



NURESTÂNI LANGUAGES

NURESTÂNI LANGUAGES, five languages constituting the Nurestâni (Pers. “Nurestâni,” Engl. “Nuristani”) subgroup of the Indo-Iranian language family. The approximately 130,000 speakers of these languages inhabit Nurestân Province in northeastern Afghanistan and a few adjacent valleys in Pakistan’s Chitral District. This region lies on the southern slopes of the Hindu Kush mountains and encompasses areas drained by the Alingar River in the west, the Pech River in the center, and the Lanḡay Sin and Konar-Chitral Rivers in the east.

The Nurestâni languages have no tradition of writing. Words cited here are transcribed phonemically, following the conventions given below under “Phonology.”

The five languages and their dialects are listed here in north-to-south order. Native names are given first, followed by other names (in quotation marks) used by non-native speakers. Maps showing the locations of the languages appear on this author’s website (Strand 1997-2010), along with further linguistic details.

1. *Vâs’i-vari* (“Prasun,” “Pâruni”; abbreviated “V.”) is spoken in three dialects by the *Vâs’i* people, who inhabit the Pârun Valley in the uppermost Pech Basin. The Upper dialect is spoken in the uppermost village, *Şup’u*, the Lower dialect is spoken in the lowermost village, *Uş’üt*, and the Central dialect is spoken in the valley’s four middle villages. (Buddruss 2004; Morgenstierne 1949; Strand 1973, 1997-2010.)



2. *Kâmk'ata-Mumkšt'a-vari* (“Kati,” “Bashgali”; abbr. “K.”) is spoken by the *Kât'a*, *Mum'o*, *Kšt'o*, and *K'om* peoples, who range across the breadth of Nurestân. The more numerous *Kâta* speak two main dialects, Western and Eastern *Kât'a-vari*. Western *Kâta-vari* is spoken in the upper *Řâmg'al* (abbr. “K.ř.”) portion of the Alingar Basin, in the *Kt'ivi* (abbr. “K.ktv.”) Valley off the upper Pech, and in the *P'eřuk* Valley off the upper Lanđay Sin. Each of these regions has its own subdialect. Eastern *Kâta-vari* (“K.e.”) is spoken in the upper Lanđay Sin and in a few villages across the border in Pakistan’s Chitral District. The Kom and the Kšto speak another dialect, *Kâmv'iri* (“Kamdeshi”; abbr. “K.km.”) or *Kšt'aviri* (“Kushtozī”), depending on which ethnic group one addresses; the Kom inhabit the lower Lanđay Sin and upper Konar basins, including a small portion of southern Chitral. Three territorially incorporated peoples, the *Jâš'a*, the *Bini'o*, and the *Jâmc'o*, live among the Kom and speak *Kâmviri*. The Kšto inhabit one village in the Ničangal Valley, *Kštořm*, surrounded by Kom, and another village, *D'u~gul*, where they speak *Kalaša-alâ*. The Mumo inhabit a tract of the Lanđay Sin between the Kom and Eastern *Kâta*. They speak a transitional dialect, *Mumv'ari*, which shows a phonological overlay of *Kâmviri* on a *Kâta-vari* base. (Ghulâmullâh 1966, Grjunberg 1980; Strand 1997-2010.)

3. *Âšku~vi:ri* (“Ashkun”; abbr. “A.”) is spoken by the *Âšku~* people in separate dialects in the *Bâzâigal*, *Mâse:gal* (*Kolâtâ*, abbr. “A.m.”), and *Titin* Valleys off the middle Alingar, and by the *Saņu* (abbr. “A.s.”) and *Grâmsaņâ* peoples of the middle Pech Valley, each with their own dialect. (Buddruss 2006; Morgenstierne 1929, 1934, 1952; Strand 1997-2010.)

4. *Kalaša-alâ* (“Waigali”; abbr. “Kal.”) is spoken by the *Kalaša* people of the Vaigal Valley off the lower Pech and to the east in the *Veligal*, *Ćüki*, and *D'u~gul* Valleys off the middle Konar. A major dialect division separates the *Varjan* people of the upper Vaigal Valley from the *Ćima-Nišeī* people of the lower valley. Minor dialect divisions separate each village, especially among the *Varjan*, where the chief subdialect, *Vâ-alâ* (abbr. Kal.v.) is that of the *Vâi* (also called *Vâ* or *Vai*) people of the village of *Vâigal*. The *Ćima-Nišeī* dialects, each centered on one of four villages in the lower Vaigal basin, show influence from *Âšku~* dialects to the west. Chief among them is *Nišeī-alâ* (abbr. “Kal.n.”), spoken in *Nišeigrâm*. In addition, *Kalaša-alâ* is spoken by two other ethnic groups: the *Vântâ* in a hamlet in the lower Vaigal Valley and the Kšto inhabitants of *d'u~gul* in the *Dungal* (Daren) Valley. A *Kalaša-ala*-speaking colony is also reported to live in the village of *Zamiâ* in the *Āpadara* Valley, a



southern tributary of the lower Pech River. (Degener 1998; Morgenstierne 1954; Strand 1997-2010, 1999.)

5. *Tregâmi* (abbr. “Tr.”, native name unrecorded) is spoken in the three villages of Gambir, Kaţâr, and Devoz in the Tregâm Valley off the lower Pech. This little-known language appears to be a close relative of Kalaşa-alâ, with characteristics borrowed from neighboring Indo-Aryan languages and some Kâta-vari influence. (Morgenstierne 1952.)

On the basis of shared phonological developments, V. and K. form a Northern Nurestâni cluster, and A., Kal., and Tr. form a Southern Nurestâni cluster; however, V. has been isolated from phonological and grammatical developments common to the “core” group of A., K., and Kal., and the grammar of Tr. has been influenced by nearby Indo-Aryan languages.

Phonology. Common to all Nurestâni languages are the phonemic contrasts between the vowels *a*, *â*, *u*, *o*, *i*, and *e*. To these are added the fronted vowels *ü* in Kal., K.km., and V. (diphthongized to *iu* in K.ktv), *ö* in Kal. and V., and *ã* in Kal. *a* represents the unmarked vowel in each language, etymologically descendant from ancient **a*. In A., K., and V. it is phonetically a high central vowel [i], while in Kal. and Tr. (probably) it is a low central vowel. In K. it is palatalized after palatal consonants; thus *č’a* ‘how many,’ pronounced [či]. *a* contrasts with *â*, which is articulated with the jaw more open. *o* is close in A., open in K. In K. and V. vowels are articulated before they are vocalized, so that initial *u*, *ü*, and *i* have homorganic on-glides, non-phonemic in K.km., but promoted to phonemic status in other dialects through later vowel changes (cf. K.km. *ušpa* with initial [ʷ] but K.ktv. *vaš’up* ‘horse’). Nasalized vowels are indicated by a following *~*. In some dialects vowel length is distinctive; it is represented by postvocalic *:* in A., V. and Kal.v., and by postvocalic *a* in K.km. Accent is indicated by ‘ before a vowel in K. and V., where it is phonemic. In the remaining groups (including Tr.?) accent apparently falls automatically on the last syllable of a word’s stem.

Among the consonants there are contrasts between labial, dental (apico-dental), “palatal” (lamino-alveolar), retroflex (apico-alveolar), and velar sets; within the sets are contrasts between stops (oral and nasalized), affricates, spirants, flaps, laterals, and approximants. Contrasts of voiced vs. voiceless stops, affricates, and spirants appear in all dialects, but voicing is positionally determined for some consonants in K.km. Among the labials *v* is a spirant, labio-dental in K.km. but bilabial in K.ktv.; in the other languages it is labio-



dental before front vowels, bilabial ([w] or [β]) elsewhere. *ć* and *ź* are dental affricates. The latter contrasts with the voiced spirant *z* in A., minimally with *z* (from loanwords) in K.km. and perhaps in V.; in the remaining dialects an earlier **ź* has become *z*. Palatal consonants include the affricates *č* and *ǰ* and the spirant *š*. In A., V., and Tr., the voiced palatal spirant *ž* is phonemic; in K.km. it is an allophonic alternant of *š*. The status of the palatal approximant *y* as a phoneme distinct from *i* is doubtful. Retroflex consonants are indicated by underposed dots: *ṭ*, *ḍ*, *ḥ*, *ḷ*, *ṣ*, *ḷ*, *ṛ*, and *ṇ*. *r* is an upward flap, contrasting with the forward flap *ɾ*, which is phonemic in A., Kal., and Tr. (?), but allophonic of *ṭ* in K.km. A retroflex approximant *ɻ*, similar to English prevocalic *r* without lip-rounding, is found in all dialects except V.; but in A.m. *ɻ* has become *l*. *ɻ* is phonemic in K., but probably allophonic of *r* in A. and *ɾ* in Kal. In A. and Kal. there is a strongly nasalized retroflex flap, *ṛ̃*, which is an intervocalic allophone of *ṛ* in Kal. and may be so in A., but it may be minimally contrastive with *ṛ* in A. in word-final position (cf. A.s. *voṅ* ‘irrigation weir,’ the lone example, vs. *kāṇ* ‘arrow’). The etymologically corresponding sound in K. is a nasalized retroflex approximant, also symbolized as *ṛ̃*, which phonemically contrasts with *ṛ*. The possibility of a retroflex *ḷ* in A. and V. needs further investigation. In K.km. there is a voiceless retroflex lateral affricate [ɭ], which is phonemically a cluster of the consonants *ṭl*. Among the velars, K.km. *k* has an intervocalic flapped allophone [ḳ]; the allophonic status of a similar voiced spirant in V. is unclear. The velar nasal *ŋ* is phonemic in all languages. Syllable juncture, indicated by *ˑ*, is phonemic in at least K. and Kal. An underscore (*_*) between words indicates close transition, which produces various articulatory assimilations, while = between words indicates fusion with elision of a phoneme.

