + fifteen.” The numerals from 11 to 19 are expressed as “ten-one” and so on, rather than in the earlier form “one-ten.” The syntax is characterized by the word order subject-object-verb in most of the languages, avoidance of indirect speech, and the like (see Edel’man, 1968, pp. 77-98 and maps 2, 5-14).

BIBLIOGRAPHY

G. Buddruss, “Nochmals zur Stellung der Nuristan-Sprachen des Afghanischen Hindukush,” in *MSS* 36, 1977, pp. 19-38.

D. I. Edel’man, *Dardskie yazyki* (The Dardic languages), Moscow, 1965.

Idem, *Osnovnye voprosy ling-visticheskoï geografii (na materiale indoiranskikh yazykov)* (Fundamental problems of linguistic geography [based on the materials of the Indo-Iranian languages]), Moscow, 1968.

Idem, *The Dardic and Nuristani Languages*, Moscow, 1983 (with extensive bibliography). G. Fussman, *Atlas linguistique des parlers dardes et kafirs*, 2 vols., Paris, 1972.

G. A. Grierson, *Linguistic Survey of India* I/1-2, Calcutta, 1927-28; VIII/2, Calcutta, 1919.

G. Morgenstierne, *Indo-Iranian Frontier Languages*, 4 vols. Oslo, 1929-67.

Idem, “Indo-European k’ in Kafiri,” *NTS* 13, 1945, pp. 225-38.

Idem, *Irano-Dardica*, Wiesbaden, 1973.

Idem, “Languages of Nuristan and Surrounding Regions” in K. Jettmar, ed., *Cultures of the Hindukush*, Wiesbaden, 1974, pp. 1-10.

I. M. Oranskij, “Indo-Iranica IV. Tadjik (Régional) *Buruĵ* ‘Bouleau,’” in *Mélanges linguistiques offerts à Émile Benveniste*, Paris, 1975, pp. 435-40.

R. Schmitt, ed. *Compendium Linguarum Iranicarum*, Wiesbaden, 1989.



R. B. Shaw, "On the Ghalchah Languages (Wakhí and Sari-ḳolí)," *Journal and Proceedings of the Asiatic Society of Bengal* 45, 1876, pp. 139-278.

I. M. Steblin-Kamenskii, *Ocherki po istorii leksiki pamirskykh yazykov. Nazvaniya kul'turnykh rasteniï* (Outlines for the lexical history of the Pamir languages. Names for cultivated plants), Moscow, 1982.

R. L. Turner, *A Comparative Dictionary of the Indo-Aryan Languages*, 2 vols., London, 1963-66.

Table 1. Origin of the Dardic languages and their relation to other Aryan languages