

Chapter 10

Gender typology and gender (in)stability in Hindu Kush Indo-Aryan languages

Henrik Liljegren

Stockholm University

This paper investigates the phenomenon of gender as it appears in 25 Indo-Aryan languages (sometimes referred to as “Dardic”) spoken in the Hindu Kush-Karakorum region – the mountainous areas of northeastern Afghanistan, northern Pakistan and the disputed territory of Kashmir. Looking at each language in terms of the number of genders present, to what extent these are sex-based or non-sex-based, how gender relates to declensional differences, and what systems of assignment are applied, we arrive at a micro-typology of gender in Hindu Kush Indo-Aryan, including a characterization of these systems in terms of their general complexity. Considering the relatively close genealogical ties, the languages display a number of unexpected and significant differences. While the inherited sex-based gender system is clearly preserved in most of the languages, and perhaps even strengthened in some, it is curiously missing altogether in others (such as in Kalasha and Khovar) or seems to be subject to considerable erosion (e.g. in Dameli). That the languages of the latter kind are all found at the northwestern outskirts of the Indo-Aryan world suggests non-trivial interaction with neighbouring languages without gender or with markedly different assignment systems. In terms of complexity, the southwestern-most corner of the region stands out; here we find a few languages (primarily belonging to the Pashai group) that combine inherited sex-based gender differentiation with animacy-related distinctions resulting in highly complex agreement patterns. The findings are discussed in the light of earlier observations of linguistic areality or substratal influence in the region, involving Indo-Aryan, Iranian, Nuristani, Tibeto-Burman, Turkic languages and Burushaski. The present study draws from the analysis of earlier publications as well as from entirely novel field data.

Keywords: Afghanistan, animacy, complexity, Dardic, gender pervasiveness, Indo-Aryan, Kashmir, non-sex-based gender, Pakistan, sex-based gender.



1 Introduction

At the very northern fringe of the Indo-Aryan world (approximately what lies north of the 34th parallel) we find a group of languages that historically and culturally are somewhat outside the sphere of the main Indo-Aryan languages of the subcontinent (Masica 1991: 20–21). Geographically, this group is wedged in between Iranian on its western side and Tibeto-Burman on its eastern side, and the distance to the Turkic belt of Central Asia is negligible at its farthest extension, even if it is not immediately adjacent. This extremely mountainous and multilingual region (see Figure 1), lies where the territories of Afghanistan, Pakistan and India-administered Kashmir meet. Henceforth, I will refer to this region as the Hindu Kush.¹ Apart from the languages and genera already mentioned, this region is also home to Nuristani – a third, but numerically small, branch of Indo-Iranian (Strand 1973: 297–298) – and to the isolate Burushaski.

The languages in question have been subject to a great deal of debate as to whether they are truly Indo-Aryan, constitute a genealogical unit of their own, or represent (perhaps along with the Nuristani languages) a transitional group between Indo-Aryan and Iranian. A term frequently used collectively for these languages is “Dardic”. However, few modern linguists use this term as anything else than a convenient umbrella term for a group of languages that are characterized – but not equally so – by a few salient retentions from previous stages of Indo-Aryan (Morgenstierne 1974: 3), but also have some contact-related developments in common (Bashir 2003: 821–822). Contact in that case includes mutual contact between the various Indo-Aryan linguistic communities as well as significant contact with adjacent communities belonging to other genera (Liljegren 2017). This non-committal line is also taken here regarding this grouping, but in order to avoid a stronger interpretation of “Dardic” than warranted, the term is abandoned in favour of Hindu Kush Indo-Aryan (HKIA) (Liljegren 2014: 135; Heegård Petersen 2015: 23), again without any claim of classificatory significance in the traditional sense. While the region for quite some time has been identified as particularly interesting in terms of areality and language contact (Emeneau 1965; Skalmowski 1985; Masica 1991: 43; Masica 2001: 259), and a number of features have been suggested as characteristic (Bashir 1988: 392–420; Bashir 1996; Bashir 2003: 821–823; Èdel’man 1980; Èdel’man 1983: 35–59; Fussman 1972: 389–399; Tikkanen 1999; 2008; Baart 2014; Toporov 1970), relatively little detailed and systematic areal-linguistic research has been carried out so far.

¹Strictly speaking, this region only partly overlaps with the Hindu Kush mountain range, while also overlapping with the Karakorum and the westernmost extension of the Himalayas.

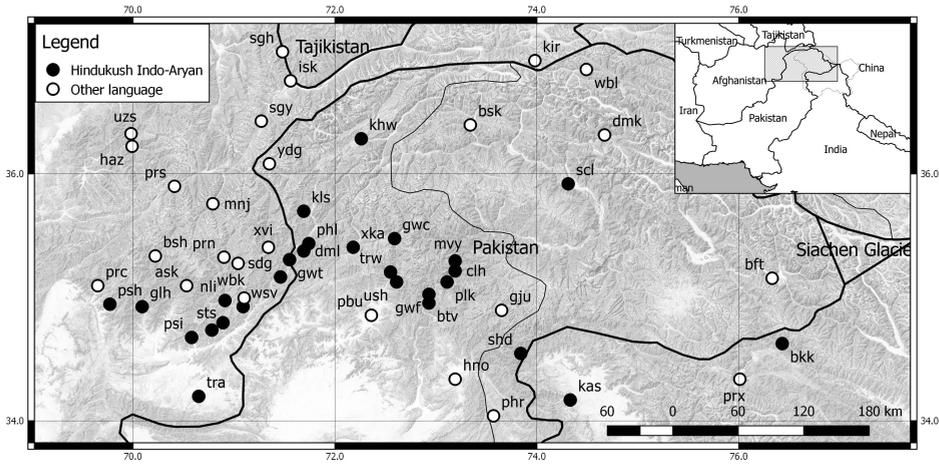


Figure 1: The Hindu Kush-Karakoram region with languages plotted (see Table 1 for an explanation of the 3-letter codes)

Regarding the ancestral nominal system, evidenced in Old Indo-Aryan as well as in Middle Indo-Aryan, it encompassed three gender values: masculine, feminine and neuter. In the Indo-Aryan world in general, these three values are only preserved in the modern languages in the southern part of the subcontinent, whereas a simplified two-value system (masculine vs. feminine, mainly as a result of neuter collapsing with masculine) dominates the large central and western parts. Such distinctions have altogether vanished in the northeast (Masica 1991: 217–223). The somewhat unexpected distribution and display of grammatical gender in the languages at the northern and western frontier of Indo-Aryan (viz. the Hindu Kush) was pointed out by Emeneau (1965: 68–71) half a century ago, but apart from Morgenstierne’s (1950: 19–20) tabulation, no systematic attempt has to my knowledge been made to account for gender distribution and manifestation across HKIA. This study tries to rectify that by showing the results of a survey of the following gender-related features – partly inspired by a number of contributions to the *World atlas of language structures (WALS)* – for each HKIA language for which there is data:

- The presence and number of gender categories (as evidenced by agreement patterns), and their basis, whether sex-based or non-sex-based (Corbett 2013a,b).
- The pervasiveness of gender, i.e. how gender is manifested in each language system in terms of (the types and numbers of) indexed domains.

- The assignment criteria at work: whether semantic or semantic-formal (Corbett 2013c).
- The presence and manifestation of pronominal gender (Siewierska 2013).

In the process of discussing and summarising these results, particularly in terms of the relative complexity of these systems, and in the light of areal patterning, a micro-typology of gender in HKIA emerges:

- The inherited sex-based system is largely preserved, but has disappeared in two of the languages at the Northwestern fringe of the Hindu Kush and is possibly eroding in a few other languages spoken in the same part of the region.
- An animacy-based system (almost exclusively marked on copulas or copula-based verbal categories) characterizes a number of the western-most languages of the region. In some cases it co-exists with a sex-based system; in others it occurs instead of a sex-based system or has contributed to a restructuring of the system as a whole.
- Gender is deeply entrenched (reflected in more target domains) in the East, i.e. in the languages spoken in areas contiguous with the main Indo-Aryan belt, whereas such pervasiveness is fading out toward the West.
- The results also suggest a weaker tendency toward semantic transparency in the gender systems in the North and a reinforcement of formal assignment, along with object agreement, in the South.

2 Hindu Kush Indo-Aryan and other languages in the region

Today, there are 28 distinct HKIA languages, i.e. languages identified as “Dardic” by the language catalogue Ethnologue (Eberhard et al. 2019), spoken in the region, the great majority of them on Pakistani soil or in areas of Kashmir now under Pakistani control. At least six clusters of related languages can be identified, mainly going with Bashir (2003: 824–825) and the classification used in Glottolog (Hammarström et al. 2018), although the definitive placement of a few of the individual languages is still pending (Dameli, Tirahi and Wotapuri-Katarqalai). All HKIA languages are presented in Table 1, roughly according to their geographical distribution, from west to east in a crescent-like fashion (see Figure 1). No

attempt has been made here to represent relatedness below the level of these six groupings.

Some of these groupings are tighter, i.e. internally less diverse, than others. This is one reason why they sometimes are treated as single languages with a number of dialects rather than as groupings of separate languages. That especially applies to Kashmiri, Shina and Pashai. The relatedness between the two Chitral group languages, Khowar and Kalasha, is also apparent from a number of features that single these two out from the rest of HKIA. The latter two were assumed by Morgenstierne (1932: 51) to represent the first wave of Indo-Aryan settlers moving in from the lowlands in the South.

If we, for the sake of simplicity, define the Hindu Kush region as the window between the longitudes 34 and 37 N and the latitudes 69 and 77 E, another 25 languages are spoken here. At least four other languages (or continua), traditionally described as belonging to sub-branches of Indo-Aryan with their geographical centres outside of the Hindu Kush region, are also found in the Hindu Kush region, or their geographical extension overlaps to a considerable extent with it: Hindko [hno], Pahari-Pothwari [phr], Gojri [gju] and Domaaki [dmk]. Hindko and Pahari-Pothwari are essentially part of a Punjabi macro-language extended far beyond the region, and as such they represent the closest main Indo-Aryan neighbour of HKIA. Gojri is the language of nomadic or semi-nomadic Gujurs, spoken in pockets throughout the region and beyond. The closest linguistic relatives of Rajasthani Indo-Aryan Gojri is, however, to be found at a considerable distance from the present region, deep into the main belt of Indo-Aryan. The closest relatives of Domaaki are likewise to be found in the plains of North India. Domaaki, however, is interesting from an areal point of view; as the language of a small enclave of musicians and blacksmiths surrounded by locally dominant speaker groups of Shina and Burushaski, it has during its 200–300 years in the area acquired a number of features typical of HKIA (Weinreich 2011: 165–166).

A number of the surrounding languages in the West are Iranian. Pashto [pbu] and Dari [prs], the two representing two completely different branches of Iranian, are both important lingua francas in parts of the region and well beyond. Dari is essentially the standard or literary type of Eastern Persian used in Afghanistan, while various names occur in reference to regional or local varieties, such as Tajik in north-eastern Afghanistan and neighbouring Tajikistan. Some of those may very well be considered languages in their own rights, e.g. Hazaragi [haz]. Most of the other Iranian languages (all very distantly related to either Pashto or Dari) are relatively minor, with a local scope only; in Afghanistan, Parachi [prc], Munji [mnj], Sanglechi [sgy], Ishkashimi [isk] and Shughni [sgh]; in Pakistan, Yidgha [ydg], basically a dialect of the same language as Munji; in Pakistan and

Table 1: Hindu Kush Indo-Aryan languages (with 3-letter ISO codes and the areas and countries where they are spoken), arranged in sub-groupings

Group	Language	code	Area (Country)
Pashai	Northwest Pashai	[glh]	Kabul, Kapisa, Konar, Laghman, Nurestan (Afg)
	Southwest Pashai	[psh]	Kabul, Kapisa (Afg)
	Southeast Pashai	[psi]	Nangarhar, Laghman (Afg)
	Northeast Pashai	[aee]	Konar, Nangarhar (Afg)
Kunar	Shumashti	[sts]	Konar (Afg)
	Grangali	[nli]	Konar, Nangarhar (Afg)
	Gawarbati	[gwt]	Konar (Afg), Chitral (Pak)
	Dameli	[dml]	Chitral (Pak)
Chitral	Kalasha	[kls]	Chitral (Pak)
	Khowar	[khw]	Chitral, Gilgit-Baltistan (Pak)
Kohistani	Tirahi	[tra]	Nangarhar (Afg)
	Wotapuri-Katarqalai	[wsv]	Nurestan (Afg)
	Gawri (Kalami)	[gwc]	Upper Dir, Swat (Pak)
	Torwali	[trw]	Swat (Pak)
	Indus Kohistani	[mvy]	Kohistan (Pak)
	Gowro	[gwf]	Kohistan (Pak)
	Chilisso	[clh]	Kohistan (Pak)
	Bateri	[btv]	Kohistan (Pak)
Shina	Mankiyali	[nlm]	Mansehra (Pak)
	Sawi	[sdg]	Konar (Afg)
	Palula	[phl]	Chitral (Pak)
	Kalkoti	[xka]	Upper Dir (Pak)
	Ushojo	[ush]	Swat (Pak)
	Kohistani Shina	[plk]	Kohistan (Pak)
	Kundal Shahi	[shd]	Jammu & Kashmir (Pak)
	Shina (Gilgiti)	[scl]	Gilgit-Baltistan (Pak), Jammu & Kashmir (Ind)
Kashmiri	Brokskat	[bkk]	Jammu & Kashmir (Ind)
	Standard Kashmiri	[kas]	Jammu & Kashmir (Ind), Jammu & Kashmir (Pak)

Afghanistan as well as in adjacent areas of Tajikistan and China, Wakhi [wbl] is spoken.