Grammar. Major grammatical traits common to the Nurestāni languages are described in the following paragraphs. Examples are from Kām viri unless otherwise noted. Vāsi-vari and Tregāmi depart somewhat from the grammatical pattern of the other languages, notably in their lack of a “split-ergative” pattern of verbal subject reference (see below).

Syntax. Typical sentences in the Nurestāni languages first depict the items of discourse and their spatial interrelationships, and then they depict a change in the items; that is, verbs normally come at the end of a sentence, as in most Indo-Iranian languages. Attributes precede the nouns they qualify. Attributive subordinate clauses are rare, their function usually being rendered by participial phrases preceding the nouns they qualify; e.g., *āska t’ua d’us*



vāñi mānša ‘*oa-sa.*, lit. “that thee-by yesterday seen man come-one-is” ‘The man you saw yesterday has arrived,’ less commonly *aska mānš’a k’âa d’us t’ua vāñ’i-sa, âsk’a oa-sa.*, lit. “that man what yesterday thee-by seen-one-is, that come-one-is.” Subordinate clauses otherwise function in the following ways. With postverbal particles, they depict past or hypothetical scenes, as in ‘*o~ć g’um_to, t’ü â-k’i n’â_âsaš* ‘When I went, you weren’t there’ or ‘*o~ć g’um bo, t’ü di ieloš* ‘If I go, you should go, too.’ With the postposed absolutive (*kti*) of the verb ‘make,’ in the sense of ‘said,’ they depict quotations (always direct) and causes, as in ‘*i~a i’e~ sta_âsa kti giĵa_kârō*. ‘He said that he has to go,’ or *gâtr’a bi-sam_kti n’â go*. ‘He didn’t go because he was tired.’ Experiential sentences depict the internal states or compulsions of a backgrounded-case experiencer, as in ‘*i~a ‘oata bo*. ‘I’m hungry’ (literally, “me-by hunger happened.”) or ‘*i~a i’e~ sta_âsa*. ‘I have to go’ (“me-by going is.”).

Nouns. Nouns inflect for case and, in the oblique cases, for plural number. Inherent nominal categories that affect agreement and reference include gender and kinship. Nouns of quality or quantity may be used either attributively or substantively; e.g., ‘*oala* ‘big’ or ‘big one,’ *d’ü goa~ gâc*. ‘Give me two cows’ vs. *d’üa~ gâc*. ‘Give me two.’ Number systems are vigesimal rather than decimal.

Case. Case suffixes are added to nouns to indicate backgrounding, genitivity, or instrumentality. Backgrounding, typically indicated by the “oblique” suffix *-a* (V. *-š*), plural *-a~* (V. *-ân*), places a noun’s referent spatially in the cognitive background relative to unmarked nouns, which appear in the foreground as subjects of discourse. Except in Tr. and V., backgrounded nouns act as definite patients of non-past-tense transitive verbs or as agents of past-tense transitive verbs, as explained below under “Verbs.” Tr. distinguishes between patient (“accusative”) and (past tense) agentive forms: *žu~* ‘me’ vs. *žepe* ‘by me’; a distinction borrowed from neighboring Indo-Aryan languages. In V. backgrounded nouns represent definite patients only; agents are unmarked for case.

Distinct genitive singular forms appear only with some types of nouns; otherwise, the genitival function is indicated by unmarked or backgrounded plural forms: *mānš’e totas* ‘the man’s father,’ with genitival *-e*, plural *mānš’oa~ totaņas* ‘the men’s fathers,’ but *ĵuk totas* ‘the girl’s father,’ with no genitival suffix. In addition to indicating close possession, as in *âsk’e ĵuk* ‘his girl,’ genitive forms are used alone or with postpositions and prepositions to indicate various types of locality; e.g., *mānš’e meš* ‘with the man,’ *pât’apke* ‘on



the gun' (cf. *tâpk'a* 'gun'), *kâmn'ile* 'Kom Lake' (place name; cf. *nil'a* 'lake').

Instrumental forms indicate a noun as conduit of action; e.g., *tâpk'ea~vino* 'he was shot,' literally, "gun-through hit-was."

Number. A noun's plural number is shown by case endings and concordant verb forms: *'ušpoa~v'iana* 'he's beating the horses,' *'ušpa vin'â* 'the horses were beaten.' A few nouns are inherently plural: *dâr'i* 'beard,' *bâ·s'ano* 'clothes.'

Gender. Substantive nouns are inherently either masculine (unmarked) or feminine, their gender being typically shown by other forms, rather than by the basic form of the noun itself (although some feminine nouns show gender through the suffix *-ik*, as in *purd'ik* 'old woman' vs. *purd'uk* 'old man'). Some classes of qualitative nouns mark feminine singular forms with the suffix *-i* (e.g., *'oali âmo* 'big house,' *'oali* 'big one [fem.]' vs. *'oala mânsa* 'big man'). In addition to such attributive gender marking, feminine singular gender is shown by concordant verb forms (*'ušpa vin'i* 'the horse [fem.] was beaten' vs. *'ušpa vin'o* 'the horse [masc.] was beaten') and by case endings (*'ušpoa v'iana* 'he's beating the horse [fem.]' vs. *'ušpea v'iana* 'he's beating the horse [masc.]').

Kinship Nouns. Characteristic of the Nurestâni languages is the obligatory marking of possessor for kinship nouns; e.g., *bř'o* '(my) brother' (unmarked), *bř'oša* 'your brother,' *bř'os* 'his/her brother,' *bř'oñas* 'their brothers.' Possessor suffixes differ from dialect to dialect.

Nominal Location. Nouns are spatially located through pronouns, directional pointers, and locations relative to other items. Place names have inherent location. Once a noun is so located, it becomes *definite*.

Pronouns. Pronouns have separate forms for singular and plural number. Third person pronouns show a contrast of perceived vs. non-perceived, with a further contrast of proximal vs. distal for the perceived ones. Pronouns show different stem forms in singular backgrounded and genitival cases; e.g., V. *and'eš*, K.ktv. *i'e*, K.km. *'i~a*, A.s. *y'u~*, Kal.n. *u~*, Tr. *žu~* 'me'; V. *üt'eš*, K.km. *t'ua*, K.ktv., Kal.n. *t'u*, A.s., Tr. *t'o* 'thee'; V. *t'eš*, K.ktv. *k'u* 'whom,' K.km. *k'ua*, A.s. *ko*, Kal.n. *ko~*. Unmarked pronominal forms appear in [Table 1](#).

Directional Location. Characteristic of the Nurestâni languages is their system of directional location, which is built on sets of *directional pointers*. Such pointers may precede verbs of motion to indicate direction of motion, and they



may combine with *locational suffixes* to indicate directional locations.

As illustrated by Kâmviri, pointers come in functional pairs: *speaker-relative* (*vâ-* ‘hither,’ *pâ-* ‘thither’), *earth-relative* (*‘u* ‘up,’ *‘ü* ‘down’), *barrier-relative* (*ât’a* ‘in,’ *bâr’a* ‘out’), and *flow-relative* (*n’i* ‘downward, downstream,’ *č’a* ‘upward, upstream’). *vâ-* and *pâ-* may combine with a bound *pointer extension* that indicates earth-relative direction (*-ñu* ‘and up,’ *-a* ‘and down,’ *-e* ‘horizontally away,’ *-ü* ‘around, over’) to form a compound pointer; e.g., *pâñ’u* ‘thither and up,’ *v’âa* ‘hither and down,’ *p’ü* ‘thither and over.’

