Grammatical gender in Hindukush languages

An areal-typological study

Julia Lautin
Grammatical gender in Hindukush languages

An areal-typological study

Julia Lautin

Abstract

In the mountainous area of the Greater Hindukush in northern Pakistan, north-western Afghanistan and Kashmir, some fifty languages from six different genera are spoken. The languages are at the same time innovative and archaic, and are of great interest for areal-typological research. This study investigates grammatical gender in a 12-language sample in the area from an areal-typological perspective. The results show some intriguing features, including unexpected loss of gender, languages that have developed a gender system based on the semantic category of animacy, and languages where this animacy distinction is present parallel to the inherited gender system based on a masculine/feminine distinction found in many Indo-Aryan languages.

Keywords

Grammatical gender, areal-typology, Hindukush, animacy, nominal categories
Grammatiskt genus i Hindukush-språk

En areal-typologisk studie

Julia Lautin

Sammanfattning


Nyckelord

Grammatiskt genus, arealtypologi, Hindukush, animacitet, nominala kategorier
Contents

1. Introduction .................................................................................................................. 1
   1.1 Purpose and research questions ............................................................................. 1

2. Background ................................................................................................................... 2
   2.1 Areal typology ........................................................................................................ 2
   2.2 The Greater Hindukush ......................................................................................... 2
      2.2.1 Geography ....................................................................................................... 2
      2.2.2 Linguistic setting ............................................................................................. 3
      2.2.3 Areal typology of the Greater Hindukush ...................................................... 4
   2.3 Grammatical gender ............................................................................................... 5

3. Method and data ......................................................................................................... 7
   3.1 Language sample and data sources ........................................................................ 7
      3.1.1 Burushaski [bsk] ............................................................................................ 8
      3.1.2 Dameli [dml] .................................................................................................. 9
      3.1.3 Domaaki [dmk] ............................................................................................. 9
      3.1.4 Gawri [gwc] .................................................................................................. 9
      3.1.5 Kalasha [klb] ................................................................................................ 9
      3.1.6 Kashmiri [kas] ............................................................................................. 10
      3.1.7 Kati [bsh] ..................................................................................................... 10
      3.1.8 Kohistani Shina [plk] .................................................................................... 10
      3.1.9 Northern Pashto [pbu] .................................................................................. 11
      3.1.10 South-eastern Pashai .................................................................................. 11
      3.1.11 Purik [prx] .................................................................................................. 11
      3.1.12 Wakhi [wbl] ............................................................................................... 12
   3.2 Analysis ................................................................................................................... 12

4. Results ......................................................................................................................... 13
   4.2 Assignment systems and controller genders .......................................................... 13
      4.2.1 Burushaski ..................................................................................................... 13
      4.2.2 Dameli ......................................................................................................... 14
      4.2.3 Domaaki ........................................................................................................ 14
      4.2.4 Gawri .......................................................................................................... 15
      4.2.5 Kalasha ........................................................................................................ 15
      4.2.6 Kashmiri ...................................................................................................... 15
      4.2.7 Kati .............................................................................................................. 15
      4.2.8 Kohistani Shina ............................................................................................ 16
      4.2.9 South-eastern Pashai ................................................................................... 16
      4.2.10 Northern Pashto ......................................................................................... 17
      4.2.11 Purik .......................................................................................................... 17
      4.2.12 Wakhi ....................................................................................................... 17
List of tables and illustrations

Figure 1. The Greater Hindukush and its languages (Copyright © 2016 by Henrik Liljegren) ........... 3
Figure 2. The languages in the Greater Hindukush ........................................................................ 4
Figure 3. The geographical distribution of the languages in the sample (Copyright © 2016 by Henrik Liljegren) ............................................................................................................. 7
Figure 4. Summary of the distribution of grammatical gender in the language sample (Copyright © 2016 by Henrik Liljegren) ............................................................................................................................ 22