All of the five to six Nuristani languages are spoken in a geographically confined area in Afghanistan's Nurestan Province, close to the Pakistan border (with some spill-over into adjacent Chitral): Kati [bsh], Kamviri [xvi] (more correctly a dialect rather than a separate language from the aforementioned), Waigali [wbk], Ashkun [ask], Tregami [trm] and Prasun [prn]. Two Turkic languages are spoken at the northern periphery of the region: Uzbek [uzs] and Kirghiz [kir]; and in the East two with each other closely related Tibeto-Burman languages are found: Balti [bft] and Purik [prx]. The already-mentioned language isolate Burushaski is spoken in the extreme North of Pakistan's Gilgit-Baltistan region.

3 Sample and data

The sparsity of data points in large-scale typological enterprises such as *WALS* stresses the need for different selectional criteria when it comes to areal-typological or micro-typological studies. For instance, three of the *WALS* features (30A, 31A, 32A) that deal with gender include in their 257-language sample only five of the languages spoken in the Hindu Kush (Burushaski, Kashmiri, Kirghiz, Pashto and Uzbek), and of them only one (Kashmiri) is a HKIA language (Corbett 2013a,b,c). For the feature surveying pronominal gender (44A), the corresponding figures are 2 (Burushaski and Kashmiri) and 1 (Kashmiri), respectively, in a world-wide 378-language sample (Siewierska 2013).

It was therefore the aim of this survey to draw data from as many as possible of the 28 above-mentioned HKIA languages, rather than trying to identify and justify a smaller sample. This posed some challenges, as the quality and amount of documentation vary greatly from language to language. However, by combining available published descriptions with my own field data from a variety of languages in the region, it has been possible to find out which are the main characteristics and values (as presented in §1) for as many as 25 of them. I saw a definite need to exclude Gowro, Chilisso and Mankiyali due to lack of adequate data, but this should probably not distort the overall picture in any significant way, since the preliminary analysis shows that at least Gowro and Chilisso are relatively closely linked to Indus Kohistani (Bashir 2003: 874). The addition of unpublished field data was particularly important concerning the under-researched languages Bateri, Kalkoti and Ushojo. In Table 2, the sources of information for each language are specified.

Table 2: Data sources for Hindu Kush Indo-Aryan

Language	Sources
Northwest Pashai	(Morgenstierne 1967: 143–203); own data
Southwest Pashai	(Morgenstierne 1967: 45–142)
Southeast Pashai	(Morgenstierne 1967: 251–297; Lehr 2014); own data
Northeast Pashai	(Morgenstierne 1967: 205–249); own data
Shumashti	(Morgenstierne 1945)
Grangali	(Bashir 2003: 837–839; Grjunberg 1971)
Gawarbati	(Morgenstierne 1950); own data
Dameli	(Morgenstierne 1942; Perder 2013); own data
Kalasha	(Heegård Petersen 2015: 35–49; Bashir 1988); own data
Khowar	(Bashir 2003: 844–849); own data
Tirahi	(Morgenstierne 1934b; Grierson 1927: 265–327)
Wotapuri-Katarqalai	(Buddruss 1960)
Gawri (Kalami)	(Baart 1997; 1999); own data
Torwali	(Lunsford 2001; Bashir 2003: 864–869; Grierson 1929); own data
Indus Kohistani	(Hallberg & Hallberg 1999; Bashir 2003: 874–877; Lubberger 2014); own data
Bateri	(Hallberg & O’Leary 1992: 207–225, 249–251); own data
Sawi	(Buddruss 1967; Liljegren 2009: 43–48); own data
Palula	(Liljegren 2016); own data
Kalkoti	(Liljegren 2009: 43–48; Liljegren 2013); own data
Ushojo	(Decker 1992); own data
Kohistani Shina	(Schmidt & Kohistani 2008); own data
Kundal Shahi	(Rehman & Baart 2005); own data
Shina (Gilgiti)	(Bailey 1924; Degener 2008: 13–65; Radloff & Shakil 1998: 183–192); own data
Brokskat	(Ramaswami 1982; Sharma 1998)
Standard Kashmiri	(Koul 2003; Verbeke 2013: 175–211); own data

4 Gender Categories and their basis

The first question to address is whether gender a distinctive feature; and, if it is, also how many genders there are in the language. Here I align myself with the view that membership in a particular gender category in contrast with one or

more other such categories in the language in question is inherent to a noun but has to be evidenced by grammatical contrasts outside the noun itself, for instance in the form of adjectival or verbal agreement (Corbett 2014: 89–90; Hockett 1958: 231–233; Greenberg 1978: 50). Another relevant question is whether the gender system is based on, or primarily linked to, biological sex, or to something other than sex. Surveying the languages in our sample, we find (Table 3) that all of them display gender distinctions, one way or the other, with the possible exception of some dialects of NW Pashai.²

As can also be seen in Table 3, the basis for such distinctions is not the same for all of the languages. In the great majority of the languages (23 out of 25), the gender system, as it is mirrored in agreement, is clearly sex-based, having (at least) a two-way, female vs. male, differentiation at its core (as in many other Indo-Aryan languages in general). This is seen in example (1) from Ushojo, where ‘boy’ in (a) triggers masculine verb agreement, and ‘girl’ in (b) triggers feminine agreement. This masculine–feminine differentiation also extends into the inanimate realm: ‘wind’, in (c), is assigned feminine gender, and ‘coldness’, in (d), is assigned masculine gender.

(1) Ushojo (Own data)

- a. *ek phóó asíl-u, se seekel-aá yáa áal-u*
 one boy(M) be.PST-M.SG 3SG.NOM bicycle-LOC going come.PFV-M.SG
 ‘There was a boy, he came riding on a bicycle.’
 (USH-PearStoryAH:001)
- b. *ek phuí ... seekal-aá yáa mušiin tarapayá áal-i*
 one girl(F) bicycle-LOC going to.near in.direction come.PFV-F.SG
 ‘A girl... came in his direction, riding on a bicycle.’
 (USH-PearStoryAH:012)
- c. *axeér oóš ćóku bíl-i*
 finally wind(F) quiet become.PFV-F.SG
 ‘Finally the wind gave up.’ (USH-NorthwindAH:007)
- d. *maáti šídal bíl-u*
 1SG.DAT coldness(M) become.PFV-M.SG
 ‘I feel cold [lit. Coldness came to me].’ (USH-ValQuestAH:060)

²The preliminary analysis of my own data, from three NW Pashai locations (Sanjan, Alasai and Alishang) indicates the overall presence of sex-based adjectival gender agreement, whereas clear evidence of animacy-based differentiation is lacking in these particular varieties. While those findings have guided the present treatment, Morgenstierne’s (1967: 150–151, 173–176) study suggests a great deal of dialectal variation within NW Pashai as far as the presence/absence of both sex-based and animacy-based gender are concerned.

Table 3: The presence of gender (sex-based, non-sex-based) in Hindu Kush Indo-Aryan

Language	Number of genders	Sex-based gender	Non-sex-based gender
Southwest Pashai	4	✓	✓
Southeast Pashai	4	✓	✓
Northeast Pashai	4	✓	✓
Shumashti	3–4	✓	✓
Dameli	3	✓	✓
Kalasha	2		✓
Khowar	2		✓
Northwest Pashai	2	✓	
Grangali	2	✓	
Gawarbati	2	✓	
Tirahi	2	✓	
Wotapuri-Katarqalai	2	✓	
Gawri (Kalami)	2	✓	
Torwali	2	✓	
Indus Kohistani	2	✓	
Bateri	2	✓	
Sawi	2	✓	
Palula	2	✓	
Kalkoti	2	✓	
Ushojo	2	✓	
Kohistani Shina	2	✓	
Kundal Shahi	2	✓	
Shina (Gilgiti)	2	✓	
Brokskat	2	✓	
Standard Kashmiri	2	✓	

In two of the languages, Khowar and Kalasha, both belonging to the Chitral group, sex-based differentiation is entirely lacking. However, in both languages we find a two-way differentiation based on animacy, where animate nouns (including humans and higher non-human animals) are treated differently from inanimate nouns by some agreement targets. For instance, the present actual copula verb used in locational predication in Khowar has different third person

singular and plural agreement forms for animate and inanimate, respectively. That is illustrated in example (2) with the two plural forms. (The corresponding singular forms are *asúr* and *šer*.) The copula, in its various forms, is also used as an auxiliary participating in some tense-aspect formations.

(2) Khowar (Own data)

- a. *dúr-a roy asúni*
 house-LOC people(AN) be.PRS.ACT.3.AN.PL
 ‘There are people in the house.’ (KHW-PredFA:011)
- b. *kitáb ma dúr-a šéni*
 book(INAN) 1SG.GEN house-LOC be.PRS.ACT.3.INAN.PL
 ‘The books are in my house.’ (KHW-PredFA:009)

A few of the dialects of NW Pashai may also lack sex-based gender distinctions (Morgenstierne 1967: 150–151); in those cases we do not have conclusive information on the presence of animacy distinctions. In another few languages – in Dameli and Shumashti (both Kunar languages), and in several of the Pashai varieties – animacy differentiation occurs, not instead of but in addition to sex-based differentiation. However, there are reasons to regard these as two separate features (with two values each) that affect different parts (or sub-domains) of the language system, a situation that Dahl (2000: 581–582) refers to as “parallel combinations of gender distinctions”. The feminine–masculine and animate–inanimate distinctions only marginally make use of the same agreement target. In Dameli, this happens in non-verbal predication, which results in a three-way differentiation at the most: animate masculine vs. animate feminine vs. inanimate, as shown in example (3). Apart from the specific domain of non-verbal predication in Dameli, a two-way masculine vs. feminine distinction is upheld in most other parts of the grammar. It is not unlikely that a similar situation holds in Shumashti, although the data available is too scanty to draw any firm conclusions.

(3) Dameli (Own data)

- a. *i mač mruy thaa*
 PROX.AN man(M) hunter be.PRS.3M.SG
 ‘This man is a hunter.’ (DML-ValQuestHM:070)
- b. *poši koki thui*
 cat(F) asleep be.PRS.3F.SG
 ‘The cat is asleep.’ (DML-ErgSurvHM:013)

- c. *bum* *šukisan daru*
ground dry be.PRS.3SG.INAN
'The ground is dry.' (DML-ValQuestHM:068)

In Pashai (at least in SE, SW and NE), animacy and sex-based gender agreement do co-occur in one and the same clause and with one and the same referent, see the SE Pashai example in (12). That results in a four-way distinction (masculine/animate, masculine/inanimate, feminine/animate vs. feminine/inanimate).

This naturally leads over to the topic of our next section: agreement targets and the general pervasiveness of gender.

5 Agreement targets and the pervasiveness of gender

In line with the view that grammatical gender and the number of gender categories is evidenced in agreement patterns, I will use the number of agreement targets as a (somewhat crude) measure of what I call gender pervasiveness (Table 4). Here, it will be necessary to look at sex-based distinctions (masculine vs. feminine) separate from non-sex-based distinctions (animate vs. inanimate). This is not to say that they need to be regarded as two entirely distinct phenomena, but rather to underscore a general observation that sex and animacy in most cases operate at different levels and affect separate (and only peripherally overlapping) subsystems or parts of the language systems under investigation. It will be possible to make some overall generalizations along relatedness lines, although I will also point out some important variation within lower-level genealogical groupings, and for some of the languages I will also elaborate further on the relative pervasiveness within the target categories. While pronominal gender is indicated in Table 4 it will not be discussed until §7. (A tick-mark within parentheses indicates that agreement is restricted to copula verbs or copula-derived auxiliaries; a question mark after a tick-mark indicates a possible but non-conclusive presence of a gender target.)

Starting with Kashmiri, gender is very pervasive throughout the system, including adjectives, adnominal demonstratives and possessive phrases in nominal modification; verbs also show gender agreement. Person, number and gender are often conflated in a complex manner, and distinctions are, at least partly, expressed non-linearly, i.e. by vowel modification or palatalization. Example (4) demonstrates agreement in adjectival inflection; as can be seen in this example, gender distinctions are upheld in the singular as well as in the plural.

Table 4: Agreement targets for gender (sex-based, animacy-based) in Hindu Kush Indo-Aryan

Language	Gender targets									
	Sex-based					Animacy-based				
	<i>verb</i>	<i>adj</i>	<i>dem</i>	<i>poss</i>	<i>pron</i>	<i>verb</i>	<i>adj</i>	<i>dem</i>	<i>poss</i>	<i>pron</i>
Standard Kashmiri	✓	✓	✓	✓	✓					
Shina (Gilgiti)	✓	✓	✓		✓					
Brokskat	✓	✓	✓		✓					
Kundal Shahi	✓	✓								
Kohistani Shina	✓	✓			✓					
Ushojo	✓	✓			✓?					
Palula	✓	✓	✓		✓					
Kalkoti	✓	✓								
Sawi	✓	✓								
Indus Kohistani	✓	✓		✓						
Gawri (Kalami)	✓	✓		✓						✓
Torwali	✓	✓						✓?		
Bateri	✓	✓								
Tirahi	✓	✓		✓						
Wotapuri-Katarqalai	✓	✓								
Gawarbati	✓	✓		✓						
Grangali		✓								
Shumashti	✓	✓				(✓)				
Dameli	✓	✓		✓		(✓)		✓		✓
Southwest Pashai	✓	✓				(✓)				
Southeast Pashai	✓	✓				(✓)				
Northeast Pashai	✓	✓				(✓)				
Northwest Pashai	✓	✓				(✓)?				
Kalasha						(✓)				
Khowar						(✓)				

(4) Standard Kashmiri (Koul 2003: 915)

- a. *n'uul* *kooṭh*
blue.M.SG coat(M)
'a blue coat'
- b. *niil* *kooṭh*
blue.M.PL coat(M)
'blue coats'
- c. *niiṣ* *kəmiiz*
blue.F shirt(F)
'a blue shirt'
- d. *niiṣ-i* *kəmiiz-i*
blue.F-PL shirt(F)-PL
'blue shirts'

In Kashmiri, gender agreement is part of the paradigm of all major verbal categories apart from the future tense. As in Indo-Aryan in general, gender differentiation became part of the verbal paradigm as participial forms were introduced and proliferated as carriers of core tense-aspect categories during the Middle Indo-Aryan stage (Pirejko 1979: 481–482; Klaiman 1987: 61–64). In a development associated with that, the transitive subject ended up non-nominatively coded while the verb (reinterpreted as part of a finite verb construction) agreed with the nominatively coded direct object (Masica 1991: 341–346). This was the establishment of a split ergative system still in existence in various versions in many Indo-Aryan languages, including many HKIA languages (Liljegren 2014).