Locational suffixes may follow a pointer to indicate a specific location (adjectival *-ra*; e.g., *n’ira* ‘[the one] downstream,’ adverbial *-re*; *n’ire* ‘downstream there’), a nonspecific region (*-ñ*; e.g., *n’iñ* ‘downstream [somewhere]’), a region that extends away or outward (*-ma*; e.g., *n’ima* ‘away downstream’), plus some others. The semantic range of such suffixes varies among the languages.

The directional systems in the other Nurestâni languages are similar to that of Kâmviri, although data on Tr. are lacking. V. has a richer system of pointers, and both V. and A. allow prefixing of directional pointers to 3rd-person pronouns.

Item-Relative Location. Nouns are located relative to other items through one or two prepositions, numerous postpositions, and certain conventional locations.

The prepositions *pa-* in K. and A., *a-* in Kal., and *tu-* in V. have a general locative function: ‘to,’ ‘at,’ ‘in.’ They are inseparable prefixes to the nouns that they point to, so that possessives must stand before the prefixed forms; e.g., K.km. *‘i~ pâmo* (< *pa-âm’o*), V. *am târak* (< *tu-vâr’ak*) ‘to my house.’ In K. there is also a comitive preposition K.km. *s’e*, K.ktv. *s’â* ‘along with.’

The range of locations indicated by postpositions varies among the languages. In K.km. the major postpositions include *to* ‘to; at; in,’ *ste* ‘from,’ *meş* ‘with,’ *nâ-i~* ‘without,’ and *düje* or *ke~* ‘for.’ Postpositions follow genitival forms of nouns that show them.; e.g., *ask’o to* ‘to her’ (lit. ‘her place’), *tâpk’e meş* ‘with a gun,’ *t’ot ke~* ‘for Father.’

Conventional locations relative to a noun include *č’eř* ‘top,’ *kâr’ü* ‘base,’ *m’ük* ‘front, face,’ *pt’i* ‘back,’ and *‘or* ‘side.’ These are “possessed” by the noun, which is in the genitive case, and they require the preposition *pa-*; e.g., *m’is*



pâcerš ‘on (top of) the table’ (lit. “table’s to-top”), *âsk’e pâmük* ‘in front of him’ (lit. “his to-face”).

Verbs. Finite verbs indicate changes in a *subject* noun, which stands in the foreground of the speaker’s cognitive spatial depiction of discourse. Finite verbs consist of a verbal *stem* plus a pronominal suffix that indicates the subject; e.g., *v’iana-m* ‘I am hitting’; subject suffixes are illustrated for the verb ‘is’ in Table 3 (see below). Verbal stems are built on a verbal *base*, which is optionally followed by a modal/aspectual suffix; e.g., *v’ia-na-*, with progressive-mode suffix *-na*. A verbal base includes a verbal *root*, optionally followed by one or two “causative” suffixes (*-ov*) that indicate remote agency; e.g., *vi’oanam* < *via-ov-na-m* ‘I am getting someone to hit (something).’ Denominative verbal bases are built on a noun plus a verbalizing suffix (*-a*), which optionally may be followed by causative suffixes; e.g., *t’op* ‘heat’ + *-a* > *tâp’-a-* as in *tâp’anam* ‘I’m getting hot’ (intransitive), *tâp’oanam* ‘I’m heating (something) up,’ *tâp’avoanam* (< *top-a-ov-ov-na-m*) ‘I’m getting (someone) to heat (something) up.’ The verbal root encodes the cognitive prototype of the verb, which may be augmented by directional indicators or nouns that precede the verb to form an enlarged verbal prototype; e.g., *bâr’a* ‘out’ + *v’ia-* ‘hit’ > *bâr’a via-* ‘beat’ (a heart, lit., “out hit”), *ç’o via-* ‘shout’ (lit., “shout hit”). In addition, various other types of adverbs may appear before a verb to indicate spatial and temporal location, direction, manner, and negation, and speaker-hearer status particles may follow a finite verb (see below).

Modes and Participles. Verbal stems by themselves function as participles, which are attributive to their subject nouns. Stems distinguish four prototypical modes through three participial suffixes. In addition to an unmarked *immediate* mode, there is a *progressive* mode (*-na*, not used as an attributive participle in K.km.), a *conceptual* mode (*-la*), and a *past* mode (typically *-i*, but with irregular past stems for some verbs). The immediate, progressive, and past modes divide the progression of time into three zones: the progressive mode depicts the subject as it progresses through the verbal event, at the time of “now”; the immediate mode depicts the subject in the immediate future, which emanates from the subject’s current progression through time or from the speaker’s wish for the subject to be in such a state (a *command* mode); and the past mode depicts the subject as affected by a past event, with the speaker’s perspective on the event looking *retrospectively* into time.

Additionally, an *evidential* mode appears in forms consisting of past-stem plus



sta ‘being’: *tâp’i sta* ‘[the one that has] gotten hot.’ Here *sta* represents the subject in its post-action state, providing the evidence for a previous change.

Each of the participial stems must be followed by the feminine marker *-i* when depicting feminine-singular subjects. The *emphatic* marker *-o* follows a non-feminine stem to indicate the speaker’s vivid perception or imagination of the verbal event: *tapi’om* ‘I did indeed get hot,’ *tâp’alo* ‘it should indeed get hot.’ With the verb *âs’a-* ‘is’ it depicts a *realizational* (“mirative”) mode, indicating former ignorance or skepticism turned to vivid certainty: *čâṭ’a_âsom* ‘now I see that I was stupid.’ In K.ktv. a feminine emphatic marker *-e* appears with feminine-singular stems in forms parallel with non-feminine forms in *-o*: *k’uliesi* (< *kula-i-e=âsi*) ‘she would have done it,’ vs. K.km. *k’ulsi* or *k’utsi* (< *k’ula-i=âsi*), without the *-e*, which may have merged with the preceding *-i* in earlier times.

Forms Compounded with Auxiliary Verbs. Stems may be combined with the auxiliary verbs *âsa-* ‘is’ and *bu-* ‘happen’ to depict additional modes of cognitive processing. Forms of *âsa-* may be morphophonemically fused to a stem (indicated by =), or they may be independent words, as in K.km *v’ialâsam* ‘I would have hit’ (< *via-la-o=âsa-m*) vs. *v’ialaasam* (< *via-la âsa-m*) ‘I’m going to hit.’ For Kâmvirî the numerous finite verbal forms that may be generated from combinations of participles and auxiliaries are shown in Table 2 ([part 1](#); [part 2](#)).

The three simple participles depict perceptual (progressive), experiential (past), or prototypical (conceptual) events. An auxiliary verb combined with a participle serves to project the verbal prototype depicted by the participle into the speaker’s cognitive image of what he or she is saying. Among the fused forms of *âs’a-*, the past form *âs’i-* ‘was’ projects progressive and conceptual verbal prototypes into the past to depict past continuous or imagined resultant outcome, respectively. The present form *âs’a-* ‘is’ projects feminine 1st- and 2nd-person-singular subjects into the cognitive image, giving them a kind of cognitive distance that downplays their saliency in accordance with cultural preferences. The conceptual form *âs’ala/i-* ‘might be’ projects verbal prototypes into a probable imaginary world. Independent forms of *âsa-* project verbal prototypes into a conceptualized external world to indicate planned action or continuing resultant state of an action. As an auxiliary verb, *bu-* ‘happen’ projects a verbal prototype as a happenstance in a hypothetical world to depict hypothetical and suppositional modes of thought.



Non-Finite Forms. In addition to the above-mentioned participles, other nonfinite verbal forms are formed with suffixes to the verbal base. These include an *absolutive*, marked with *-ti*, indicating an actor in a state after having completed the verbal action (e.g., *v'iatī* 'having hit'), a *gerund*, marked with *sta* following the verbal base and indicating the nominalization of a verbal prototype; an *infinitive*, marked with *-̃*, indicating the completed state of a change as a goal of action. With nasalized markers similar to that of the infinitive are the *precedential adverb*, marked with *-̃u* added to an emphatically-marked past stem to indicate the time preceding the completion of a change, expressed negatively (with *n'â* 'not') as the time during which the change did not occur (*t'ua n'â oasõu šaŋ'e bo*. 'It happened before you came.');

and a *destination noun*, marked with *-ñ*, indicating a change as a goal of motion (*vâll'oañ gu-sa*. 'He went to call on [someone]'). There is also a *location noun*, marked with *-tõ* 'place,' indicating the location at which a change occurs, an *impulse noun*, marked with *-ik*, indicating an uncontrollable internal impulse in an experiencer (*ĩa j'ek bo*. 'I have to sit down.' ["me-by sitting-down-impulse happened"]) and an *adverb of motion*, marked with *-m*, indicating a change concomitant with motion: *kân'am oaso*. 'He came laughing.'