Table 1. Sample languages and primary data sources ......................................................................... 8
Table 2. General overview of the languages gender assignment systems and agreement targets .......... 13
Table 3. Case marking of common nouns (Heegård Petersen, 2006: 53) ............................................. 15
Table 4. Layer 1 case suffixes in Kohistani Shina nouns (Schmidt and Kohistani, 2008: 42-47) .......... 16
### Abbreviations

Animate and inanimate glosses has been consistently changed to AN or INAN, regardless of the gloss in the original sources. This primarily affects the glosses for Kalasha, Dameli and Pashai.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN</td>
<td>Animate</td>
<td></td>
</tr>
<tr>
<td>INAN</td>
<td>Inanimate</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>Person</td>
<td>[prx]</td>
</tr>
<tr>
<td>P/F</td>
<td>Present/future</td>
<td></td>
</tr>
<tr>
<td>PAST</td>
<td>Past tense</td>
<td>[dmk]</td>
</tr>
<tr>
<td>PST</td>
<td>Past tense</td>
<td>[dml]</td>
</tr>
<tr>
<td>PCP</td>
<td>Participle</td>
<td>[gwc]</td>
</tr>
<tr>
<td>PL</td>
<td>Plural</td>
<td>[psi]</td>
</tr>
<tr>
<td>POSS</td>
<td>Possessive</td>
<td>[pbu]</td>
</tr>
<tr>
<td>PRS</td>
<td>Present</td>
<td>[wbl]</td>
</tr>
<tr>
<td>PS</td>
<td>Possessive Suffix</td>
<td>[bsk]</td>
</tr>
<tr>
<td>PSTPRF</td>
<td>Past perfect</td>
<td>[kas]</td>
</tr>
<tr>
<td>SG</td>
<td>Singular</td>
<td>[plk]</td>
</tr>
<tr>
<td>STV</td>
<td>Stative</td>
<td>[psi]</td>
</tr>
<tr>
<td>TRZ</td>
<td>Transitive suffix</td>
<td>[kas]</td>
</tr>
<tr>
<td>X</td>
<td>X-class</td>
<td>[kas]</td>
</tr>
<tr>
<td>Y</td>
<td>Y-class</td>
<td>[kas]</td>
</tr>
<tr>
<td>Z</td>
<td>Z-class</td>
<td>[kas]</td>
</tr>
<tr>
<td>I</td>
<td>Inferential</td>
<td>[kls]</td>
</tr>
<tr>
<td>INF</td>
<td>Infinitive</td>
<td>[kas]</td>
</tr>
<tr>
<td>I</td>
<td>Inferential</td>
<td>[kls]</td>
</tr>
<tr>
<td>INF</td>
<td>Infinitive</td>
<td>[kas]</td>
</tr>
<tr>
<td>I</td>
<td>Inferential</td>
<td>[kls]</td>
</tr>
<tr>
<td>INF</td>
<td>Infinitive</td>
<td>[kas]</td>
</tr>
<tr>
<td>I</td>
<td>Inferential</td>
<td>[kls]</td>
</tr>
<tr>
<td>INF</td>
<td>Infinitive</td>
<td>[kas]</td>
</tr>
</tbody>
</table>

**ISO 639-3 Language codes**

- **bsh**: Kati
- **bsk**: Burushaski
- **dmk**: Đōmaaki
- **dml**: Dameli
- **gwc**: Gawri
- **kls**: Kalasha
- **kas**: Kashmiri
- **plk**: Kohistani Shina
- **psi**: Southeastern Pashai
- **pbu**: Northern Pashto
- **wbl**: Wakhi
- **prx**: Purik
1. Introduction

This is a comparative study of a number of languages in the Greater Hindukush area from an areal linguistic point of view. It aims to investigate grammatical gender systems in the languages of this area, located in north western Afghanistan, northern Pakistan and Kashmir. There are about fifty different languages belonging to several genera all spoken in this mountainous area, in a way isolated from outside interference which has caused them to keep many archaic features. They are simultaneously innovative due to intense language contact and extensive multilingualism. This makes the area especially interesting from an areal typological view.