Gender is generally also very pervasive in the Shina group (Shina (Gilgiti) to Sawi in Table 4), although it varies between the individual languages. None of them manifest gender agreement in possessive modification. In Gilgiti Shina, Brokskat and Palula, adjectives, adnominal demonstratives and verbs are targets of gender agreement, whereas it is limited to adjectives and verbs in the rest of the languages classified as Shina. The pervasiveness of gender within the verbal paradigms varies to a great extent, and is partly related to considerable differences in verbal alignment patterns. Gilgiti Shina and Kohistani Shina, the two varieties that together constitute “Shina proper”, are characterized by consistent accusative verbal alignment in combination with ergative case marking (see example 5). A number of Shina enclaves farther to the West instead show an aspectual split between ergatively aligned clauses in the perfective (see example 6), in which the verb agrees in gender and number with the direct object, and

accusatively aligned clauses in the non-perfective. In Shina proper, gender agreement is largely conflated with person-marking, whereas in the Western varieties, gender- and number-inflected verb forms (based on participles) have largely replaced person-inflected forms.

- (5) Gilgiti Shina (Own data)
ro baál-se khirkí phuṭ-eég-u
 REM.M.SG boy(M)-ERG window(F) break-PFV-3M.SG
 ‘The boy broke the window.’ (SCL-ValQuestAH:025)
- (6) Palula (Own data)
phoo-á darúri phooṭéel-i
 boy(M)-OBL window(F) break.PFV-F
 ‘The boy broke the window.’ (PHL-ValQuestNH:025)

In addition to the categories surveyed in this section, gender agreement in Palula is also extended or copied to e.g. adjuncts in predicatively used adverbial phrases. In (7), the scalar modifier *bíid-* ‘much’ agrees with the feminine noun head of the subject.

- (7) Palula (Own data)
asíi iskuúl bi asaám the bíid-i dhúura hín-i
 1PL.GEN school(F) also 1PL.ACC to much-F distant be.PRS-F
 ‘Our school is also very far away for us.’ (PHL-OUR:016)

In none of the Kohistani languages are adnominal demonstratives targets of gender marking. On the other hand, gender differentiation is part of possessive modification in at least two of the languages. Examples are provided from Indus Kohistani in (8).

- (8) Indus Kohistani (Lubberger 2014: 62, 82)
- a. *zãĩ bakàr*
 1PL.POSS.F goat(F)
 ‘our goat’
- b. *zãã baá*
 1PL.POSS.M house(M)
 ‘our house’

Manifestation of gender in the verbal paradigm is not necessarily much less pervasive than in the languages of the Shina group, but it tends to be more challenging in terms of description. It is to a greater extent non-segmental in Kohistani than in Shina. A case in point is the Kohistani language Gawri (a.k.a. Kalam Kohistani) which historically has lost most of its gender-specific endings (both on the nouns themselves and on their agreement targets) as well as its suffixing plural or case-marking. It has, however, preserved the distinctions themselves up to a point, in the form of vowel modifications and/or distinct tonal patterns, as can be seen in example (9).

(9) Gawri

- a. Inflection of nouns (H=high tone, LH=low to high, HL=high to low, L=low) (Baart 1999: 36)

SG.NOM		PL.NOM/SG.OBL/PL.OBL		
<i>šak</i>	H	<i>šäk</i>	HL	‘piece of wood’ (M)
<i>dätär</i>	LH	<i>dätär</i>	L	‘cooking frame’ (M)
<i>naar</i>	H	<i>neer</i>	HL	‘root’ (F)
<i>därin</i>	LH	<i>därin</i>	L	‘ground’ (F)

- b. Gender and number agreement on adjectives (Baart 1999: 19; p.c. Muhammad Zaman Sagar)

<i>raan</i>	<i>poo</i>	<i>rään</i>	<i>lukuṭor</i>
good.M.SG	boy	good.M.PL	boy.PL
‘good boy’		‘good boys/children’	
<i>reen</i>	<i>bire</i>	<i>reen</i>	<i>likiṭeer</i>
good.F	girl	good.F	girl.PL
‘good girl’		‘good girls’	

- c. Gender and number agreement on verbs (conflated with aspect marking) (Baart 1999: 19; p.c. Muhammad Zaman Sagar)

<i>poo</i>	<i>bäč-an-t</i>	<i>lukuṭor</i>	<i>bäč-än-t</i>
boy	go-IPFV.M.SG-PRS	boy.PL	go-IPFV.M.PL-PRS
‘The boy is going.’		‘The boys are going.’	
<i>bire</i>	<i>bäč-en-t</i>	<i>likiṭeer</i>	<i>bäč-en-t</i>
girl	go-IPFV.F-PRS	girl.PL	go-IPFV.F-PRS
‘The girl is going.’		‘The girls are going.’	

Masculine and feminine agreement forms are clearly distinguished in all of the major tense-aspect categories in Gawri and Torwali, either inflectionally or by vowel alternation. However, a high degree of levelling seems to have taken place in Indus Kohistani; and most likely in Bateri too. In Indus Kohistani and Bateri, transitive verbs (or at least most of them) are invariant in the simple past (i.e., there is no agreement with any of the arguments). In addition, the application of the ergative marking of the transitive subject is variable. In Bateri, a nominative vs. ergative contrast is possibly missing altogether with full nouns, as evidenced in example (10).

(10) Bateri (Own data)

- a. *yak muuʃ as-uu*
 one man(M) be.PST-M.SG
 ‘There was a man.’ (BTV-PearStoryMB:001)
- b. *muuʃ ɖaʌŋ sand-id*
 man(M) stick make-PST
 ‘The man made a stick.’ (BTV-ValQuestMB:085)

In the Kunar group, the targets of sex-based gender differentiation are adjectives, verbs and, in the case of Gawarbatı and Dameli, possessive modifiers. The sentences in (11) illustrate some of those agreement patterns in Gawarbatı: possessive and verbal (copula) agreement with a feminine noun in (a), possessive agreement with a masculine noun in (b), and adjectival and verbal agreement with a feminine noun in (c). Verbal agreement that takes gender into account is rather restricted in Gawarbatı: it occurs only with intransitive verbs, and for third person singular. As seen in (b), the transitive subject in the past (perfective) is ergatively marked, while verbal agreement is accusatively aligned.

(11) Gawarbatı (Own data)

- a. *woi ʃekura-an-i awaaz then-i*
 PROX.SG boy(M)-POSS-F voice(F) be.PRS-3F.SG
 ‘This is a boy’s voice.’ (GWT-NPhonNU:071-4)
- b. *ʃekuri-e kitaab-an-a faʃaa daal-us*
 girl-ERG book(M)-POSS-M leaf(M) tear-PST.3SG
 ‘The girl tore the page from the book (lit. the book’s leaf).’
 (GWT-ValQuestAS:032)

- c. *pol-i tekuri hans-ui*
 small-F girl(F) laugh-PRS.3F.SG
 ‘The little girl laughed.’ (GWT-ValQuestAS:057)

As already mentioned in §4, an added distinction between animate and inanimate occurs in Dameli and Shumashti. While animacy influences lexical or constructional choices on various levels of Dameli, the only purely paradigmatic contrasts that depend on animacy values are those of the copula verb (Perder 2013: 121–125), as illustrated above in example (3), and of demonstratives. However, it is highly uncertain whether the inanimate copula is at all used as an auxiliary in verbal predication in any of the tense-aspect categories in Dameli. More interestingly, Perder (2013: 51–55) observes what seems to be an ongoing restructuring of the entire gender system, a point to which we shall return in the next section when discussing assignment criteria.

In Pashai, sex-based gender is again relatively pervasive, although limited in its manifestation to adjectives and verbal agreement. As in Dameli, there is an additional layer of animacy-based differentiation in the verbal paradigm. Lehr (2014: 255) describes (for SE Pashai) how the masculine vs. feminine distinction is upheld throughout the past and perfective parts of the verbal paradigm, a contrast that is present in first, second as well as in third person. The additional animate vs. inanimate distinction, on the other hand, is limited to the verbal system (2014: 256–257), occurring only in non-verbal predication and in the (participial-based) present perfect category. The three sentences in (12) are all examples of the present perfect: the main verb agrees in person with the subject, in sex-based gender with the object, and the auxiliary agrees in sex-based as well as non-sex-based gender and person with the object.

(12) SE Pashai (Lehr 2014: 290, 297)

- a. *pari-y kelaa kaṭ-ee=šeer-a ne-l-aw-aa-e*
 Pari(F)-OBL boy(M) cot-OBL=head-LOC sit-TRZ-STV.PTC-M-POSS.3SG
aaS
 be.AN.M.PRS.3
 ‘Pari has seated the boy on the cot.’
- b. *miy maada-y doa be ka-w-aa-e*
 DEM.SG.OBL woman(F)-OBL prayer(M) too do-STV.PTC-M-POSS.3SG
š-i
 be.INAN.PRS-3
 ‘This woman has made a prayer.’

- c. *mam pelek meez-ee=šeer-a ĵe-w-i-m*
 I cup(F) table(F)-OBL=on-LOC place-STV.PTC-F-POSS.1SG
š-i
 be.INAN.PRS-3
 ‘I have placed the cup on the table.’

Finally, both of the two Chitral group languages, Khowar and Kalasha, entirely lack any sex-based gender in their agreement patterns. Grammatical differentiation between animate and inanimate nouns is manifested, but only in the verbal paradigm. It occurs in those verbal categories that are constructed with a copula-based auxiliary, such as in the Kalasha example in (13): here, the animate as well as the inanimate forms occur, each along with the main verb ‘hit’. Kalasha expresses animate vs. inanimate differentiation in five of its nine main tense-aspect categories (Bashir 1988: 60–72), but because of its consistent accusative alignment with subject agreement (as compared to the pattern of direct object agreement in Pashai), the frequency of inanimate marking is in effect rather low. A similar situation holds for Khowar (Bashir 1988: 123–133). Thus, the centrality of the animacy contrasts that these tense-aspect systems allow for could in fact be questioned.

- (13) Kalasha (Heegård Petersen 2015: 250)
gheri tya-y a-aw=e, tasa ek bab-as
 again hit-PFV.PTC AUX.AN.ACT-3SG=when 3SG.REM.OBL a sister-OBL.SG
gulin-a tya-y š-iu.
 lap-LOC hit-PFV.PTC AUX.INAN-PRS/FUT.3SG
 ‘When he hit [the ball] again, it was hit into her sister’s lap.’

It seems that whereas sex-based gender generally is deeply entrenched in the languages that have it, and is clearly evidenced in many of the inflectional paradigms, the non-sex based type of gender differentiation that we saw examples of in a few of the languages is indexed in considerably fewer domains and is thus affecting, in each case, a rather limited domain of the language system. The question remains open as to whether those contrasts should be seen as instances of mere (lexical) co-occurrence restrictions, instead of truly grammatical contrasts. We may also regard the occurrence of animacy distinctions in these languages as examples of overdifferentiated targets (Corbett 1991: 168–169), probably more so in the languages with parallel combination of distinctions (Dameli, Shumashti and the Pashai varieties) than in the languages with non-sex based distinctions only (Khowar and Kalasha).

6 Assignment criteria

Determining the assignment criteria for gender in individual languages is a less straightforward matter, even for much more well-known languages with large corpora available. For this reason, the following is meant only as a very tentative assessment, and the results of the assessment is therefore not reduced to a simple table representation. Although the focus will be on the languages for which there is a more comprehensive description in place, it remains beyond the present investigation to lay down precise assignment rules for any of these.

For all the languages that have a sex-based two-term system, i.e. the large majority of HKIA, gender is with high consistency assigned according to natural sex as far as nouns denoting humans and other higher animates, particularly domestic animals, are concerned. Below this cut-off point between higher and lower animates (or possibly between animates and inanimates), semantics is a much less reliable indicator, although some outstanding semantic properties beside sex will be mentioned in connection with the discussion of individual languages. But it also seems clear that formal (i.e. non-semantic) criteria do play a non-trivial role in some of the languages in assigning inanimate and lower animate nouns to the masculine and feminine classes, respectively. In a historical perspective, the present two-term systems is the result of the masculine and the neuter categories of the former three-gender system having merged (Masica 1991: 221). This, however, is not mirrored in a totally unbalanced feminine to masculine ratio, as might be expected. Instead, there is a relatively even distribution; in Palula, there were 58 per cent masculine and 42 per cent feminine nouns in a database comprising about 1,300 nouns, and in a Gawri list of 2,000 nouns, the percentages were 60 and 40, respectively (Baart 1999: 82), and inanimates and lower animates of both genders are numerous.