Non-finite forms show additional modal distinctions when used with certain verbs. A contrast between internal and external compulsion appears between *ĩa j'ek bo* 'I have to sit down' [internal compulsion] vs. *ĩa j'e sta bo* 'I have to sit down [external compulsion].' Infinitives combine with a small set of verbs to give specialized meanings; most importantly, 'e- 'go' forms a *passive*," as in *v'iã e-* 'get hit' ("go to [one's own] hitting"), and *b'â-* 'attain, be able' forms a *potential*, as in *v'iã b'â-* 'be able to hit' ("attain hitting"). In passive formations with *e-* a subject's lack of control is indicated by remote agency of the infinitive; compare the remote, uncontrollability of *jâñ'oã go* 'he got killed' [through no fault of his own] and the closer, controllable situation in *jâñ'ã go* 'he got himself killed' [through his own fault].

Comparison of Verbal Systems. The K.km. verbal system shares a prototypical base with the other dialects of K., and further with those of A. and Kal., while the system of V. has a rather different foundation (see below). The core group of A., K., and Kal. partition experiential time into the prototypical modes of prospective immediate future (e.g., Kal. *-a*) and retrospective past (K., A. *-i*, Kal. *-e/i*). In those languages there is a prospective conceptual mode (*-la-*) and an evidential mode (*_sta*). A. and K. have a prospective present progressive mode



based on an ancient *-n- (K. -na, K.e. -ta [*< *-nta-*], A.s. -ŋ-).

The retrospective past forms of A., K., and Kal. are based on the ancient past passive participle ending *-i-ta-, which yields modern -i. This participle determines a *retrospective perspective* on past events, with the speaker looking backward into time at the events. The subject of this participle is the noun at the temporal end-point of the verbal action. Thus, for intransitive verbs, the subject is the verbal actor in its final state, but for transitive verbs, the subject is the patient (affected) noun in its final state. In retrospective perspective the verbal actor stands farther back in time than the affected subject/patient noun and is therefore placed in the background through oblique-case markers. Contrastingly, the non-past verbal prototypes depict a *prospective perspective* on transitive action, in which the verbal actor stands in the foreground as subject, and a verbal patient is backgrounded through an oblique case marker. Retrospective perspective, determined by the patient-focus of the ancient past passive participle, is the cognitive basis for the so-called “split-ergative” systems of verbal subject reference found throughout most of the Indo-Iranian-speaking world.

The core languages differ in their choice and use of auxiliaries. A. agrees largely with the above-outlined uses of *âs'a-* and *b'u-* in K.km. Kal.n. uses *o-* ‘is’ for the close cognitive projection of a verbal prototype, *âs-* for placing the subject into progressive mode, and *oro-* for distal past or realizational mode. The role of *bü-* as an auxiliary in Kal. requires further research. In Kal.n. the auxiliaries take on the functions of the emphatic-realizational and progressive morphemes of K.km., so that Kal.n. *o-* corresponds to K. -o, Kal.n. *oro-o-* corresponds to K.km. *âs'o-*, and Kal. *âs-* corresponds to K. -na. Present-tense forms of the auxiliary verb ‘is’ for the various languages are shown in [Table 3](#).

Vâsi-vari (Morgenstierne 1949, Buddruss 2005) presents archaisms and special developments of its verbal system which imply a longstanding remoteness from the influence of the other Nurestâni languages. There is a contrast between 1st-person-plural subject suffixes that indicates the inclusion (-m) vs. exclusion (-mš) of the hearer. All verbal forms depict change in a prospective perspective; so that past changes are not viewed retrospectively, as in the core languages, but as past changes in an actor subject, depicted by a past active participle ending in -ok. Consequentially, definite patients are always backgrounded, and the V. verbal system is not of the “split-ergative” type. The progressive suffix is -mâ (f. -mi), rather than one based on *-n-; perhaps this form shares an origin with the K.km. directional zone marked with -ma (see



above), depicting outward motion. A “static participle” in *-gâ* (f. *-ik*) is perhaps functionally comparable to past participles with the evidential enclitic *sta* in the core languages. The auxiliary *as’a-* combines with the progressive, past, and static participles to depict present, past, and evidential finite verbs. Another auxiliary verb, *lâ-* ‘have,’ combines with the infinitive in *-in’ik* to depict conceptual change. Notable are the different auxiliaries used with intransitive and transitive past participles: the past participle plus a subject suffix depicts a transitive verb in the simple past mode, while a corresponding intransitive verb is depicted by the past participle plus a closely-compounded form of *as’a-*. Past perfect verbal forms consist of the past participle plus a noncompounded form of *as’a-* for intransitive verbs, while corresponding transitive verbs consist of past participle plus *lâ-*.

What little is known of Tregâmi (Morgenstierne 1952) shows an active, prospective past similar to that of V., but one that requires a special pronoun for the past actor-subject; e.g., *žepe to žya~tem* ‘I saw you’ (ibid., p. 123). Present-tense forms are built on a form with *-k*: *žã~kom* ‘I see,’ reminiscent of forms in nearby Pashaî dialects.

Enclitic Particles. A variety of enclitic particles may follow nouns and verbs to indicate conjunction, limitation, emphasis, speaker-hearer status, conditionality, or subordination. Thus, in K.km., after nouns: *že* ‘and,’ *di* ‘also,’ *de* (emphasis), *mi* ‘just; only’; after animate nouns: *o* (vocative masculine), *â* (vocative feminine), and *so* (vocative plural, after plural backgrounded case forms). After verbs may occur particles indicating the status of the speaker toward the hearer; these include assertive *o* (to male) and *e* (to female), supplicative *oa* (to male) and *ea* (to female), with rising sentence-final intonation, and interrogative ‘*â* (request for conformation), *e* (request for explanation, with half-falling intonation), and ‘*âa* (supplicative request for confirmation, with rising intonation). *bo* ‘suppose’ indicates non-past conditionality of the preceding clause; *to* indicates past conditionality or temporality of the preceding clause. Conjunction with a preceding sentence is indicated by post-verbal *čã*, which may also impart an emphatic value. Notable is the reportative particle *mma*, K.ktv. *mem*, Kal. *le*, which indicates that what the speaker is relating is hearsay.

Evolution of the Nurestâni Languages. Because of an absence of historical and archaeological evidence, it is comparative linguistics that provides the clues to the origin of the speakers of the Nurestâni languages. The evolutionary schema of the Nurestâni languages was first correctly discerned by Georg



Morgenstierne on the basis of his linguistic fieldwork during the 20th century A.D; see especially Morgenstierne 1945.

Within the Indo-European linguistic family the Nurestâni languages form a sub-group of the Indo-Iranian group, alongside the Iranian and Indo-Aryan (IA) sub-groups. The evolution of the Nurestâni languages, reconstructed internally and through comparison with other languages (e.g., Turner 1966), shows the following linguistic phases and steps:

1. Aryan Phase. Proto-Nurestâni speakers participated in most of the pronunciation processes that differentiated the speech of the early Āryas (proto-Indo-Iranian speakers) from that of other Indo-European speakers, notably:

a. Loss of nasality of the vocalic nasal vowel **n*, resulting in **a*; cf. K.km. *â-* but English “un-.”

b. Fronted vocalization of the laryngeal sound **H* (*ə*) to **i*, as opposed to **a* in other Indo-European languages; cf. K.km. *-i* [past-participial suffix] < **-i-ta-* < **-H-to-*.

c. The First Palatalization, in which the root of the tongue was raised, pushing the body of the tongue forward. This change produced lamino-alveolar affricates from non-labialized velar stop consonants, so that **k* > **č*, **g* > **j*, and **gh* > **jh*.

d. Subsequent loss of lip-rounding on the labiovelar stop consonants, so that **k^w* > **k*, **g^w* > **g*, and **g^wh* > **gh*.

e. Assimilation of **s* to the place of articulation of a preceding consonant or front vowel, so that **is* > *iš*, **čs* > **čš*, **rs* > *rš*, **ks* > *kš*.

f. However, in the earliest indication of linguistic independence, the proto-Nurestânis did not participate with the Āryas in the subsequent backing of **s* after the back vowel **u*; thus, e.g., K.km. *mus'a*, but Fārsi *muš*, Sanskrit *mûṣaka-* ‘mouse’; K.km. *ks'a-* (< **kus'a-*) ‘become instantaneously,’ but Fārsi *koš-*, Skt. *kuṣá-* ‘strike, kill’; K.km. *d'us* ‘yesterday,’ but Skt. *doṣ'ā-* (< **dausâ-*) ‘night.’