Grammaral gender has been described as being a long-lived feature in the language families where it is present, but at the same time a recessive feature since it is a system that is not easily borrowed or acquired by other languages. Studies have been carried out on other nominal categories in the area in the past, several within the same research project as this study, but no attempt for an areal-typological survey on grammatical gender. Secondary to the main goal of this investigation, it also made me aware of the gender system in the archaic Gotlandic dialect spoken on my home island. The Gotlandic three gender system is possibly still active, as opposed to the Swedish standard variety where the feminine/masculine distinction has been lost for a long time.

1.1 Purpose and research questions

The purpose of this study is to investigate and understand how the feature of grammatical gender functions and is distributed in the languages of the Hindukush area.

- Is there a grammatical gender system present in the languages and, if so, how does this system categorize nouns and attest on other parts of the clause?
- Which patterns emerge when comparing the results of the study?
- Can the patterns be explained in terms of genealogical relationships or as a result of language contact?
2. Background

In this section some key elements of this thesis will be presented. The first section contains an introduction to areal typology, in the second part the geographical area of interest is presented, both its geographical setting and linguistic situation. The third part is devoted to the definition of grammatical gender.

2.1 Areal typology

Areal typology is the area of research where linguistic typology and areal linguistics overlap. Linguistic typology is concerned with the distribution of linguistic features on a larger scale and the identification of universals of language, while areal linguistics is concerned with identifying linguistic areas and languages within such an area that share several cross-genetic features, making them similar to each other without any genealogical reason (Masica, 2001: 209-212). Linguistic areas can be the size of South Asia or as small as the sub-area of the Greater Hindukush. Areal typology combines features from linguistic typology with the linguistic areas of areal linguistics, which focuses on detailed knowledge of the distribution of a particular phenomenon in languages and dialects within a specific geographical area. Dahl (2001: 1456) defines areal typology as ‘the study of patterns in areal distribution of typologically relevant features of languages’. Features can be randomly distributed, some are universally favored, and they can be inherited. In areas where features are clustered it can be due to borrowing or substratal effects (Koptjevskaja-Tamm, 2010: 5). This in turn can be explained in terms of migration of some speech community, resettlement, language shift, language contact or long-term bi- or multilingualism (Dahl, 2001: 1458-1459). The goal in areal typology is to identify the feature, its geographical distribution and, if possible, the mechanisms and reasons behind a particular areal pattern.

2.2 The Greater Hindukush

2.2.1 Geography

The Greater Hindukush (GHK) area of northern Pakistan and Afghanistan is highly multilingual and the mountains and rivers provide natural barriers separating the linguistic varieties. This is where Pakistan, Afghanistan, Tajikistan, China and India meet (Liljegren, 2014: 134). Here, as in most places, it is difficult to establish whether a variety is a language in its own right or just a dialect. There are no generally accepted criteria to make the distinction between language and dialect, that is, how different a speech variety has to be from another to qualify as a separate language. In Europe, where the nation state provides both geographical and political units with clear cut language boundaries, this is quite unproblematic in many cases. In multilingual areas such as the Hindukush, any attempt to establish whether or not a certain speech variety is a language or a dialect is nearly impossible, especially since some of the languages in the Indo-Aryan family are more or less a dialect continuum and language contact has made neighboring languages more alike. Dialects that once were considered one language could easily evolve in opposite directions, making them mutually incomprehensible with time. Speakers might identify themselves as belonging to a specific group,
ethnically or culturally, and consider the different dialects within that group as one language, even though their own dialect might be linguistically closer to some other language.

2.2.2 Linguistic setting

In this area, five language families or sub-families are present; Indo-Aryan, Nuristani and Tibeto-Burman, the Iranian and Turkic families as well as a language isolate, Burushaski.