Although there are plenty of examples in Kashmiri of feminine nouns derived from masculine nouns by means of various semi-regular phonological processes (such as stem vowel diphthongization or fronting) these correlations between characteristic phonological features and one or the other gender are mainly restricted to higher animates: *guur* 'milkman' vs. *guuər* 'milkwoman'; *koʃ* 'boy' vs. *kəʃ* 'girl'; *kəkur* 'rooster' vs. *kəkir* 'hen'; *mool* 'father' vs. *məj* 'mother'. However, the nominal inflectional patterns of the language (see Table 5) also predict gender to a large extent. Most non-nominative case forms, for instance, have endings that are typical for masculine vis-à-vis feminine nouns (with a great deal of syncretic *i* occurring in the paradigms of feminine nouns, contrasting with differentiating forms in the paradigms of masculine nouns), often accompanied by stem alternations (with vowel fronting or palatalization in the feminine forms).

Table 5: Sample Kashmiri nominal paradigm (Koul 2003: 909)

Case	'boy' M		'girl' F	
	SG	PL	SG	PL
NOM	<i>lɔɖki</i>	<i>lɔɖki</i>	<i>kuur</i>	<i>koori</i>
DAT	<i>lɔɖkas</i>	<i>lɔɖkan</i>	<i>koori</i>	<i>kooren</i>
ERG	<i>lɔɖkan</i>	<i>lɔɖkav</i>	<i>koori</i>	<i>koorev</i>
GEN	<i>lɔɖki</i>	<i>lɔɖkan</i>	<i>koori</i>	<i>kooren</i>

In the Shina group, many of the languages have sizeable subclasses of masculine and feminine nouns with gender-typical endings, mostly *o/u/a* with masculine nouns, and *i* with feminine nouns. But again, similar to what was noted regarding Kashmiri, there is a considerable overlap between nouns with such overt gender markers and biological sex. Brokskat, a Shina language which otherwise has few overt phonological characteristics related to one or the other gender, makes use of two Tibetan-derived suffixes, *-pa/-po* and *-ma/-mo* to indicate the sex of some higher animates (see Table 6). To what extent these suffixes are used with inherited vocabulary is not clear.

Table 6: Masculine–feminine higher animate pairs in Brokskat (Ramaswami 1982: 38–39; Sharma 1998: 56–58, 80)

Masculine		Feminine	
<i>rgəl-po</i>	'king'	<i>rgəl-mo</i>	'queen'
<i>bäg-pa</i>	'bridegroom'	<i>bäg-ma</i>	'bride'
<i>bya-po</i>	'rooster'	<i>bya-mo</i>	'hen'
<i>əbs</i>	'horse'	<i>əspi, rgun-ma</i>	'mare'
<i>byo</i>	'boy, son'	<i>mole</i>	'girl, daughter'
<i>dudo</i>	'grandfather'	<i>dede</i>	'grandmother'
<i>čhatalo</i>	'he-goat'	<i>aav</i>	'she-goat'
<i>laanto</i>	'bull'	<i>gooli</i>	'cow'

However, for many consonant-ending nouns below the threshold for sex-based assignment, i.e. between higher and lower animates, assignment seems to a large extent arbitrary in Shina languages. Although there are clearly discernible declensional classes in e.g. Kohistani Shina, Palula and Sawi, these are not in all cases directly mapped to one or the other gender. In Gilgiti Shina, a language

where declensional differences are less clearly identifiable, there are fewer formal clues to gender assignment, and in Brokskat, where there are few phonological clues and a relatively uniform inflectional pattern, the arbitrariness seems even more noticeable as far as nouns low on the animacy scale are concerned. It is in fact likely that gender assignment in these languages to a varying extent is an intricate interplay of overlapping semantic, morphological and phonological factors, not altogether different from what we find in e.g. German (Corbett 1991: 49).

Let us take Palula as an example in terms of such a complex interplay of different assignment criteria. Starting with nominal morphology (see Table 7), Palula has three major declensional classes, characterized by plural formation with *-a*, *-i* and *-m*, respectively. The *m*-declension consists exclusively of feminine nouns (all of which end with gender-typical *i* in their singular form), whereas *a*-declension consist to 79 per cent of masculine nouns, and the *i*-declension to 70 per cent of feminine nouns. In addition, there are two minor declensions (together representing 10–15 per cent of all nouns), both exclusively masculine.

Table 7: Palula noun declensions

Decl	SG NOM	SG OBL	PL NOM	PL OBL	Relative size	M/F
<i>a</i> -decl	<i>púustu</i> 'skin (M)'	<i>púust-a</i>	<i>púust-a</i>	<i>púust-am</i>	large (~50%)	79/21
<i>i</i> -decl	<i>baát</i> 'word (F)'	<i>beet-í</i>	<i>beet-í</i>	<i>beet-íim</i>	large (~25%)	30/70
<i>m</i> -decl	<i>ṭiki</i> 'bread (F)'	<i>ṭiki</i>	<i>ṭiki-m</i>	<i>ṭiki-m</i>	large (~13%)	0/100
<i>ee</i> -decl	<i>alučá</i> 'plum (M)'	<i>alučá</i>	<i>aluč-éé</i>	<i>aluč-éém</i>	small (~8%)	100/0
<i>aan</i> -decl	<i>ḍaakú</i> 'robber (M)'	<i>ḍaaku-á</i>	<i>ḍaaku-aán</i>	<i>ḍaaku-aanóom</i>	small (~5%)	100/0

However, the amount of arbitrariness within the two “gender-divided” declensions is further reduced by taking phonological clues into account (see Table 8). About a third of the nouns in the *a*-declension have for Palula gender-typical endings in their nominative singular forms (mainly masculine nouns in *u*, and feminine nouns in *ái*). A typical property of many *i*-declension consonant-ending nouns that are assigned feminine gender is that they have a second-mora accented *aá* which very often is subject to a process of umlaut (> *ee*) in its inflected

Table 8: Gender-typical phonological properties in Palula

Masculine			Feminine		
Cu#	<i>tómbu</i> ‘trunk’, <i>šúuru</i> ‘hole’, <i>rúulu</i> ‘tear’, <i>púustu</i> ‘skin’, <i>priiṅṣu</i> ‘flea’, <i>báabu</i> ‘father’, <i>báatru</i> ‘irrigation lock’, <i>bháaru</i> ‘load’	<i>a</i> -decl	Ci#	<i>šúri</i> ‘ladder’, <i>tíki</i> ‘bread’, <i>šišáki</i> ‘ogress’, <i>phéepi</i> ‘father’s sister’, <i>nóki</i> ‘beak’, <i>múti</i> ‘arm’, <i>lúuti</i> ‘ball of yarn’, <i>béjei</i> ‘heifer’	<i>m</i> - decl
Coó#	<i>rhoó</i> ‘song’, <i>phoó</i> ‘boy’, <i>paṅoó</i> ‘slipper’, <i>muuṣoó</i> ‘elbow’, <i>baḍiloó</i> ‘male descendant of Badil’, <i>haṅoó</i> ‘egg’	<i>ee</i> - decl, <i>a</i> -decl	Cíi#	<i>rhootašú</i> ‘morning’, <i>rhaíi</i> ‘footprint’, <i>phaaṭurii</i> ‘butterfly’, <i>ačhíi</i> ‘eye’, <i>balíi</i> ‘roof end’, <i>bíi</i> ‘seed’	<i>a</i> -decl
Cá#	<i>teeká</i> ‘contract’, <i>lambá</i> ‘flame’, <i>alaaqá</i> ‘area’, <i>alučá</i> ‘plum’, <i>çaṅzá</i> ‘torch’	<i>ee</i> - decl, <i>i</i> -decl	Cái#	<i>tookrái</i> ‘basket’, <i>puṭái</i> ‘piece of meat’, <i>mulái</i> ‘radish’, <i>bhraajái</i> ‘sister-in-law’	<i>a</i> -decl
Caá#	<i>saaraá</i> ‘wilderness’, <i>raajaá</i> ‘ruler’, <i>paalaá</i> ‘leaf’, <i>aaghaá</i> ‘sky’, <i>bhalaá</i> ‘evil spirit’, <i>čoolaá</i> ‘speech, style’	<i>i</i> -decl, <i>ee</i> - decl	CaáC#	<i>aašaár</i> ‘apricot’, <i>salaám</i> (pl. <i>saleemí</i>) ‘greeting’, <i>oombaár</i> (pl. <i>oombeerí</i>) ‘canal inlet’, <i>baát</i> (obl. <i>beetí</i>) ‘word’	<i>i</i> -decl

forms (with affixes involving *i*). This is also characteristic of a good number of loan words. This is not to say that there are no exceptions to these correlations between certain vocalic properties and one of the two genders, but they are indeed few.

Another sizeable group of *a*- and *i*-declension nouns (although partly overlapping with those having gender-typical phonological properties) are assigned

gender semantically. Primarily that is by biological sex for nouns referring to humans and higher non-human animates. Word pairs referring to male and female, respectively, which have a common lexical root are frequent (see Table 9), especially in the realm of kinship. For most higher animates, the masculine is the default, and for those that have a feminine counterpart, the latter is a marked form (often part of the *m*-declension and ending in *i*), i.e. the one used only when a specification of sex is called for. However, in a few cases, the reverse holds, e.g. with ‘fox’ and ‘cat’. The semantic relationship between masculine ‘goat kid’ and its feminine counterpart ‘goat (generic)’ is again different.

Table 9: Masculine–feminine higher animate pairs in Palula

Masculine		Feminine	
<i>jáanu</i>	person	<i>jéeni</i>	female person
<i>saaróonu</i>	woman’s sister’s husband	<i>saaréeni</i>	wife’s sister
<i>phóó</i>	boy	<i>phái</i>	girl
<i>móomu</i>	mother’s father	<i>méemi</i>	mother’s mother
<i>káaku</i>	older brother	<i>kéeki</i>	older sister
<i>khaamaád</i>	owner, husband	<i>khaaméedi</i>	female owner
<i>práaču</i>	guest	<i>préечи</i>	female guest
<i>phóopu</i>	father’s sister’s husband	<i>phéepi</i>	father’s sister
<i>kučúru</i>	dog	<i>kučúri</i>	female dog
<i>bačúuru</i>	young calf	<i>bačúuri</i>	young female calf
<i>karáaru</i>	leopard	<i>karéeri</i>	female leopard
<i>iñç</i>	bear	<i>iñçi</i>	she-bear
<i>luumóo</i>	male fox	<i>luumái</i>	fox (generic)
<i>púšu</i>	tom-cat	<i>púši</i>	cat (generic)
<i>kakóok</i>	chicken	<i>kakuéeki</i>	hen
<i>čhaál</i>	goat kid	<i>čhéeli</i>	goat (generic)

Apart from this relatively straightforward correlation between sex and grammatical gender, there is another (but obviously related) correlation, namely between relative size or power and gender, primarily applied to lower animates and inanimates (as exemplified in Table 10). In these cases, the derivation of feminine nouns could be described as a type of diminutive formation. The similarity in kind

is more approximate and less predictable than with the previously exemplified higher animate pairs.

Table 10: Masculine–feminine lower animate and inanimate pairs in Palula

Masculine		Feminine	
<i>phútu</i>	fly	<i>phúti</i>	mosquito
<i>khaláaru</i>	large leather bag, made from skin of a he-goat	<i>khaléeri</i>	small leather bag, made from skin of a she-goat
<i>şúuru</i>	hole	<i>şúuri</i>	cap
<i>angúru</i>	thumb, big toe	<i>angúri</i>	finger, toe
<i>açhibáaru</i>	eyebrow	<i>açhibéeri</i>	eyelashes
<i>angóor</i>	fire	<i>angeeri</i>	charcoal

Leaving Palula and the Shina languages for now, some of the languages of the Kohistani group also have overt phonological markers, similar to the ones in the Shina group. In Indus Kohistani, *i*-endings are associated with a group of feminine nouns, and in Bateri some masculine nouns end in *-o/-u* and some feminine nouns in *-a/-ã*. In both of these cases, however, that pattern is relatively restricted and perhaps primarily relevant for feminine nouns derived from masculine nouns denoting humans, particularly applied to male–female pairings in the kinship systems of these languages. Due to historical loss of final vowel segments, the corresponding correlations in Gawri and Torwali are often only preserved in stem vowel alternations and tonal contrasts, resulting from assimilation prior to apocope. In Gawri, there is a strong correlation between feminine gender and the vowel qualities [i] and [e], and a corresponding correlation between masculine gender and the qualities [a], [æ], [o], and [u].

In the Kunar languages, there are no obvious declensional differences (plurality is for instance normally left morphologically unmarked, and case marking has little allomorphy), and nouns that have gender-typical endings are relatively few (*a*-ending masculine nouns in Dameli, Gawarbati and Shumashti; *i*-ending feminine nouns in Dameli and Gawarbati; *i*-ending or *ik*-ending feminine nouns in Shumashti). Like in many of the other groups, nouns with these overt phonological “markers” often participate in masculine–feminine pairings where the latter term is derived from the former, which frequently applies to humans or domestic animals. Although needing a more systematic study, there is

evidence suggesting that Dameli is drifting away from formal-semantic gender assignment toward purely semantic gender assignment, as strict masculine vs. feminine gender assignment is becoming restricted to nouns above the cut-off point between higher and lower animates. This is for instance manifested in the native speaker inconsistency that Perder noted while eliciting the gender of inanimate nouns (2013: 54), along with an observed pattern of a default application of masculine gender agreement between verbs and inanimate subjects (2013: 111). Together with the already-mentioned observations regarding animacy-related distinctions, it seems like we are witnessing a development in Dameli from a partly formal assignment system with two sex-based grammatical genders to a system by which gender is assigned entirely along semantic lines. In most parts of the system there is a contrast between a feminine class consisting of female higher animate nouns and a masculine class with all the remaining nouns, and in a restricted part of the system (with the copula verb as target) there is a three-way contrast between higher animate males, higher animate females and the rest. The grammatical animate-inanimate distinction in Dameli is, as far as has been observed, altogether missing in Gawarbati, leaving it with a two-way distinction and with assignment principles along the same lines as described for many of the Kohistani and Shina languages. Although the scanty material available does not give us any firm evidence, the Shumashti copula forms that Morgenstierne (1945: 255) presents us with (*in-e* ‘is M’, *in-i* ‘is F’, *šuu-e* ‘(it) is’) implies an actual four-way differentiation, although we can only assume that a hypothetical inanimate feminine form (**šuu-i* ‘(it) is F’) simply is missing in the data.