2. Early Iranian Phase. Speakers of the proto-Nurestâni languages appear to have been on the southeastern edge of the wave of Aryan expansion that



placed the earliest Iranian-speaking peoples in their present locations. Proto-Nurestānis partook in the basic phonological innovation that distinguished the Iranian speakers: they strengthened front-glottal tension to the exclusion of back-glottal tension. Strong anterior voicing, often tensed to produce acoustic noise, is the normal phonation type in today's Iranian languages, while the phonation of most of the region's Indo-Aryan languages is normally produced with some degree of posterior voicing, which under further tension produces contrastive tones and the whispery-voiced consonants ("voiced aspirates") of those languages. The predominance of front-glottal tension in the Iranian region had the following consequences:

a. The formerly whispery-voiced consonants merged with their anterior-voiced counterparts (**bh* > *b*, **dh* > *d*, **jh* > **j*, **gh* > *g*; e.g., Old Persian. *bumi-*, K.km. *b'üm*, but Skt. *bhūman-* 'earth').

b. Dentalization: the increased front tension on the glottis spread upward to bolster the already raised tongue root (from Step 1.c above), pushing the tongue further forward. The effect was to change the articulation of the affricates from lamino-alveolar to lamino-dental, so that **č* > **ć* and **ǰ* > **ź*. These dental affricates persist to this day in the Nurestāni languages, although *ź* has lost its affrication in most dialects. Examples include Kal. *ću*~ 'dog' vs. Skt. *śauna-*; K.km. *ź'o*~ 'knee' vs. Skt. *jānu-*. Meanwhile, the early Indo-Aryan speakers laxed the voiceless lamino-alveolar affricate **č* to a spirant *š*, while leaving the voiced affricates *ǰ* and *jh* undisturbed, as in the preceding examples.

After the emergence of the affricates during the First Palatalization, there were two phonological processes that probably emanated from the Iranian region to sweep across the entire Indo-Iranian-speaking area:

c. The Second Palatalization, in which speakers anticipated the fronting of a following vowel to a velar stop (from an earlier labiovelar, Step 1.d above); thus **ketwāra-* 'four' became Kal. *čatā*, Avestan *čaθwar-*, Skt. *catv'ārah*; **giHw'o-* 'alive' became K.km. *ǰ'üa-*, Av. *jīvya-*, Skt *jīva-*.

d. Loss of fronting and rounding on open vowels, resulting in the merging of earlier **e* and **o* with *a*.

At this point the proto-Nurestānis departed from Iranian influence and did not partake in the further changes that characterize the Iranian languages.



Through a series of laxing processes, Iranian speakers reduced *s* to *h*, stop consonants to spirants, and, in the east, dropped the occlusion of the dental affricates, so that **č* became *s* and **ž* became *z*; in the west, as attested by Old Persian, the dentality of those affricates was strengthened, so that **č* > *θ* and **ž* > *d*.

As Morgenstierne (e.g., 1945) first demonstrated, it was the development of the affricates produced from the First Palatalization that mainly distinguishes the Nurestāni languages from the Iranian and Indo-Aryan ones. Where Nurestāni shows *č*, Iranian shows *s* or *θ* and Indo-Aryan shows *š*; where Nurestāni shows *ž*, Iranian shows *z* or *d* and Indo-Aryan shows *ǰ* or *ǰh*; e.g., K.km. *d'uč*, Av. *dasa-*, Skt. *daša-* 'ten'; K.km. *ž'o*~, Skt. *jānu-* 'knee'; K.km. *ž'im*, Skt. *hima-* (*h* from earlier **ǰh*) 'snow.'

3. Transitional Phase. Perhaps between the time that the early Nurestānis separated from the Iranians and the time that they came under Indo-Aryan influence (v. below), another distinguishing change in the Nurestāni languages took place: loss of aspiration after voiceless stops. Indo-Aryans retained the aspirated stops, while Iranians spirantized them; thus **khand-* 'laugh' > K.km. *kān'a-* but Fārsi *xand-*; K.km. *p'ul* 'small spherical object' vs. Skt *ph'ala-*. If these changes had happened later, there would be little motivation for a loss of aspiration in an Indo-Aryan-speaking milieu that maintained it.

4. Indo-Aryan Phase. Having been initially out of range of the Indo-Aryan side of the Aryan expansion, the proto-Nurestānis subsequently entered the Indo-Aryan sphere, where they acquired many IA loanwords and participated in many of the Middle Indo-Aryan (MIA) changes that characterize the northwestern IA languages. According to their oral traditions, before they settled in Nurestān, the Nurestānis migrated from Khorasān to Kandahār to Kābol to Kāpisa to Kāma, at the confluence of the Konar and Kābol Rivers. Kāma at that time was solidly within the Indo-Aryan-speaking world; even today speakers of the Indo-Aryan language Pašaī border the district, which is now Pashto-speaking. Processes of change during this phase are revealed through comparison with Sanskrit and reconstructed Old Indo-Aryan (OIA) forms (Turner 1966); they included:

a. Retroflexion of tongue-end consonants next to *r* or after vocalic *r*. Examples with postconsonantal *r* include **čr* > *ç* as in **črāva-* > K. *ç'o* 'shout,' Skt. *šrāva-*, **ācru-* > K.km *âç'ü* 'tear,' Skt. *āšru-*, but initial *tr* and *dr* remain unchanged (K. *tr'e* 'three,' *dr'u*~ 'bow'). With preconsonantal *r* there are two outcomes. With



a following voiceless *t* the cluster *rt* > *ṭṭ*, with compensatory lengthening of the preceding vowel: **varta-* > **vâṭṭa-* (> **vâṭa* > K. *v'oṭ* 'stone'). With a following voiced consonant, the outcome is *r*: **ardu-* > K. *ar'u* 'peach,' **vârdha-* > K. *vâr'a-* 'grow,' **parṇa-* > **pâra* (> K. *p'or* 'leaf'). The development after vocalic *r* is exemplified by **bhr̥t'a* > **br̥ṭa* > K.km. *bâr'a* 'taken.'

b. Nasal latency, with the nasalization of a preceding *n* spreading across a vowel to a following stop consonant; e.g., Skt. *nad'ī-* > **nandī* > K. *nan'i* 'river,' Skt. *nagara-* 'town' > **nanagara* > K. *nāṅ'ar* 'name of a village in Chitral.'

c. Weakening or loss of single intervocalic stop consonants. Thus, labials become *v*: *n'apât-* > K.km. *nâv'o* 'grandson,' **â-bhr̥t'a* > K. *âv'aṛa* 'brought.' Dentals are lost: *-ita* (passive-participle ending) > K. *-i* (retrospective verbal suffix), *pad'a-* > K. *p'u* 'step'; *godh'ûma-* > A.s. *go:m*, Kal.n., Tr., K. *g'um* 'wheat'; however, *d(h)* becomes *l* in V.: *godh'ûma-* > V. *uly'um*. Affricates remain unchanged. Retroflexes become flapped *r* in A. and Kal., *r* in V., and are further reduced to *ř* in K. and Tr.: *ghoṭa-* > A.s. *goṛa*, Kal.n. *guṛa* 'horse,' Tr. *goṛa*; *mṛḍâ-* > **mṛḍâ-* > V. *mira*, K.km. *muř'i* 'clay.' Velars are lost: **çâka-* > Kal.n. *çâ* 'kind of bean' K. *ç'o* 'greens,' Skt. *šâka-*, **â-gascâ-* > K. *âçâ-* 'come,' Skt. *âgaccha-*. Intervocalic *n* and *ṇ* merge to become a nasal flap in A. and Kal. and a nasal approximant in K.: Skt. *kāṇâ-* 'one-eyed' > A. *kâña*, K.km. *kâñ'a* 'blind'; but *n* is retained in V.: Skt. *stenâ-* > V. *ištn'e*, K.km. *št'âña* 'thief.'

d. Simplification of consonant clusters. Consonants in clusters assimilate to each other and reduce, in most cases, to a single consonant. In intervocalic clusters of stops, the second one predominates: *ut-pâtaka-* > K. *up'a* 'scorpion'; *saptâ* > K. *s'ut* 'seven,' *citta-* > K. *č'it* 'will,' *abhukta-* > K.ktv. *âvt'a*, *kiṭṭâla-* 'pot' > K.km. *kṭ'ol*, A.s. *čičâl* 'stomach'; **makkāṭa-* > K.km. *mâk'aṛ* 'monkey,' etc; also *âtmân-* > K. *âm'u* 'oneself.' Postconsonantal *v* shows various outcomes: loss, occlusion with anaptyxis (initially in A.) or assimilation (V.), or retention with anaptyxis (K.); e.g., *catv'âraḥ* > A.s. *çâtâ:*, Kal., Tr. *çatâ*, K.ktv. *štav'o*, V. *čp'u* 'four'; similarly, in initial clusters, with occlusion and anaptyxis in A.: **sva-çura-* > *šipasu*, but loss in K.ktv. *saçî'uṛ* 'father-in-law,' with anaptyxis in K.: **çvâta-* > K.ktv. *çav'o* 'rhubarb,' with loss: *svâsr̥* > K. *s'us* 'sister,' **dvara-* > Kal. *dor*, K. *d'u* 'door.' Postconsonantal *y* palatalizes a preceding dental stop: *satyâ-* 'true' > K. *s'uč* 'pre-Islamic ritual utterance' *madhya-* > K.km. *pâ-m'üč* 'in the middle.' An *r* before *v* drops: *sârva-* > A.s. *savâ:k*, K.ktv. *s'uv* 'all'; before *y* it is reduced in K.km. and lost with retained fronting in Kal.: *kâryâ-* > K.km. *k'oṛ* 'to be done,' Kal.n. *kô* 'work.' Nasals are reduced in A. and lost in K. and V. before voiceless stops: *vasanta-* > Kal.n. *osunt*, A.s. *voso~t*, K.ktv. *vas'ut*, V. *usti* 'spring.'