Indo-Aryan is one of three Indo-Iranian branches, mainly spoken in South Asia by more than one billion people (Cardona and Dhanesh, 2007: 2). In the Greater Hindukush area there are several Indo-Aryan languages present (for a discussion on the sub-branches and groups of Indo-Aryan, see Masica 1991:446-463), many of which belong to the “Dardic” or northeast Indo-Aryan group. The term “Dardic” mainly refers to a geographic group consisting of seventeen languages with no proof of genealogical unity according to Strand (2013). They are divided into six subgroups; Pashai, Kunar, Chitral, Kohistani, Shina and Kashmiri. These Indo-Aryan languages have retained some ancient features and developed new characteristics different from other Indo-Aryan languages due to their relative isolation in the mountainous area of Hindukush. The relationship and history between these languages is complicated, because of their contact with each other and with languages of other genera (Bashir, 2003: 822). Following Liljegren (2014) this group will be referred to as Hindukush Indo-Aryan (HKIA) further on in this thesis, and they are treated as an areal group and not a genealogical one. Along with the HKIA languages, there are a few other Indo-Aryan languages in this area belonging to different sub-branches or groups.

Nuristan is a geographic area in the Hindu Kush Mountains of northeastern Afghanistan, located in the basins of the Aliningar, Pech, Landai Sin, and Kunar rivers (Strand, 2011). The area was formerly known as Kafiristan meaning ‘land of infidels’ but since the conversion to Islam the name Nuristan ‘land of light’ is used. The languages spoken in the area belong to a separate group, that Morgenstierne (1961) first established was a third branch of Indo-Iranian; the other two being the Iranian and the Indo-Aryan branches. Both Strand (1973) and Degener (2002) agree to this classification. Since the languages are spoken more or less within the boundaries of the Nuristan area, they are called Nuristani languages. They share features with both the Indo-Aryan and the Iranian branch, probably due to intense language contact with both groups during different times, and this has

Figure 1. The Greater Hindukush and its languages (Copyright © 2016 by Henrik Liljegren)
made them difficult to classify (Degener, 2002). The language family consists of 5 languages: Kati, Wasi-leri/Prasun, Kalaša-alâ/Waigali, Tregâmî and Åskun.

Tibeto-Burman is one of two branches in the Sino-Tibetan language family, the other one being Chinese. Ethnologue (Paul, Simons and Fennig, 2016) lists 441 languages in this branch. The sub-grouping is far from finished, but so far it consists of six groups; Lolo-Burmese, Bodic (Tibetic), Sal, Kuki-Chin-Naga, Rung and Kanenic, as well as several smaller sub-groups and not yet classified languages (Thurgood and LaPolla, 2003: 7-21). The Tibetic/Bodic sub-group consists of around 50 languages spoken, of which Ladakh, Balti and Purik are spoken in the Hindukush area (Tournadre, 2013: 117).

The Iranian language family is one of three Indo-Iranian branches, and some 150-200 million people speak Iranian languages from Central Turkey, Iraq and Syria in the west to Pakistan and China in the east. They are spoken from central Caucasus and Tajikistan in the north, all the way down to the Persian Gulf (Windfuhr, 2009: 1,5). The Iranian languages in the Greater Hindukush area are: Dari, Hazaragi, Ishkashimi, Northern Pashto, Parachi, Sanglechi and Shughni (Windfuhr, 2009), as well as Munji, Wakhi and Yidgha that belong to the Pamir branch of the Iranian languages. The Ethnologue (Paul, Simons and Fennig, 2016) lists 41 Turkic languages divided into 5 sub-groups. In the Hindukush area, two Turkic languages are spoken; Southern Uzbek with almost 3 million speakers, belonging to the Eastern group, and Kyrgyz which is a Western Turkic language spoken by 4.3 million people.

Burushaski is considered a language isolate since there is no definite evidence for any genealogical relationship with any other living language (Tikkanen, 1988).

2.2.3 Areal typology of the Greater Hindukush

Several areal features are relevant for the Greater Hindukush area. One feature of particular interest when explaining grammatical gender is split ergativity. A language can be accusative or ergative on several levels, including agreement with the verb, i.e. if the subject or object of a transitive clause triggers verb agreement, and which agreement pattern the intransitive verb shows.