The patterns observed for most parts of the other groupings can also be seen in Pashai. Here, too, there are certain endings associated with one or the other gender. In SE Pashai, for instance, *-i* or *-ek* is typical of feminine nouns and *-aa* of masculine. While the feminine *i*-ending is found with many inanimate nouns, there are many regular alternations involving gendered pairs where the masculine form with *-aa* contrasts with a feminine form with *-ek*. But again, there are numerous nouns that are either masculine or feminine that have none of these overt phonological markers. Nor is there much in terms of declensional differences. The only clear distinction in plural marking is instead related to humanness or animacy. The choice of copula and auxiliary forms is, like in Dameli, entirely governed by semantics. This gives us in effect a system of two sex-based genders, masculine and feminine, each with two sub-genders, animate and inanimate.

The assignment in the languages of the Chitral group, which are entirely void of any sex-based distinctions, goes only along semantic lines, where the auxiliary

use in the verbal paradigms reflects an animate vs. inanimate distinction. Certain local case markers only occur with inanimate nouns and not with animate nouns (Heegård Petersen 2006: 53; Bashir 2003: 844). However, it is doubtful whether this can be considered a primary assignment criterion.

7 Pronominal gender

A separate issue, but also necessary to mention in the context, is the presence of pronominal gender distinctions in Hindu Kush Indo-Aryan. In pronominal gender (see Table 4) we find some interesting differences, partly going along sub-classification lines. Even in this case, it is more instructive to differentiate between sex-based distinctions and non-sex-based (i.e. animacy-based) distinctions. Interestingly, so far, no combination of the two (in the same domain) has been noted for any individual language. Note, that only personal pronouns (or demonstratives used as third person pronouns) have been taken as diagnostic in this case.

Only in two of the subgroups do we find evidence for differentiating personal pronouns for masculine and feminine referents (including non-human animates and inanimates), in Kashmiri and in at least four of the Shina languages. These languages all have a two-term system, a masculine third person pronoun contrasting with a feminine, so that even reference to inanimates makes use of one of the two according to their grammatical gender. The differentiation is limited to singular reference and third person, whereas the same term is used for masculine plural and feminine plural alike. Gender is also neutralized in some of the case forms. For instance, Kohistani Shina (14), has separate feminine (a) and masculine (b) ergative pronouns for perfective transitive constructions, whereas there is only one third person singular form used in non-perfective transitive constructions (c) or in intransitive clauses (d).

(14) Kohistani Shina (Schmidt & Kohistani 2008: 181, 217, 247, 224)

- a. *séso* *asór* *ṭiki* *d-eég-i*.
 3F.SG.ERG.PFV 1PL.DAT bread give-PFV-3F.SG
 ‘She gave us food.’
- b. *sési* *ráaty-oo* *kom* *th-áa-o*.
 3M.SG.ERG.PFV night-ABL work do-PFV-3M.SG
 ‘He worked all night.’

- c. *ses* *dōóchi* *ágo* *çiç-eé* *táam*
 3SG.ERG.IPFV tomorrow headshawl embroider-CV complete
th-úu.
 do-FUT.3F.SG
 ‘She will finish embroidering the headshawl tomorrow.’
- d. *sa* *ruleé* *b-eé* *boĵ-áa-n-i.*
 3SG.NOM disguise be-CV go-IPFV-AUX.PRS-3F.SG
 ‘She goes (there) disguised.’

Within the Shina group, there are four different patterns (see Table 11). In Gilgiti Shina and in Brokskat, both nominative and ergative have distinct masculine and feminine forms. In Kohistani Shina (as illustrated above), this distinction is upheld in the (perfective) ergative but is neutralised in the nominative (and elsewhere). In Palula, the opposite holds, and it is in the nominative that gender is differentiated whereas it is neutralised in the ergative (and elsewhere). In Sawi, Kalkoti, Kundal Shahi and possibly in Ushojo, no pronominal gender differentiation is made at all. Kashmiri, the only other HKIA language that makes pronominal gender distinctions, displays the same pattern as Gilgiti Shina does.

Table 11: Pronominal third person gender distinctions in Shina languages

	Nominative		Ergative	
	Masc.	Fem.	Masc.	Fem.
Gilgiti Shina	<i>ro</i>	<i>re</i>	<i>ros</i>	<i>res</i>
Kohistani Shina		<i>sa</i>	<i>sési</i>	<i>séso</i>
Palula	<i>so</i>	<i>se</i>		<i>tíi</i>
Sawi		<i>see</i>		<i>ti</i>

Pronominal differentiation related to animacy is found in a few individual languages belonging to different subgroups. Different pronouns for animate and inanimate reference, respectively, are used in Gawri, as in example (15), in Dameli and possibly also in Torwali.

(15) Gawri (Baart & Sagar 2004: 35, 52)

- a. *ääs* *sä* *äsēē* *duu isaal* *yeeš.*
 3SG.OBL.VIS.AN with 3SG.VIS.POSS.F two women come.PFV.F.PST
 ‘Both his wives had also come with him.’

- b. *abdul häq-ēē än mäy yärääz nãāt*
 Abdul Haq-POSS.F 3SG.OBL.VIS.INAN in interest is.not
 ‘For Abdul Haq, there is no interest in it.’

Curiously, such a distinction is not found in the two languages that otherwise make the most systematic use of animacy distinctions in their agreement patterns, Kalasha and Khowar. For the latter, see example (16).

(16) Khowar (Own data)

- a. *awá ho mar-ít-am*
 1SG.NOM 3SG.DIST.OBL kill-PST.ACT-1SG
 ‘I killed him.’ (KHW-PronDemAA:010)
- b. *tu ho paš-ís-an-a*
 2SG.NOM 3SG.DIST.OBL see-2SG-PRS/FUT.SPC-Q
 ‘Can you see that? [the speaker pointing to an object a few feet away]’ (PronDemAA:018)

8 Gender complexity

Based on the findings in §4–§7, a cautious attempt is made at measuring the relative complexity of the gender systems in HKIA, guided by the complexity metric as laid out by Di Garbo (2016), based on the three following dimensions of complexity: the number of values, the number and nature of assignment rules, and the amount of formal marking, as previously proposed by Audring (2014). In order to arrive at a more significant internal differentiation between the HKIA languages than would otherwise be the case, the metrics were slightly adjusted (see Table 12) as compared to Di Garbo’s. Di Garbo’s features related to manipulable assignment and cumulative exponence, were for instance not taken into account here, partly due to non-applicability to the languages of my sample, partly due to unavailability of comparative data. In the case of the values dimension, a language with four or more genders receives the maximum score (instead of those with 5 or more), and in the case of indexation domains, a language with five or more targets receives the maximum score (instead of those with 4 or more). It is therefore important to note that the scores are primarily intended to provide a relative (i.e. sample-internal) measure (min=0, max=1) rather than being comparable in a wider cross-linguistic sense.

complexity languages.

This metric has been applied to each of the HKIA languages, resulting in the ranking displayed in Table 13. For some of the languages, the number of genders

Table 12: Gender complexity metric (as applied to HKIA)

Complexity dimension	Values	Score
Number of genders	Two genders	0
	Three	0.5
	Four or more	1
Number/nature of assignment rules	Semantic or formal	0
	Semantic + formal	1
Number of target domains	One target domain	0
	Two	0.25
	Three	0.5
	Four	0.75
	Five or more	1

(see Table 3) varies between dialects or is not entirely clear from the descriptions available. In those cases, the highest number in a range was used in the calculation. As for the number of target domains (see Table 4), no differentiation was made between sex-based and non-sex-based agreement. To counter a too literal interpretation of the individual complexity scores, the languages have been grouped into three complexity categories: those scoring *up to and including* 1/3 are LOW gender complexity, those scoring *more than 1/3 up to and including* 2/3 are MEDIUM, and those scoring *more than 2/3* are HIGH gender

In the high complexity category we find three of the four Pashai languages and Shumashti, i.e. the only languages in our sample where we may (although far from conclusively) speak of four genders, or rather systems in which animacy and sex-based differentiation overlap; and Kashmiri, the latter a two-gender system characterized by a high number of target domains. At the other extreme, that is the low complexity category, we find Khowar and Kalasha, the only two languages in our sample with a purely semantic two-way (animate-inanimate) differentiation, as well as Grangali, a masculine-feminine-gender language characterized by having only a single agreement domain. The remaining 17 languages are all of medium complexity according to this metric.

However, it is important to point out that there are other (less measurable) factors, not included in the present metric, that contribute to the overall complexity of individual gender systems, such as the interplay between different assignment criteria (briefly mentioned in §6), declensional differences that do not map directly onto gender distinctions, and the conflation of gender and other grammatical categories (e.g. number and case).

Table 13: HKIA languages ranked for complexity

Rank	Language	Complexity score	Complexity category
1	SW Pashai	0.75	High
1	SE Pashai	0.75	
1	NE Pashai	0.75	
1	Shumashti	0.75	
2	Kashmiri	0.67	
3	Gawri	0.58	Medium
3	Indus Kohistani	0.58	
3	Brokskat	0.58	
3	Palula	0.58	
3	Shina (Gilgiti)	0.58	
4	Tirahi	0.50	
4	Torwali	0.50	
4	Dameli	0.50	
4	Gawarbati	0.50	
4	Ushojo	0.50	
4	Kohistani Shina	0.50	
5	NW Pashai	0.42	
5	Bateri	0.42	
5	Wotapuri-Katarqalai	0.42	
5	Kalkoti	0.42	
5	Kundal Shahi	0.42	
5	Sawi	0.42	
6	Grangali	0.33	Low
7	Khowar	0.00	
7	Kalasha	0.00	

9 Distribution and areal-linguistic implications

The findings presented above enable us to present at least some general tendencies in the geographical distribution of gender properties (see Figure 2).

First, a sex-based gender system with the two values masculine and feminine is the default for Hindu Kush Indo-Aryan. Such a system is found throughout the region, from east to west. However, two exceptions were noted, Khowar and Kalasha, where sex-based differentiation is lacking altogether. Both are situated at the northwestern periphery of the Hindu Kush region, representing the ultimate frontier of Indo-Aryan in general. Furthermore, it is in an adjacent area to those two languages that we find Dameli, a language where sex-based gender is described as being on the retreat. In at least some dialects of NW Pashai, another language spoken in the western-most part of Hindu Kush, sex-based gender may be altogether absent. Non-sex-based gender, or more specifically gender distinctions that have a contrast between animate and inanimate at their core, are also represented in the region, but only clearly so in the western part of the region. Two of the languages with such a basis are, again, Khowar and Kalasha. In a few other languages spoken in the vicinity of the former two – most prominently in varieties of Pashai – an animacy-based system overlaps with a sex-based system. However, the targets for such gender distinctions are often kept distinct.

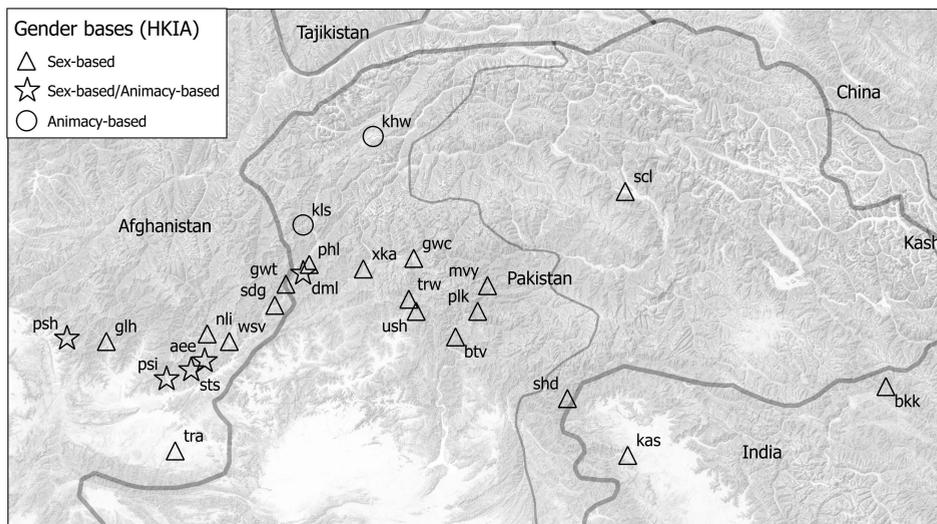


Figure 2: Gender bases in HKIA languages

Second, gender is generally deeply entrenched in those languages that have a sex-based system. Especially in Kashmiri and Shina, i.e. the languages mainly

spoken in the eastern part of the region, gender agreement is displayed with a wide range of targets. In a number of those languages, it is intertwined with person agreement in their verbal morphology, and we also noted some examples of gender agreement being extended to further targets. Kashmiri and some of the Shina languages have gender agreement with demonstratives, and it is only in these languages that we also find sex-based pronominal gender. Gender in some of the Kohistani languages, spoken in the central part of the region, is almost equally pervasive. However, the lack (or loss) of direct object agreement in a few of those languages and the subsequently lower frequency of gender agreement with noun phrases low in the animacy hierarchy may in the long run weaken the masculine–feminine differentiation in parts of the vocabulary where sex plays no role in assignment. Accusative verbal alignment, along with relatively few agreement targets, is probably in some ways related to the erosion of sex-based gender in the Kunar languages in the western part of the region.