Post-nasal voiced stops are absorbed in Kal., K., and, in later final position, in A., but in V. it is the nasal that is lost: **kanda-* > A.s. *kanda*, K.kt. *kan'a* ‘tree,’ Kal.n. *kana* ‘twig’; **khand-* > K.ktv. *kan'e-*, V. *wyōida-* ‘laugh’; *stamba-* ‘post’ > K. *št'um*, V. *ištobu* ‘tree(-trunk), stem’; *k'āṇḍa-* > A., Kal. *kāṇ*, K. *k'oṇ* ‘arrow.’

e. Anticipation of *r*. In medial clusters of *r* plus non-apical consonant, the *r* is anticipated to the consonantal onset of the preceding syllable: *dīrghá-* > Kal. *drigala*, K.km. *draṇ'aña*, V. *ḷign'i* ‘long’; *kárman-* > A.s. *kram* ‘work.’ Sporadic anticipation of postconsonantal *r* in this period appears in *támisrâ-* ‘darkness’ > K.km. *trâms̄'a* ‘twilight,’ **vikriṇâti* ‘sells’ > K.km. *vř'e č* (with *č* from another source); but such anticipation is normal in K. today, as seen in forms with the locative prefix: *pa-* + *gř'om* > **břag'om* > K.km. *břâk'om* ‘in the village.’ (Development of **r* is fully treated in Morgenstierne 1947, Hamp 1968.)

f. Development of vocalic *r*. Vocalic *r* became *a*, *â*, *i*, or *o* depending on environment or dialect; e.g., *rṣabhá-* > A., K.ktv. *âṣ'a* ‘bull,’ *kṛtá-* > K.ktv. *kař'a* ‘done,’ *r̥kṣa-* > A., K. *íc*, Kal. *oc*, Tr. *woc* ‘bear,’ *kṛṣí-* > K., Kal. *k'iṣ* ‘ploughing.’

g. Assimilation in diphthongal sequences. The sequences *aya*, *ava* are reduced to the open front or back vowels *e* and *o*, respectively: *tráyah* > K., Kal. *tr'e* ‘three’; *náva* > A.s. *no*, > K., Kal. *n'u* ‘nine.’

h. Loss of word-final syllables, as illustrated by most of the preceding examples; but final *u* and *i* in some instances may have been retained, as seen in **ácru-* ‘tear,’ **ardu-* ‘peach,’ and *nad'î-* ‘river,’ cited above.

The many changes and borrowings of this phase have lead some linguists to erroneously classify the Nurestâni languages as belonging to the Indo-Aryan subgroup of Indo-Iranian; but as can be seen from the changes in Phases 1 to 3 above, the proto-Nurestâni speakers were well on their way to linguistic independence before they fell under Indo-Aryan influence.

5. Nurestân Phase. According to local traditions, the Nurestânis were expelled from the Kâma region during the Ghaznavid invasion of Nangarhâr in the early 11th century CE, ending up as refugees in the middle reaches of the Konar, Pech, and Laghmân Valleys before dispersing to their current abodes in the isolated valleys of the Hindu Kush. In the millennium since the Nurestânis entered Nurestân, their languages have undergone further changes that distinguish today’s dialects. Notable among these are the progressive raising of the tongue’s dorsum, with concomitant fronting and rounding of stressed



vowels. Syncope of close vowels occurs in K. and V. Anticipation of fronting and rounding to vowels in preceding syllables (umlauting) is characteristic of Kal. and V. Consonants in A. have undergone changes reiterating those of Phases 1 and 2 above, while those in V. have been lost initially, resulting in virtually no comprehension of V. by speakers of other Nurestâni languages. Many developments during this phase are obscured by interdialectal borrowings, but sequences of phonological changes and their underlying processes can be reconstructed as follows:

a. Assimilations and dissimilations of successive syllable onsets. Successive consonantal syllable-onsets tend to form an articulatory unit, with one feature selected to dominate both onsets. When the second onset contains an affricate or a stop, a preceding affricate may be deaffricated, or a preceding spirant may be affricated; both onsets are usually produced with the same articulator. Resulting onsets may coalesce through later syncope (Step 5.d below). Such assimilations appeared sporadically in individual dialects, some spreading over wider areas and others remaining locally confined. Examples include:

i. Initial affricates lose spirancy to become *d* or *t*: **žizû*- ‘tongue’ (cf. Avestan *hizû*-, Skt. *juhû*-, *jihv’â*-) > **dižû*, becoming K.ktv. *d’iz*, K.km. *d’íc*, V. *luz’uk* (with suffix *-k*), but without dissimilation in A. *žu*: (< **žižu*: < **zizû*); Kal. *žip* is an IA loanword; **žasta*- ‘hand’ (cf. Skt. *hásta*-) > **dasta* > A., Kal. *dost*, K.km. *d’üst*, V. *l’ust*, if this is not an early borrowing of Fārsi *dast*; **ćiča*- (< **ćikya*-, cf. Skt. *šikyà*- ‘carrying-sling’) > **ćíca*- > K.km. *tić’a* ‘goatskin sack,’ but K.ř. *ćac’a*, K.ktv. *sać’a*; Skt. *jyeṣṭá*- ‘eldest’ > A.s. *diṣṭo*, Kal. *düṣṭö*, but K. *jeṣṭa*; MIA loanword **ḷandra* (< OIA *yantrá*- ‘device’) > **ḷandra* > **ḷāṇḍr’a* > K.km. *ḷraṇr’a* [ḷraṇḍr’i] ‘hand-mill.’

ii. Initial spirants assimilate following occlusion to become affricates: **šrčil’a* ‘slack’ (< OIA **šrthilá*-) > V. *čič’il*, Kal. *čičila*, A. *ćíčila*, K.km. *čil’a* ‘soft’ (with syncope), but with no occlusion in K.ktv. *šač’ala*; Skt. *sūc’ī*- ‘needle’ > **čuči* > K. (*čam*-)*č’ač*, A. (*âr*-)*ćus*; **sasč’anka* ‘coagulated’ (< **sasčanaka* < OIA **sa~styânaka*-) > A.m. *čučo:η* ‘buttermilk solids,’ A.s. *čoço:η*, but K.km. *sač’üη*, K.ktv. *saći’uη*.

Other processes of this type include:

iii. dissimilation of *st* (and later, *čt*) to *št* in K. and Kal.; e.g., K. *št’um* ‘tree’; K. *št’or* ‘quiver,’ K.km. *št’o* ‘4.’



iv. Voicing assimilation on the locative prefix *pa-* in K., which becomes *ba-* before voiced consonants; e.g., K.km. *pâkṭ'ol* 'in the stomach' but *bâg'otr* 'on the upper arm.'

b. Dorsal raising. For vowels, progressive dorsal raising determines the following sequence:

i. Dorsal raising starts in A., K., and V., where *a* > [ə] and, in A., *o* > [ɔ] with a resulting slight fronting of *u*. But in A. and K. an initial *a* is strengthened by opening to *â*, so that no initial *a* appears in those languages.

ii. In Kal., K.km., and V. further dorsal raising and fronting pushed accented *u* forward to *ü* (*iu* in K.ktv.) while raising *o* to *u*; e.g. *duv'â* > Kal., K.km. *d'ü*, V. *l'ü* 'two,' but A. *du*; Skt. *doṣ'ā-* 'night,' Kal., K. *d'us*, V. *ul'us* 'yesterday,' but A. *dos*; *ôṣ'âdhi-* 'herb' > **oṣ'a* > K. *uṣ'a* 'medicine.'

iii. In A., Kal., and K. accented *a* is rounded, except when it was in an ancient suffix (e.g., *-aka*). In A. and Kal. it becomes *o*; with further dorsal raising it becomes *u* in K.; e.g., *saptá* > **s'at* > A., Kal. *s'ot*, K. *s'ut* 'seven'; loanword **aṣpa* 'horse' (cf. Skt. *áṣva-*) > K.km. *uṣpa*.