If the subject of a transitive verb is treated in the same way as the subject of an intransitive verb, the pattern is accusative. If the subject of the intransitive verb behaves in the same way as the transitive object, it is ergative. Applied to the feature of grammatical gender, this means that the accusative pattern triggers
agreement on the verb with nouns acting as either transitive or intransitive subjects, or only to the transitive object. The ergative triggers verb agreement with the intransitive subject and transitive object, or only with the transitive subject. In the GHK area it is common for languages to show ergative agreement in some tense/aspect forms, and accusative agreement in others (Liljegren, 2014).

2.3 Grammatical gender

The definition of gender in this thesis is based primarily on the work of Corbett (1991; 2013a; 2013b; 2013c). The defining characteristic of grammatical gender is agreement, that is, if a language has a gender system, it will show agreement dependent on nouns of different types. There has to be evidence of gender outside of the nouns themselves. Corbett agrees with Hockett (1958: 231), who simply states “genders are classes of nouns reflected in the behavior of associated words”.

If a language has a gender system, the most common number of genders is two. In Indo-European languages many, such as German, have three. The German NP’s kleiner Hund ‘small dog’, schwarze Katze ‘black cat’ and schnelles Pferd ‘fast horse’ belong to masculine, feminine and neuter gender respectively. Some have reduced their number of genders to two, including Swedish and Danish where the former masculine and feminine genders merged into what is now called the neuter gender, resulting in an neuter/gender distinction. In the Swedish NP en röd bil ‘a red car’ the head noun bil is in neuter gender, and in ett rött hus ‘a red house’, hus in the neuter, and the indefinite article and the adjective agrees with the head noun in gender. In other languages, such as French and Spanish, the feminine and masculine genders have been retained but the neuter merged into either, or both of them. A few, such as English, have lost their gender altogether.

In Africa there are languages traditionally described as having many noun classes, but since there is no difference in dividing nouns into what have been called genders in Indo-European languages, among others, or the noun classes described by linguists focusing on African or Caucasian languages. The term gender is now used to cover both. Many African languages have extensive noun class systems, which functions in the same way as gender, dividing nouns into different categories, mostly on semantic grounds. The agreement with every part of the clause make noun classes a core feature of the grammar in many languages that have them. The following examples are from Swahili, a Bantu language spoken in eastern and southern Africa:

1) m-toto m-dogo a-mefika  
   CL.1-child CL.1-little CL.1-arrived  
   ‘The little girl arrived’

2) ki-kapu ki-dogo ki-mefika  
   CL.7-basket CL.7-little CL.7-arrived  
   ‘The little basket arrived’ (Nurse and Philipppson, 2003: 111)

Bantu languages, such as Swahili, often show alliterative concord, such as in example 2). The noun includes a form, a marker, which is identical to the marker found on the agreement targets, and markers on different agreement targets are the same. As seen in example 1) the noun marker and the agreement markers are not always completely alliterative, these systems can show irregularities and are not always consistent (Corbett, 1991: 116-119). There are separate noun classes for singular and plural forms. Humans belong to class 1 when in singular and the prefix mu- is added to the noun; mu-ntu ‘person’. In plural the nouns referring to humans are assigned to class 2 and have the prefix ba- as in ba-ntu ‘people’.
The main criterion of a gender system is agreement (Corbett, 1991: 105), and it can therefore be distinguished from other systems, such as classifiers. It is possible, but unusual, for a language to have both gender and classifiers. Further, this distinguishes gender from systems where there is a difference depending on the biological sex of the speaker; lexical contrasts denoting males and females, as in the Finnish examples *jumala* ‘god’ and *jumalatar* ‘goddess’, are not gender distinctions because of their lack of agreement with other parts of a phrase or clause. This sex-distinction is also found in English.

There is always some kind of semantic core in a gender system, based on animacy of some sort; human or animate. It is common to divide nouns into genders by the referents biological sex, some languages denote only human females and males, and some include the sex of (higher) animals. Where biological sex is not the criterion used for dividing nouns into genders, there are other ways to express biological sex if needed. The residual non-human or non-animate nouns can be further divided into classes, by semantic or formal criteria.