In Kashmiri, Kohistani and Kunar, possessive modifiers are frequently targets of gender agreement. Pashai, at the western extreme, shows a diverse picture when it comes to gender pervasiveness. As mentioned before, gender may be lost altogether in some varieties at the western periphery of Pashai; whereas in e.g. SE Pashai, where direct object agreement in parts of the paradigm co-occurs with subject agreement in transitive clauses, such distinctions are frequently displayed also for inanimates. The grammatical pervasiveness of animacy-based gender is nowhere near the pervasiveness of sex-based gender, and its targets are almost invariably restricted to copula verbs and auxiliaries. The (split-)ergative pattern with object agreement in SE Pashai is possibly a factor that may point to a higher frequency of actual and potential contrasts in animacy being expressed than in the solidly accusative languages Khowar and Kalasha.

Third, when it comes to assignment criteria, the usual pattern for the sex-based systems is one of straightforward semantic assignment for humans and higher animates, and a combination of various factors (semantic, morphological and phonological) involved in the assignment of gender for lower animates and inanimates. In the animacy-based systems or sub-systems, geographically almost exclusively found at the western end of the region, semantics is the sole criterion. It also seems likely that a shift from largely non-semantic gender, such as the one in most of the Indo-Aryan languages, to largely semantic gender, is taking place in Dameli (and possibly also in Shumashti).

As already noted, speakers of Hindu Kush Indo-Aryan languages are and have been in contact with speakers of a number of other languages spoken in the region. Let us therefore take a look at these other languages and genera, in order to relate the above findings to areality beyond Indo-Aryan.

Other Indo-Aryan languages. In all four of the region's non-Hindu Kush Indo-Aryan languages (Hindko, Pahari-Pothwari, Gojri and Domaaki), we find a sex-based two-term system typical of Indo-Aryan (Rehman & Robinson 2011; Weinreich 2011; Kogan 2011; Losey 2002: 105–201). Apart from the obvious semantic assignment of humans and other higher animates according to biological sex, lower animates and inanimates are found in the masculine and feminine classes alike. Like in many of the HKIA languages, at a minimum, a sub-set of nouns have overt phonological markers; and at least in Gojri and Domaaki, there is a certain co-variation between gender and declensional class membership. All four languages display gender agreement with adjectives and verbs, and in addition adnominal demonstratives agree in gender in Gojri and Domaaki, and possessives in Gojri. Only Gojri shows evidence of pronominal differentiation. There are no targets of any non-sex-based agreement in any of these languages, and no observed pronominal differentiation related to animacy.

These languages are (apart from the small Domaaki enclave in the far North) mainly spoken in the southeastern part of the region, and conform in all major aspects to the pervasive sex-based gender patterns found in the HKIA languages in the same part of the region, i.e. Kashmiri and various Shina and Kohistani varieties. It is fair to assume a high level of prolonged language contact between at least Kashmiri and one or more of the languages of the Punjabi continuum, whether known as Pahari, Pothwari or Hindko, and possibly also between some of the eastern Kohistani languages and Hindko. However, in most of the areas where there is some overlap between speakers of HKIA and speakers of other Indo-Aryan languages, there is no clear dominance relationship, perhaps with Hindko-dominated parts of Pakistan-held Kashmir as an exception (Rehman 2011: 219). Both Gojri and Domaaki are examples of low-status languages vis-à-vis almost any other language communities that they have been in contact with (Losey 2002: 2–4; Weinreich 1999), and in spite of some intra-regional variations related to the relative socioeconomic status of the Gujar community (Hallberg & O'Leary 1992: 98–99, 143–144), there is no evidence of any significant influence exerted by Gojri on any of the HKIA languages.

Iranian languages. Iranian languages are predominantly found in the western half of the outlined region. They belong to different groupings, and their presence, and relative influence, in the area are of very different time depths. Of the nine Iranian languages represented, only three – Pashto, Shughni and Munji/Yidgha – display a sex-based gender system of some kind (Bashir 2009; Èdel'man & Dodykhudoeva 2009a,b; Kieffer 2003; 2009; Morgenstierne 1938: 110–167; Robson & Tegey 2009; Skjaervø 1989; Windfuhr & Perry 2009). In Munji/Yidgha, gender as a whole is probably in radical decline. In Shughni, the gender categories

show evidence of having restructured as to form a system of semantic classes rather than primarily being assigned on the basis of sex. Only in Pashto, which is also the language in the closest long-time contact with Indo-Aryan, do we find a two-term system akin to the typical Indo-Aryan one, with adjectives, verbs and adnominal demonstratives as agreement targets, and a certain co-variation between gender and declensional membership. Pashto and Shughni are the only Iranian languages in the sample that express pronominal gender. The rest of the Iranian languages of the region have long since lost the sex-based gender systems (masculine–feminine–neuter and masculine–feminine) that characterised their proto-languages (Skjaervø 2009b: 71; Skjaervø 2009a: 204; Yoshida 2009: 288; Durkin-Meisterernst 2009: 242–243). Although animacy distinctions are not part of agreement morphology, animacy does play a role in various forms of Persian, as certain plural allomorphs are found almost exclusively with animate nouns (Windfuhr & Perry 2009: 431), and animacy or humanness, along with register, also governs pronominal choices (2009: 435).

It is notable that it is exactly in the transitional area between Iranian and Indo-Aryan, i.e. in the western-most part of the region, that we find both a number of Iranian languages without gender, and those HKIA languages and dialects that have either lost sex-based gender altogether, or are in the process of shifting away from a primarily sex-based system to a system where animacy distinctions are becoming grammaticalized alongside an existing sex-based system. The gender-reduced systems are found primarily in the northwest, and the systems with overlapping sex-based gender and animacy in the southwest. There is possibly a correlation between gender-preserving Pashto being the most influential language of wider communication in the southwest and the retention of a masculine–feminine contrast in e.g. most Pashai and Kunar varieties. This is in contrast with the Chitral languages, which show evidence, in many parts of their language systems, of long-standing and far-reaching contact with gender-reduced Iranian languages in particular, and with a larger Central Asian contact zone in a more general sense (Bashir 1996: 176–177). Of particular interest is the now historical but crucial contact between speakers of HKIA Khowar and Iranian Wakhi. While Wakhi of today is the less influential of the two in areas where they overlap, the relationship was most likely of a symmetrical kind in a remote past, as evidenced in cross-borrowing of basic vocabulary (Morgenstierne 1936; Morgenstierne 1938: 441–442; Bashir 2007: 208–210). Different varieties of gender-less Persian, whether literary Persian, Dari or Tajik, have also had a significant (and recent) impact on the languages of Chitral and adjacent areas across the Afghanistan border in the northwestern corner of the Hindu Kush region, as a learned language and a lingua franca.

Nuristani languages. In three of the five Nuristani languages we find a two-term system of the Indo-Aryan type: in Waigali (Degener 1998: 39–91), Ashkun (Morgenstierne 1929; Morgenstierne 1934a; Morgenstierne 1952; Buddruss 2006; Grjunberg 1999) and Kati/Kamviri (Strand 2015; Èdel'man 1983: 59–71), whereas its presence in Prasun is doubtful (Morgenstierne 1949; Buddruss & Degener 2017: 69). The available data for the remaining language, Tregami, is insufficient to draw any conclusions (Morgenstierne 1952). Only Kati/Kamviri displays pronominal gender differentiation.

Although there is evidence for Nuristan and the Nuristani languages as an ancient centre of small-scale diffusion (Liljegren & Svård 2017), Nuristani stands in most aspects, especially in more recent times, at the receiving end of contact-induced change, especially from Iranian Pashto and Persian (Degener 2002: 103). As far as gender is concerned, the possible erosion in Prasun may be attributable to the same areal influences from adjacent and influential gender-deprived Iranian languages, as was already suggested above in regard to the HKIA Chitral languages.

Turkic languages. There is a general absence of gender distinctions in Turkic languages, whether as overt markers of nouns or as an agreement feature (Kornfilt 2009: 530). Neither are there any pronominal distinctions in these languages. This is equally true of the two Turkic languages, Uzbek (Boeschoten 1998) and Kirghiz (Kirchner 1998), spoken by populations at the northern periphery of the Hindu Kush region.

There is no present-day overlap, or at best marginally so, between any of the HKIA communities and any of the relatively nearby Turkic-speaking groups. However, it has been suggested that at least the northern-most fringes of the Hindu Kush, together with the Pamirs and perhaps a larger region to the North, form a contact area (Èdel'man 1980; Payne 1989: 423), or alternatively a transit zone between South and Central Asia (Tikkanen 2008: 253), and it is not wholly farfetched to consider Turkic as a component of it. Bashir (1988: 402–421) points out several grammatical features (e.g. inferentiality), primarily in Kalasha and Khowar, with Turkic as their ultimate source, either mediated by certain Iranian Pamir languages or the result of a Turkic substrate. Besides, as Johanson (2013: 104) remarks, the role of Turkic in the massive gender loss in Iranian at large is yet to be fully explored.

Tibeto-Burman languages. Similar to what was said about Turkic, gender in its canonical sense is not a feature generally present in Tibeto-Burman. That is also largely true of Purik, a Tibeto-Burman language spoken at the south-eastern periphery of the region, although there are traces of derivational morphemes indicating male or female sex (Zemp 2013: 118–127). In closely related

Balti (Bielmeier 1985: 81; Read 1934: 4), the other Tibeto-Burman language represented in the region, we find to a larger extent such markers, postposed to some nouns denoting humans or other animates, signalling the sex of the person or animal referred to: *po* or *pho* for male, and *mo* or *ngo* for female (see §6 for formally and functionally similar markers in Brokskat). This type of sex marking or gender marking on the nouns themselves, without any reflexes in agreement patterns, should not be confused with grammatical gender as we have defined it here. In the same vein, an entirely semantically transparent pronominal differentiation can be made in Balti between human male, *kho*, human female, *mo*, and everything else (or when the sex is unknown), *do* (Read 1934: 12–13; Bielmeier 1985: 76).

It is primarily the Shina languages in the East that show traces of interaction with Tibeto-Burman (unless we, along with Tikkanen 1988: 305, consider the possibility that some of the peculiarities of Kashmiri vis-à-vis other Indo-Aryan languages might be attributed to an ancient Proto-Tibetan or Sinitic substrate). Presently, only some groups of speakers of Gilgiti Shina type varieties in Baltistan and the Brokskat community can be said to stand in any such direct and significant contact relationships, and it is only in the latter case that Tibetan plays the role of an influential donor language. It seems likely that the relationship has been more symmetrical in the past; alternatively, we would have to assume a major Tibetan substrate in the eastern Shina-speaking area. That would for instance explain agent-marking (as well as some of its formal reflexes) in Gilgiti as well as in Kohistani Shina (Liljegren 2014: 162–163; Bailey 1924: 211; Hook & Koul 2004: 213–214). In the gender domain, however, Tibeto-Burman contacts do not seem to have led to any loss or restructuring in adjacent HKIA languages, although we lack substantial information on gender assignment in Tibetan loan vocabulary in Brokskat. The continued (and perhaps strengthened) use of overt sex-marking for higher animates in Balti, and not in Purik, seems to point to Shina influences on Balti, and not the other way around.

Burushaski. In the northern part of the region, in close proximity to Indo-Aryan Shina, Indo-Aryan Khowar and Iranian Wakhi, the language isolate Burushaski is spoken. Burushaski has four genders, which makes it the language with the largest number of genders in the entire region. Although the number of differentiating values differs greatly from one part of the grammar to another, or from one target to another (including demonstratives, numerals, verbs, possessives and to some extent adjectives), there is a maximum four-way differentiation between human masculine (HM), human feminine (HF), and two non-human categories that traditionally have been given the labels x and y (Willson 1996: 8–9; Berger 1998: 33–34). Somewhat simplified HM is human male, HF is human female,

x is non-human animate, and y inanimate. However, in reality the relationship between the genders x and y is not quite as straightforwardly related to animacy; x includes not only animals but also fruit and some other count nouns, whereas y is the gender of abstract notions and mass nouns, but also includes e.g. trees and buildings (Yoshioka 2012: 32–33). Burushaski displays verbal agreement in gender and number with the subject as well as with the direct object of transitive clauses, as can be seen in example (17), the first by means of a suffix and the latter by means of a prefix.

- (17) Burushaski (Willson 1996: 17)
hilés-e dasín-mo r toofá-muts píiš ó-t-imi
boy-ERG girl-OBL.F to gift(x)-PL.ABS present 3PL.X-do-3SG.HM.PST
'The boy presented gifts to the girl.'

Gender is also pronominal, but in that case HM and HF are normally neutralised, whereas x and y both have distinct forms of pronominally used demonstratives (Berger 1998: 81–82).

As Burushaski represents one of the oldest, possibly the very oldest surviving, linguistic layer in the Hindu Kush region,³ it is particularly interesting from an areal point of view. While occupying a very modest territory today, the precursor of Burushaski, or other languages perhaps (but not necessarily) closely related to Burushaski, in all likelihood had a wider geographical scope before the advent of Indo-Iranian languages. It has been suggested that such substratal influence underlies some features found across Iranian, Indo-Aryan and Burushaski (Tikkanen 1988; 1999; Bashir 1988: 408–420; Êdel'man 1980). Bashir in particular attributes the gender development in the Chitral languages to Burushaski rather than to Iranian, emphasizing the emergence of animacy-based contrasts. Along the same lines, Payne (1989: 423), mainly referring to Êdel'man's proposed convergence area, attributes the shift from formal-semantic to "purely" semantic assignment in Iranian Pamir languages to a substratum related to or similar to Burushaski, with special reference to a strikingly similar four-way differentiation in Iranian Yazghulami (female human, male human, animal and inanimate), a language situated in today's Tajikistan, only marginally outside the Hindu Kush region as defined here.