iv. In K. accented *â*, when not followed by *i*, is rounded to a phonetically open *o*, continuing the trend started with the rounding of *a*, as in K.km. *gr'om* but A., Kal. *grâm*, V. *g'am* 'community'; K.ktv. *âz'or* but A. *âžâ:r*, Kal. *âžâr* 'four hundred.'

v. In Kal. and Kâta-vari, but not in Kâmviri, nasalization is enhanced by velar and dorsal raising, so that nasalized *o* > *u*; e.g., K.e. *u~ć*, but K.km. *o~ć* 'I', Kal. *punč*, but A. *po~č* 'five.' In K.ktv. velar raising in all nasal environments is taken to the point where the velum is closed off to exclude vocalic nasality; e.g., *'uze* 'I,' *âč'e* vs. K.km. *âče~* 'eye,' *zar'a* vs. K.km. *žar'a* 'red'; also next to nasal stops: *gr'um* 'community,' *m'uč* vs. K.km. *m'oč* 'man,' *n'uṭ* vs. K.km. *n'oṭ* 'dance,' but not always in final open syllables: *ân'o* 'meat,' *âṇ'o* 'fire.'

vi. In V. raising of the tongue extends to earlier **e*, yielding *i*: *m'iza* 'urine,' (cf. Skt *meha-*), *iṣtn'e* 'thief' (cf. Skt. *stená-*), with syncope of the resulting *i*.

c. Initial-syllable opening ("vridhhi"). In K.km. an *a* in initial syllables is strengthened by opening to *â*, K.km. *vâs'ut*, vs. K.ktv. *vas'ut* 'spring,' *kâlṣ'a* vs. K.ktv. *kaṣ'a*, Kal. *kalaṣa* 'Kalasha.'



d. Syncope. A characteristic of K. and V. is the syncope of pretonic close vowels between voiceless obstruents and certain other consonants, as illustrated by K.km. **šup'iš* > *šp'iš* 'drizzle [noun]' and **šupiš'a-* > *šupš'a-* 'drizzle [verb].' In K. syncope produces many complex initial consonant clusters, e.g., K.km. *pṭ'i* 'back,' *ṭk'u* 'nail,' *čk'ara* 'whey solids,' *kṣṭ'a* 'unadulterated,' *trk'iṭ* 'knucklebone,' *pštr'a* 'broken.' In K.km. there is assimilation of resulting homorganic clusters: K.km. *č'ūr* vs. K.ktv. *sači'ur* 'father-in-law,' K.km. *čal'a* vs. K.ktv. *šač'ala* 'soft.' In K.ktv. syncope is apparently subphonemic with rounded vowels after (velar?) stops, the rounding being retained on the stop; e.g. Kal. *kuṭa*, K.ktv. *kuṭ'a* [k^wṭ'i]; K.km. *kṭ'a* 'lame,' Kal. *küşü*, K.ktv. *kuṣ'i* [k^wṣ'i], K.km. *kṣ'ü* 'bean.' In V. syncope occurs ubiquitously in interior syllables, less frequently in initial syllables; e.g., *pš'ik* 'cat,' *psn'ok* 'thing,' *kšč'u* 'crippled,' *kṣ'u* 'left,' *čne-* 'sneeze.' Notable is the treatment of **duṣit'â* 'daughter,' with early internal syncope before the loss of *t* in V.: **duṣit'a* > **d'ušta* > **l'ušta* > *l'üšt*, but with spreading of *u*, loss of *t*, and subsequent syncope in the first syllable of the other languages: **duṣut'a* > **duṣ'ua* > **d'j'ua* > Tr. *ju*;, K., Kal. *ṣ'ü*; and with intermediate dentalization in A.: **duṣut'a* > **duṣ'ua* > **duž'ua* > *zua* > A.s. *zu*;

e. Epenthesis. Epenthesis before an earlier initial *st* occurred throughout Nurestân, except perhaps in Tr.; e.g. A. *istu*~, Kal. *üstüm* 'pillar,' V. *ištob'u* 'stem.' In K. the epenthetic vowel was later lost, but left traces with the locative prefix *pa-* in K.km.: *št'ü*~ 'pillar,' but **pa-ištü*~ > *p'eštü*~ 'to the pillar' (not ***pâšt'ü*~; but cf. K.ktv. *pšti'urâ*). In V. epenthesis occurred before all initial clusters beginning with *s* plus occlusive, as well as before initial continuants and clusters with subsequently lost *r* and *v*: *üşk'öp* 'bridge,' cf. Skt. *skambha-*, *üşp'u*~ 'flute' < K. *šp'o*~, *ul'us* 'yesterday,' cf. Skt. *doṣ'â-*, *un'ü* 'new,' cf. K.km. *nu-i*~, *üč'ü* 'horn' < **črû-*, *uč'âpar* 'rhubarb' < **čvâtvara-*. The basic epenthetic vowel is *i* in each language, with added anticipatory rounding in Kal. *ü* and V. *ü* and *u* (see next).

f. Vocalic-component spreading. Anticipation of vocalic fronting to a preceding vowel is usual in Kal., producing the front vowels *ä* and *ö*, and grammatical alternations in stem vowels with a following feminine suffix *i*; e.g., Kal.n. *dä* 'beard,' cf. K.km. *dâr'i*, *sö* 'sun,' cf. Skt. *s'ürya-*; *čüväli* 'walker' (fem.) vs. *čüvala* (masc.), *čaṭäki* 'sharp' (fem.) vs. *čaṭaka* (masc.). In V. there is anticipation of both fronting and rounding; e.g., *iž'i*~ 'eye,' cf. K.km. *âč'e*~, *üč'ü* 'tear,' cf. K.km. *âč'ü*, plus the forms cited above. In K. fronting extends through following velar consonants to produce subphonemic palatalization (e.g., K.km.



v'âik [v'akʷ] 'lamb'); similarly in V., with a wider range of following consonants (*üšk'öp* [üşky'öp] 'bridge,' *ıştob'u* [ištʷob'u] 'stem,' *ütն'ok* [ütny'ok] 'cultivated field'), the progressive palatalization is probably non-phonemic.

g. Voiced deaffrication. In all dialects except K.km. and Tr. the voiced apico-dental affricate *ž* laxes to *z*: K.km. *ž'otr* 'kinsman' but K.ktv. *z'otr* 'affine,' A.s. *zâ:tr*, V. *z'ât*; K.km. *ž'u* 'milk' but K.ktv. *zu*, A.s. *zo*; Kal.n. *zor*, K.km. *ž'im* 'snow,' Tr. *žim* but K.ktv., Kal. *z'im*, V. *z'ima*.

h. Developments of *r*. Initial **r* is strengthened to *ṛ* in A. and V. (further to *ṛ* when palatalized in V.), *j* in Tr., bolstered with *v* to *vř* (phonemically *vr* ?) in Kal., and weakened to *ř* in K.: e.g., A. *zo:-kânda*, Tr. *jo*, Kal. *vřo*, K. *ř'u* 'deodar,' A.s. *zât-â:r*, V. *ž'eṭ*, Tr. *jâtr*, K. *ř'otr*, Kal.n. *vâtr* (< **vřâtr*) 'night.' Postconsonantal **r* after non-apicals is reduced to *ř* in A. (phonetically) and K.; e.g., A. *brâ* [břa] 'younger brother,' K. *bř'o* 'brother,' A. *grâm* [gřam], K. *gr'om* 'community.' It is lost in V. except after a dental stop, where its retroflexion is absorbed or subsequently palatalized before a front vowel: *b'â* 'Brother!,' *g'am* 'community,' *uṭ'us*, K. *tr'us* 'avalanche,' *ḍui*, K *dr'u* 'head-hair,' *čü'u*, K.km. *tr'ua* 'yoghurt,' *jign'i*, K.km. *draṇ'aña* 'long,' *wyâč'i*, K.km. *v'etr* 'fairy'; but the loss after non-dentals must have followed Step 5.j.i below. Final postvocalic **r* is lost in some cases; e.g., Kal. *dor* but K. *d'u* 'door,' Kal. *zor* but A. *zo*; K.km. *ž'u* 'milk.' In K.ktv. *ř* metathesizes or drops before *i*: K.ktv. *křu·i*, K.km. *kuř'i* 'dog,' *ka·i* [kiy'i], K.km. *kâr'i* 'made' (fem.).

i. Consonantal development in Âşkuṇu. As shown by examples from A.s., A. underwent developments in its lingual obstruents that in part repeat changes that occurred in the Early Iranian Phase (2) above. The data are highly obscured by interdialectal borrowings, and it is not possible to find a sequence of changes to account for all the outcomes of earlier forms. The responsible articulatory processes were primarily deaffrication, laminalization before front vowels ("palatalization"), and "prognathizing," i.e., slightly protruding the jaw while keeping the tongue's blade fixed against the back of the lower teeth, resulting in a forward shift in the point of articulation from alveolar to dental and from apico-alveolar (retroflex) to lamino-alveolar. These processes occurred perhaps in the following order:

i. Deaffrication:

a). of *ć* to *s*: *saṇu* 'person from Wâmâ' vs. K.km. *cân'u*; *sâu* 'branch' vs. K.ktv. *ć'ov*; *gas* 'length of outstretched arm' vs. K.km. *g'eć*; with further palatization



ṣ̌i:ṭ ‘fertilizer,’ vs. K.km. *ć’iṭ*, *viši* ‘twenty,’ vs. K.km. *vić’i*.

b). of *ǰ* > *ž* before *i*: *ži* ‘sinew’ vs. K.km. *ǰ’a*; *žirik* ‘shame’ vs. K.km. *ǰar’ik*.