³As pointed out to me by Johanna Nichols (p.c.), this makes perfect sense in terms of linguistic geography: a language isolated along different rivers at the highest inhabitable level is almost certainly the earlier one in and has been cut off in its former lower reaches by uphill spreads of other languages.

10 Conclusions

We are now in a position to summarise and draw some overall conclusions regarding the presence and distribution (geographically and subclassification-wise) of various gender properties in Hindu Kush Indo-Aryan (see Figure 3).

There are two types of gender systems in the HKIA languages. A fairly typical New Indo-Aryan sex-based two-gender system is present in the majority of the HKIA languages, and in five of the six subgroups. However, it is curiously missing altogether in the two Chitral group languages, Khowar and Kalasha, both spoken in the northwestern corner of the region. Here, instead, a two-way animacy-based gender differentiation is in place. Furthermore, these two types of gender systems are combined in another few HKIA languages, all of them found in the same part of the larger region, more or less adjacent to the Chitral languages. In one of the latter languages, Dameli, the inherited sex-based gender system is most likely subject to an ongoing process of erosion, and grammaticalized animacy-distinctions have emerged, although largely in complementary distribution with remaining sex-differentiation. In many of the varieties of Pashai, the western-most extension of HKIA, an animate–inanimate differentiation serves as a sub-gender distinction within the main masculine–feminine division.

As for the entrenchment of gender, we observed important differences between the sub-groups, forming a slight decline in pervasiveness moving from East to West. However, there is also a correlation between the presence of object agreement and the reinforcement of formal gender assignment (particularly applicable to inanimate nouns), with object-agreeing languages clustering in the South, while such HKIA languages are lacking altogether in the North. As for the pervasiveness of animacy-based gender, it was similarly suggested that its functional load is higher in systems with ergative verbal alignment (such as in Pashai) than in those with a purely accusative system (such as in the Chitral group), the latter a subject for more refined, preferably corpus-based, studies. Sex-based pronominal gender is a typical Eastern feature, exclusive to Kashmiri and the Shina group, whereas the evidence for animacy-based pronominal gender is scanty and does not allow for any further generalizations.

The weight that different assignment criteria have varies from language to language, and is a topic for which more detailed language-specific studies are needed. At a general level, there is a correlation between primarily sex-based gender and semantic-formal assignment criteria on the one hand; and a correlation between animacy-based gender and more straightforward semantic assignment criteria on the other. While gender in Indo-Aryan in general often involves

declensional differences (Masica 1991: 219), this is not a general tendency in the HKIA languages.

As far as overall complexity is concerned, a few of the HKIA languages stand out, either as being of higher than average complexity or of lower than average complexity. Languages of the first kind are primarily found in the south-westernmost part of the region; these are a handful of languages in which sex-based and animacy-based gender overlap while their targets remain largely distinct. In a single language, Kashmiri, spoken in the south-easternmost part of the region, high complexity is instead related to a high number of target domains. The languages of the second kind are those two (Kalasha and Khowar) in which gender is exclusively animacy-based, and another language (Grangali) in which agreement has been reduced to a single target domain.

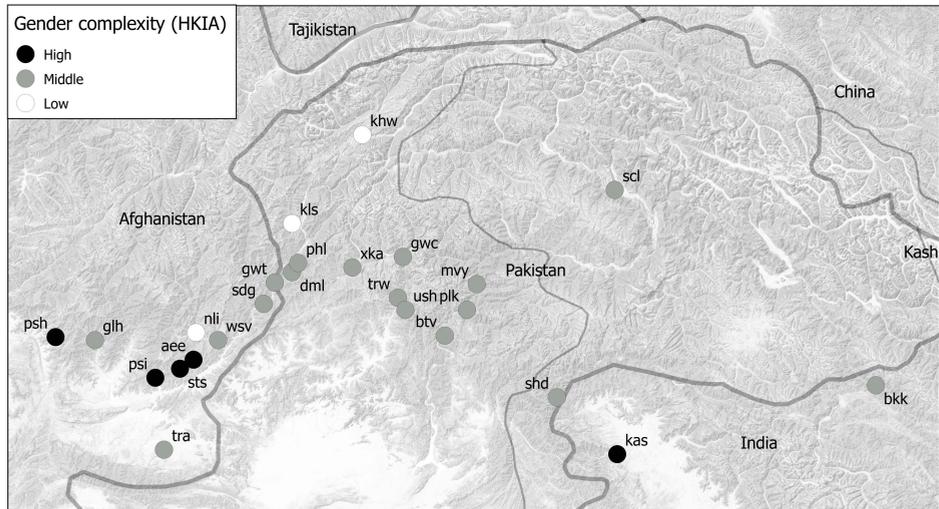


Figure 3: Gender complexity in HKIA languages

The geographical distribution of gender properties within HKIA is clearly parallel to cross-genera distribution within the region. Adjacent to the main (non-HK) Indo-Aryan continua to the Southeast as well as to Pashto, one of the more important gender-preserving Iranian languages, in the South, is where we find the most pervasive sex-based gender systems in HKIA. At the other end, i.e. the Northwest, the gender-less or gender-reduced HKIA languages border with the larger Iranian-dominated region of West and Central Asia, where sex-based gender is a rare or eroding feature, in its turn adjacent to the Turkic belt of inner Asia where gender is altogether lacking. This patterning is clearly in line with Nichols'

(2003: 303) characterization of gender as a stable feature, but only as long as related languages with inherited gender are geographically clustered. We can thus expect to find that languages that have lost this feature are indeed neighbours of one another or are surrounded by non-related languages. This makes sense if we consider Morgenstierne's (1932: 51) hypothesis that the common ancestor of the two "sex-less" languages Khowar and Kalasha represents the earliest northward migration of Indo-Aryans into this region. For a prolonged period this language must have been a relatively minor component in an area where non-Indo-Aryan (perhaps Burushaski-related, or now entirely lost) languages dominated (Tikkanen 1988; Parpola 2002: 92–94), at the time isolated from the rest of the Indo-Aryan varieties from which today's HKIA languages derive. It is also fair to assume that groups of speakers of some of those other languages shifted to a Khowar-Kalasha-type language once it became a more influential element in its new environment.

Perhaps, but not necessarily, related to this is the presence of animacy-based or other semantically highly transparent gender in the North and Northwest, with Burushaski being an obvious example. While animacy-based lexical differentiation with areal manifestation very well could be the result of borrowing, it is harder to imagine such a scenario for the copula or auxiliary agreement patterns in Shumashti and in the Chitral and Pashai languages (the forms themselves also reflecting a common source); instead we have to posit either very old substratal effects, or an internal development reinforced by similar differentiations already in place in neighbouring, and at the time influential, languages. The Dameli inanimate copula form is interesting as it bears no resemblance to the forms in the other HKIA languages (cf. examples (2), (3), (12) and (13)); instead it seems to have been recruited from inherited vocabulary (Morgenstierne 1942: 138). This topic, however, deserves a great deal of more detailed research, also taking data from the Pamir region (to the North of the Hindu Kush) into account.

Acknowledgements

I would like to thank Alla-ud-din, Bahrain, Swat, for his help with digitizing questionnaire data; Noa Lange, Stockholm, for assistance in processing and annotating audio and video recordings; and Maria Koptjevskaja Tamm and Ekaterina Melin, Stockholm, for making the contents of publications in Russian available to me. I am also thankful for the many helpful suggestions offered by Johanna Nichols, the volume editors and four anonymous peer-reviewers.

A special thanks to a number of speakers of various HKIA languages participating in four collaborative elicitation workshops organized in Islamabad and Kabul in the time period 2015 to 2017, thus contributing valuable first-hand data.

This work is part of the project *Language contact and relatedness in the Hindu-kush region*, supported by the Swedish Research Council (421-2014-631).

Special abbreviations

The following abbreviations are not found in the Leipzig Glossing Rules:

ACT	active	REM	remote
AN	animate	SPC	specific
CV	converb	STV	stative
INAN	inanimate	TRZ	transitivizing suffix
H	human	VIS	visible
PTC	participle	x	class x (gender in Burushaski)

References

- Audring, Jenny. 2014. Gender as a complex feature. *Language Sciences* 43. 5–17.
- Baart, Joan L. G. 1997. *The sounds and tones of Kalam Kohistan: With wordlist and texts* (Studies in Languages of Northern Pakistan 1). Islamabad: National Institute of Pakistan Studies, Quaid-i-Azam University & Summer Institute of Linguistics.
- Baart, Joan L. G. 1999. *A sketch of Kalam Kohistani grammar* (Studies in Languages of Northern Pakistan 5). Islamabad: National Institute of Pakistan Studies & Summer Institute of Linguistics.
- 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.
- Baart, Joan L. G. & Muhammad Zaman Sagar. 2004. *Kalam Kohistani texts* (Studies in Languages of Northern Pakistan 9). Islamabad: National Institute of Pakistan Studies & Summer Institute of Linguistics.
- Bailey, T. Grahame. 1924. *Grammar of the Shina (ṣiṇā) language* (Prize Publication Fund 8). London: The Royal Asiatic Society.
- Bashir, Elena L. 1988. *Topics in Kalasha syntax: An areal and typological perspective*. University of Michigan dissertation.

- Bashir, Elena L. 1996. 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 1), 167–179. Karachi: Oxford University Press.
- Bashir, Elena L. 2003. Dardic. In George Cardona & Danesh Jain (eds.), *The Indo-Aryan Languages*, 818–894. London: Routledge.
- Bashir, Elena L. 2007. Contact-induced change in Khowar. In Heather Bolton & Saeed Shafqat (eds.), *New perspectives on Pakistan: Visions for the future*, 205–238. Oxford: Oxford University Press.
- Bashir, Elena L. 2009. Wakhi. In Gernot Windfuhr (ed.), *The Iranian languages*, 825–862. London: Routledge.
- Berger, Hermann. 1998. *Die Burushaski-Sprache von Hunza und Nager 1. Grammatik* (Neuindische Studien 13). Wiesbaden: Harrassowitz.
- Bielmeier, Roland. 1985. *Das Märchen vom Prinzen Čobzañ: Eine tibetische Erzählung aus Baltistan: Text, Übersetzung, Grammatik und westtibetisch vergleichendes Glossar* (Beiträge zur tibetischen Erzählforschung 6). Sankt Augustin: VGH Wissenschaftsverlag.
- Boeschoten, Hendrik. 1998. Uzbek. In Lars Johanson & Éva Á. Csató (eds.), *The Turkic languages*, Reprint edition, 357–378. London: Routledge.
- Buddruss, Georg. 1960. *Die Sprache von Wotapur und Katarqala: Linguistische Studien im afghanischen Hindukusch* (Bonner Orientalische Studien 9). Bonn: Selbstverlag des Orientalischen Seminars der Universität Bonn.
- Buddruss, Georg. 1967. *Die Sprache von Sau in Ostafghanistan: Beiträge zur Kenntnis des dardischen Phalūra* (Münchener Studien zur Sprachwissenschaft Beiheft [Supplement] M). Munich: Kitzinger in Kommission.
- Buddruss, Georg. 2006. Drei Texte in der Wama-Sprache des afghanischen Hindukusch. In Mikhail N. Bogoljubov (ed.), *Indoiranskoe yazykoznanie i tipologija jazykovykh situacii, sbornik statei k 75-letiju professora A. L. Grjunberga (1930-1995)*, 177–200. St. Petersburg: Nauka.
- Buddruss, Georg & Almuth Degener. 2017. *Materialien zur Prasun-Sprache des afghanischen Hindukusch 2. Grammatik* (Harvard Oriental Series 84). Cambridge: Harvard University Press.
- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2013a. Number of genders. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/30>.
- Corbett, Greville G. 2013b. Sex-based and non-sex-based gender systems. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of*

- language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/31>.
- Corbett, Greville G. 2013c. Systems of gender assignment. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/32>.
- Corbett, Greville G. 2014. Gender typology. In Greville G. Corbett (ed.), *The expression of gender*, 87–130. Berlin: Mouton De Gruyter.
- Dahl, Östen. 2000. Elementary gender distinctions. In Matti Rissanen, Terttu Nevalainen & Mirja Saari (eds.), *Gender in grammar and cognition*. Vol. 2: *Elementary gender distinctions: Manifestations of gender* (Trends in Linguistics. Studies and Monographs 124), 577–593. Berlin: de Gruyter.
- Decker, Sandra J. 1992. Ushojo. In Calvin R. Rensch, Sandra J. Decker & Daniel G. Hallberg (eds.), *Languages of Kohistan* (Sociolinguistic Survey of Northern Pakistan 1), 65–80, 193–205. Islamabad: National Institute of Pakistani Studies & Summer Institute of Linguistics.
- Degener, Almuth. 1998. *Die Sprache von Nisheygram im afghanischen Hindukusch*. Wiesbaden: Harrassowitz Verlag.
- 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.
- Degener, Almuth. 2008. *Shina-Texte aus Gilgit (Nord-Pakistan): Sprichwörter und Materialien zum Volksglauben, gesammelt von Mohammad Amin Zia*. Wiesbaden: Harrassowitz.
- Di Garbo, Francesca. 2016. Exploring grammatical complexity crosslinguistically: The case of gender. *Linguistic Discovery* 14(1). 46–85.
- Durkin-Meisterernst, Desmond. 2009. Khwarezmian. In Gernot Windfuhr (ed.), *The Iranian languages*, 336–376. London: Routledge.
- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2019. *Ethnologue: Languages of the world, twenty-second edition*. Online version. Dallas: SIL International. <http://www.ethnologue.com>, accessed 2019-5-6.
- 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. 2009a. Shughni. In Gernot Windfuhr (ed.), *The Iranian languages*, 787–824. London: Routledge.

- Èdel'man, Džoi Iosifovna & Leila R. Dodykhudoeva. 2009b. The Pamir languages. In Gernot Windfuhr (ed.), *The Iranian languages*, 773–786. London: Routledge.
- Emeneau, Murray B. 1965. India and historical grammar. Lecture at Annamalai University in 1959. Annamalainagar.
- Fussman, Gérard. 1972. *Atlas linguistique des parlers dardes et kafirs*. Paris: École française d'Extrême-Orient.
- Greenberg, Joseph H. 1978. How does a language acquire gender markers? In Joseph H. Greenberg, Charles Ferguson & Edith Moravcsik (eds.), *Universals of human language*. Vol. 3: *How does a language acquire gender markers?: Word structure*, 47–82. Stanford: Stanford University Press.
- Grierson, George A. 1927. *Linguistic survey of India*. Vol. 1.1: *Introduction*. Calcutta: Government of India, Central Publication Branch.
- Grierson, George A. 1929. *Torwali: An account of a Dardic language of the Swat Kohistan* (Prize publications fund 11). London: Royal Asiatic society.
- Grjunberg, Aleksandr L. 1971. K dialektologii dardiskikh jazykov (glangali i zemiaki). In N. A. Dvoriankov (ed.), *Indijskaja i iranskaja filologija: Voprosy dialektologii*, 3–29. Moscow: Nauka.
- Grjunberg, Aleksandr L. 1999. Ashkun jazyk. In Džoi Iosifovna Èdel'man (ed.), *Jazyki mira. Dardskie i nuristanskie jazyki*, 126–131. Moscow: Indrik.
- Hallberg, Calinda E. & Clare F. O'Leary. 1992. Dialect variation and multilingualism among Gujars in Pakistan. In Calvin R. Rensch, Calinda E. Hallberg & Clare F. O'Leary (eds.), *Hindko and Gujari* (Sociolinguistic Survey of Northern Pakistan 3), 91–194, 259–302. Islamabad: National Institute of Pakistan Studies & Summer Institute of Linguistics.
- Hallberg, Daniel G. & Calinda E. Hallberg. 1999. *Indus Kohistani: A preliminary phonological and morphological analysis* (Studies in Languages of Northern Pakistan 8). Islamabad: National Institute of Pakistan Studies & Summer Institute of Linguistics.
- Hammarström, Harald, Robert Forkel & Martin Haspelmath (eds.). 2018. *Glottolog 3.2*. Jena: Max Planck Institute for the Science of Human History. <http://glottolog.org/>.
- Heegård Petersen, Jan. 2006. *Local case-marking in Kalasha*. University of Copenhagen dissertation.
- Heegård Petersen, Jan. 2015. Kalasha texts – with introductory grammar. *Acta Linguistica Hafniensia* 47(Supplement 1). 1–275.
- Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.
- Hook, Peter Edwin & Omkar N. Koul. 2004. Case as agreement: Non-nominative subjects in Eastern Shina, non-dative objects in Kashmiri and Poguli, and labile subjects in Kashmiri and Gujarai intransitive inceptives. In Peri Bhaskararao

- & Karumuri Venkata Subbarao (eds.), *Non-nominative subjects*, vol. 1, 213–225. Amsterdam: John Benjamins.
- Johanson, Lars. 2013. *Structural factors in Turkic language contacts*. London: Routledge.
- Kieffer, Charles M. 2003. *Hazāra iv. Hazārāgi dialect – Encyclopaedia Iranica*. <http://www.iranicaonline.org/articles/hazara-4>, accessed 2014-5-14.
- 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*, Reprint edition, 344–356. London: Routledge.
- Klaiman, Miriam H. 1987. Mechanisms of ergativity in South Asia. *Lingua* 71(1–4). 61–102. DOI:10.1016/0024-3841(87)90068-4
- Kogan, Anton I. 2011. Potkhokhari jazyk. In Tatiana. I. Oranskaja, Julia V. Mazurova, Andrej A. Kibrik, Leonid I. Kulikov & Aleksandr Y. Rusakov (eds.), *Jazyki mira: Novye indoarijskie jazyki*, 516–527. Moskva: Academia.
- Kornfilt, Jaklin. 2009. Turkish and the Turkic languages. In Bernard Comrie (ed.), *The world's major languages*, 519–544. London: Routledge.
- Koul, Omkar N. 2003. Kashmiri. In George Cardona & Danesh Jain (eds.), *The Indo-Aryan languages*, 895–952. London: Routledge.
- Lehr, Rachel. 2014. *A descriptive grammar of Pashai: The language and speech community of Darrai Nur*. Chicago: University of Chicago PhD Dissertation.
- Liljegren, Henrik. 2009. The Dangari tongue of Choke and Machoke: Tracing the proto-language of Shina enclaves in the Hindu Kush. *Acta Orientalia* 70. 7–62.
- Liljegren, Henrik. 2013. Notes on Kalkoti: A Shina language with strong Kohistani influences. *Linguistic Discovery* 11(1). 129–160. DOI:10.1349/PS1.1537-0852.A.423
- 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.
- Liljegren, Henrik. 2016. *A grammar of Palula* (Studies in Diversity Linguistics 8). Berlin: Language Science Press.
- 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
- Liljegren, Henrik & Erik Svärd. 2017. Bisyndetic contrast marking in the Hindukush: Additional evidence of a historical contact zone. *Journal of Language Contact* 10(3). 450–484.

- Losey, Wayne E. 2002. *Writing Gojri: Linguistic and sociolinguistic constraints on a standardized orthography for the Gujars of South Asia*. University of North Dakota MA thesis.
- Lubberger, Beate. 2014. *A description and analysis of four metarepresentation markers of Indus Kohistani*. Grand Forks, North Dakota: University of North Dakota MA thesis.
- Lunsford, Wayne A. 2001. *An overview of linguistic structures in Torwali: A language of northern Pakistan*. University of Texas at Arlington MA thesis.
- 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. 1929. The language of the Ashkun Kafirs. *Norsk Tidsskrift for Sprogvidenskap* 2. 192–289.
- Morgenstierne, Georg. 1932. *Report on a linguistic mission to North-Western India* (Serie C III - 1). Oslo: Instituttet for sammenlignende kulturforskning.
- Morgenstierne, Georg. 1934a. Additional notes on Ashkun. *Norsk Tidsskrift for Sprogvidenskap* 7. 56–115.
- Morgenstierne, Georg. 1934b. Notes on Tirahi. *Acta Orientalia* 12. 161–189.
- Morgenstierne, Georg. 1936. Iranian elements in Khowar. *Bulletin of the School of Oriental Studies, University of London* 8(2/3). 657–671.
- Morgenstierne, Georg. 1938. *Indo-Iranian frontier languages*. Vol. 2: *Iranian Pamir languages (Yidgha-Munji, Sanglechi-Ishkashmi and Wakhi*. Oslo: Universitetsforlaget.
- Morgenstierne, Georg. 1942. Notes on Dameli: A Kafir-Dardic language of Chitral. *Norsk Tidsskrift for Sprogvidenskap* 12. 115–198.
- Morgenstierne, Georg. 1945. Notes on Shumashti: A Dardic dialect of the Gawar-Bati type. *Norsk Tidsskrift for Sprogvidenskap* 13. 239–281.
- Morgenstierne, Georg. 1949. The language of the Prasun Kafirs. *Norsk Tidsskrift for Sprogvidenskap* 15. 186–334.
- Morgenstierne, Georg. 1950. *Notes on Gawar-Bati*. Oslo: Det Norske Videnskaps-Akademi.
- Morgenstierne, Georg. 1952. Linguistic gleanings from Nuristan. *Norsk Tidsskrift for Sprogvidenskap* 16. 117–135.
- Morgenstierne, Georg. 1967. *Indo-Iranian frontier languages*. Vol. 3: *The Pashai language, 1, Grammar*. Oslo: Universitetsforlaget.

- Morgenstierne, Georg. 1974. Languages of Nuristan and surrounding regions. In *Cultures of the Hindukush: Selected papers from the Hindu-Kush Cultural Conference held at Moesgård 1970*, vol. 1 (Beiträge zur Südasienforschung), 1–10. Wiesbaden: Steiner.
- Nichols, Johanna. 2003. Diversity and stability in language. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 283–310. Oxford: Blackwell.
- Parpola, Asko. 2002. From the dialects of Old Indo-Aryan to Proto-Indo-Aryan and Proto-Iranian. In Nicholas Sims-Williams (ed.), *Indo-Iranian languages and peoples* (Proceedings of the British Academy 116), 43–102. Oxford: Published for the British Academy by Oxford University Press.
- Payne, John R. 1989. Pamir languages. In Rüdiger Schmitt (ed.), *Compendium linguarum Iranicarum*, 417–444. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Perder, Emil. 2013. *A grammatical description of Dameli*. Stockholm University PhD Dissertation.
- Pirejko, L. A. 1979. On the genesis of the ergative construction in Indo-Iranian. In Frans Plank (ed.), *Ergativity: Towards a theory of grammatical relations*, 481–488. London: Academic Press.
- Radloff, Carla F. & Shakil Ahmad Shakil. 1998. *Folktales in the Shina of Gilgit: Text, grammatical analysis and commentary* (Studies in Languages of Northern Pakistan 2). Islamabad: National Institute of Pakistan Studies & Summer Institute of Linguistics.
- Ramaswami, N. 1982. *Brokskat grammar* (CIIL grammar series 8). Mysore: CIIL.
- Read, Alfred F. C. 1934. *Balti grammar*. London: James G. Forlong Fund.
- Rehman, Khawaja A. 2011. Ergativity in Kundal Shahi, Kashmiri and Hindko. In Mark Turin & Bettina Zeisler (eds.), *Himalayan languages and linguistics: Studies in phonology, semantics, morphology and syntax*, 219–233. Leiden: Brill Academic Publishers.
- Rehman, Khawaja A. & Joan L. G. Baart. 2005. A first look at the language of Kundal Shahi in Azad Kashmir. *SIL Electronic Working Papers* 2005-008.
- Rehman, Khawaja A. & Mark A. Robinson. 2011. Khindko jazyk. In Tatiana. I. Oranskaja, Julia V. Mazurova, Andrej A. Kibrik, Leonid I. Kulikov & Aleksandr Y. Rusakov (eds.), *Jazyki mira: Novye indoarijskie jazyki*, 527–537. Moskva: Academia.
- Robson, Barbara & Habibullah Tegey. 2009. Pashto. In Gernot Windfuhr (ed.), *The Iranian languages*, 721–772. London: Routledge.
- Schmidt, Ruth Laila & Razwal Kohistani. 2008. *A grammar of the Shina language of Indus Kohistan* (Beiträge zur Kenntnis südasiatischer Sprachen and Literaturen 17). Wiesbaden: Harrassowitz.

- Sharma, Devidatta. 1998. *Tribal languages of Ladakh. Part One: A concise grammar and dictionary of Brok-skad* (Studies in Tibeto-Himalayan languages 6). New Delhi: Mittal Publications.
- Siewierska, Anna. 2013. Gender distinctions in independent personal pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/44>.
- Skalmowski, Wojciech. 1985. The linguistic importance of the Dardic languages. *Journal of Central Asia* 8. 5–15.
- Skjaervø, Prods Oktor. 1989. Pashto. In Rüdiger Schmitt (ed.), *Compendium linguarum Iranicarum*, 384–410. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Skjaervø, Prods Oktor. 2009a. Middle West Iranian. In Gernot Windfuhr (ed.), *The Iranian languages*, 196–278. London: Routledge.
- Skjaervø, Prods Oktor. 2009b. Old Iranian. In Gernot Windfuhr (ed.), *The Iranian languages*, 43–195. London: Routledge.
- Strand, Richard F. 1973. Notes on the Nuristani and Dardic languages. *Journal of the American Oriental Society* 93(3). 297–305.
- Strand, Richard F. 2015. *Richard Strand's Nuristân site: Lexicon and grammar*. <http://nuristan.info/lngFrameG.html>, accessed 2016-4-13.
- Tikkanen, Bertil. 1988. On Burushaski and other ancient substrata in northwestern South Asia. *Studia Orientalia* 64. 303–325.
- 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 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 & Educational Research.
- Verbeke, Saartje. 2013. *Alignment and ergativity in new Indo-Aryan languages*. Berlin: De Gruyter Mouton.
- Weinreich, Matthias. 1999. Der Domaaki-Dialekt von Nager. In Georg Buddruss (ed.), *Studien zur Indologie und Iranistik. Band 22*. 203–214. Reinbek: Verlag für Orientalistische Fachpublikationen.

- Weinreich, Matthias. 2011. Domaaki jazyk. In Tatiana. I. Oranskaja, Julia V. Mazurova, Andrej A. Kibrik, Leonid I. Kulikov & Aleksandr Y. Rusakov (eds.), *Jazyki mira: Novye indoarijskie jazyki*, 165–194. Moscow: Academia.
- Willson, Stephen R. 1996. Verb agreement and case marking in Burushaski. *Work Papers of the Summer Institute of Linguistics, University of North Dakota* 40. 1–71.
- Windfuhr, Gernot & John R. Perry. 2009. Persian and Tajik. In Gernot Windfuhr (ed.), *The Iranian languages*, 416–544. London: Routledge.
- Yoshida, Yutaka. 2009. Sogdian. In Gernot Windfuhr (ed.), *The Iranian languages*, 279–335. London: Routledge.
- Yoshioka, Noboru. 2012. *A reference grammar of Eastern Burushaski*. Tokyo: Tokyo University of Foreign Studies PhD Dissertation.
- Zemp, Marius. 2013. *A historical grammar of the Tibetan dialect spoken in Kargil (Purik)*. Berne: University of Berne PhD Dissertation.