Deaffrication of *ž* previously occurred in Step 5.g above.

ii. Palatalization before *i*:

a). of *k* to *č*: *čilâ* ‘cheese,’ (with further dentalization) vs. K.km. *kil’âr*; *čiṭâl* ‘stomach’ (via **čiṭâl*, with further palatalization) vs. K.ktv. *ḳti’ol*, K.km. *ḳt’ol*.

b). of the apical spirants *s* > *š* and *z* > *ž*: *šikâ*: ‘fat,’ vs. K.km. *ski’o*; *žim* ‘snow’ vs. K.ktv. *z’im*; also *ṣ̌i:ṭ* and *viši* above.

iii. Prognathizing, resulting in:

a). Dentalization of the alveolars *č* > *ć*, *ǰ* > *ž*, and *š* > *s*: *čâm* ‘skin’ vs. K. *č’om*; *ćiatr* ‘carved design’ vs. K.km. *č’etr*; also *čilâ* above; *žâda* ‘other’ vs. Kal.n. *ǰâta*; *žâl’âi* ‘duck’ vs. K.ktv. *ǰâl’âi*; *žit* ‘body’ vs. Kal.n. *ǰit*; *sâl* ‘stable’ vs. Kal.n. *šâl*; *sâli* ‘rice plant’ vs. K.km. *šâl’i*; but before *i* the laminal spirants of Steps 5.i.i.b and 5.i.ii.b remain.

b). Laminalization of retroflexed *ç* > *č* and *ş* > *š*: *čila* ‘abomasum’ vs. K. *çal’a*; *šij* ‘horn’ vs. K., Kal. *ş’ij*; also *čiṭâl* above.

However, this sequence of processes leaves unaccounted the change of *ǰ* > *z* seen in *zu*: ‘daughter’ vs. K.km. *ǰü*, *zâliak* ‘omasum’ vs. K.km. *ǰalik*, *pamaz* ‘in the middle’ vs. K.km. *pâm’üč* (< **pa-maj*).

j. Consonantal Development in *Vâsi-vari*. The most aberrant of the Nurestâni languages is *V.*, which, in addition to the changes noted in previous steps, has undergone loss of initial consonants and various fronting and voicing assimilations. The analysis of sound changes in *V.* is due to Morgenstierne (1949) and includes:

i. Lenition of initial occlusion, so that initial *k*, *g*, *ǰ*, *t*, and *p* were lost, *b* > *v*, *č* and *ž* (v. Step 5.g) > *z*, *č* > *ž*, while initial (and medial) *d* had earlier become *l* (Step 4.c above); e.g., *ip’a* ‘Kâta’ < **kitva* < **kitivâ* < **kântivâ*, cf. K. *kt’ivi*, Pashto *kântiw’â* (place name); *uly’um* ‘wheat plant,’ cf. Skt. *godh’ūma-*; *z’o* ‘cow dung,’ cf. K.km. *ć’u*; *ešte* ‘elder,’ cf. Skt. *jyēṣṭha-*; *y’u* ‘thou,’ cf. K.km. *t’ü*; *y’â* ‘father,’ cf. Skt. *pit’ā* ‘father’; *v-* ‘become,’ cf. Skt. *bhū-*; *žim’a* ‘iron,’ cf. K. *čam’a*. Initial consonants before *r* remained, with the *r* subsequently lost (v. Step 5.h above).



ii. Intervocalic voicing of affricates: Intervocalic voiceless affricates, otherwise unaffected by the changes in Step 4.c, become voiced: *lez'e* '10,' cf. K. *d'uč*; *uz'â* 'herding,' cf. K.km. *pâc'o*; *ürj'uk* 'light,' cf. K. *ř'uč*; but the development of -č- is ambiguous.

iii. Palatalization of retroflex stops: In the environment of front vowels, retroflex stops are palatalized, so that $t > č$ and $d > j$; e.g., *ič'i* 'bone,' cf. K. *ât'i*; *müj'ü* 'drum' < **mađü* < **mañdu*, cf. K.km. *mâñ'ü*. This change must follow the absorption of *r* after dental stops treated in Step 5.h; cf. the similar change in A., Step 5.i.iii.b.

6. Afghan-Islamic Phase. After entering Nurestān, the pre-Islamic ("Kafir") Nurestānis managed to hold off encroachments from their Muslim neighbors, until they were conquered by the Afghans in 1896 A.D. At that time the "Kafirs" were forcibly converted to Islam, and since then they have been assimilating thousands of words of Arabic, Persian, and Pashto origin into their lexicons, to the detriment of many traditional terms. Despite the changes brought about by such borrowings, the Nurestāni languages continue to thrive and do not appear to be in danger of dying out.

BIBLIOGRAPHY

Georg Buddruss, "Māra's Mühlenbau"; Analyse eines Prasun-Textes aus dem afghanischen Hindukusch," in *Hrdā mánasā: Studies presented to Professor Leonard G. Herzenberg on the occasion of his 70-birthday*, ed. Nikolai N. Kazansky et al., St. Petersburg, 2005, pp. 653-92.

Idem, "Drei Texte in der Wama-Sprache des afghanischen Hindukusch," in *Indoiranskoe yazykoznanie i tipologiya yazykovykh situatsii, sbornik statei k 75-letiyu professora A. L. Gryunberga (1930-1995)*, ed. M. N. Bogolyubov, St.-Peterburg, 2006, pp. 177-200.

Almuth Degener, *Die Sprache von Nisheygram im afghanischen Hindukusch*, Wiesbaden, 1998.



Ghulâmullâh, *Grâmar-e lesân-e nurestâni*, manuscript grammar of Kâmviri in Universitetsbiblioteket, University of Oslo, 1966.

Aleksandr L. Gryunberg, *Yazyk Kati: teksty, grammaticheskiï ocherk*, Moskva, 1980.

Eric P. Hamp, "On *r in Kafir," in *Studies in Indian Linguistics (Volume presented to Prof. M. B. Emeneau on his sixtieth birthday year)*, Poona, 1968, pp. 124-37.

Georg Morgenstierne, "The Language of the Ashkun Kafirs," *NTS* 2, 1929, pp. 192-289.

Idem, "Additional Notes on Ashkun," *NTS* 7, 1934, pp. 56-115.

Idem, "Indo-European k' in Kafiri," *NTS* 13, 1945, pp. 225-38.

Idem, "Metathesis of Liquids in Dardic," in *Festskrift til Professor Olaf Broch*, Oslo, 1947, pp. 145-54.

Idem, "The Language of the Prasun Kafirs," *NTS* 15, 1949, pp. 186-334.

Idem, "Linguistic Gleanings from Nuristân," *NTS* 16, 1952, pp. 117-35.

Idem, "The Waigali Language," *NTS* 17, 1954, pp. 146-324.

Idem, "Languages of Nuristân and Surrounding Regions," in Karl Jettmar and Lennart Edelberg, eds, *Cultures of the Hindukush: Selected Papers from the Hindu-Kush Cultural Conference Held at Moesgård 1970*, Beiträge zur Südasienforschung, Südasien- Institut, Universität Heidelberg 1, Wiesbaden, 1974, pp. 1-10.

Richard F. Strand, "Notes on the Nûristânî and Dardic Languages," *JAOS* 93/3, 1973, pp. 297-305.

Idem, "Locality and Nominal Relationships in Kamviri," in *Proceedings of the Conference on Participant Roles: South Asia and Adjacent Areas*, ed. Arlene R. K. Zide, David Magier, and Eric Schiller, Bloomington, Indiana: Indiana University Linguistics Club, 1985, pp. 48-57.

Idem, *Richard Strand's Nuristân Site*, <http://users.sedona.net/~strand>, 1997-2010.



Idem, Review of *Die Sprache von Nisheygram im afghanischen Hindukusch*, by Almuth Degener, *Acta Orientalia* 60, 1999, pp. 236-44.

Ralph L. Turner, *A Comparative Dictionary of the Indo-Aryan Languages*. London, 1966.