In some languages there is a strict semantic system that assigns nouns to a particular gender purely based on the meaning of the noun. Some of the strict semantic systems are based on a division between animate and inanimate nouns, the animate nouns can then be further divided into masculine and feminine categories; where all nouns denoting females are feminine, all nouns denoting males are masculine and the inanimates are neuter with no or few exceptions. There are languages with other categories, for instance a division between human and animate instead of one between feminine and masculine. In other languages, where the semantic system is not as strict, the residue (that is neither female nor male in the example above) is divided into genders by other, more complex, criteria, but there is no language with a strict formal assignment system; it has to be more or less combined with semantic rules. The additional information needed to assign the residue nouns to different categories can be phonological, morphological or both. In Swedish, most animates are assigned to theuter gender, and most inanimates to the neuter gender, in addition, the definite singular form is more or less overt; most nouns that take *-en* are utter, and most nouns that take *-(e)t* are neuter. This is called overt gender, i.e. the gender is visible on the noun itself. In covert systems the form of the noun is *inherent*, and there is no way to determine the gender of the noun only by the form (Corbett, 1991: 117).
3. Method and data

In the first section of this chapter, the language sample is presented together with a short presentation of each language and the primary data sources. In the second section, the analysis that will be conducted on the data is presented.

3.1 Language sample and data sources

Figure 3. The geographical distribution of the languages in the sample (Copyright © 2016 by Henrik Liljegren).

In the Greater Hindukush area there are about fifty languages from five or six genera. For this thesis, a sample of twelve languages has been chosen with respect to the number of languages in the different language families and with regard to the geography of the area. At least one grammar or language description of sufficient quality was needed for each language and because of this some languages had to be chosen above others. The area is not yet well-described and many languages remain more or less without a proper language description.

The only language isolate in the area, Burushaski, is present in this sample. About 50% of the languages in the Greater Hindukush are Hindukush Indo-Aryan, and one language from each subgroup is present in this sample: South-eastern Pashai (Pashai); Dameli (Kunar); Kalasha (Chitral); Gawri (Kohistani); Kohistani Shina (Shina) and Kashmiri (Kashmiri). There are a few other Indo-Aryan languages in the area, belonging to different groups. For this sample Domaaki was chosen, since it is somewhat well-described.

Out of the approximately 50 languages in GHK, 10 are Iranian languages; 6 languages belong to the Pamir sub-group, out of which Wakhi was chosen. The other Iranian language in this sample is Northern Pashto.

There are 17 Nuristani languages in total, out of which Ashkun, Kamviri, Kati, Prasuni, Tregami and Waigali are spoken in the Greater Hindukush. These represent about 10% of the total number of languages in the area, and Kati was chosen as a representative of Nuristani languages in the sample.
The Turkic languages in the Greater Hindukush, Southern Uzbek and Kyrgyz, are not present in this sample because of their location at the peripheral northwest corner of the area and their limited contact with the more central languages, as well as the lack of sufficient data.

<table>
<thead>
<tr>
<th>Language</th>
<th>Language family</th>
<th>Primary data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dameli [dmk]</td>
<td>HK Indo-Aryan</td>
<td>(Perder, 2013)</td>
</tr>
<tr>
<td>Gawri [gwc]</td>
<td>HK Indo-Aryan</td>
<td>(Baart, 1999)</td>
</tr>
<tr>
<td>Kalasha [klh]</td>
<td>HK Indo-Aryan</td>
<td>(Bashir, 1988) (Heegård Petersen, 2006)</td>
</tr>
<tr>
<td>Kashmiri [kas]</td>
<td>HK Indo-Aryan</td>
<td>(Koul, 2003)</td>
</tr>
<tr>
<td>Kohistani Shina [plk]</td>
<td>HK Indo-Aryan</td>
<td>(Schmidt and Kohistani, 2008)</td>
</tr>
<tr>
<td>Southeastern Pashai [psi]</td>
<td>HK Indo-Aryan</td>
<td>(Lehr, 2014) (Morgenstierne, 1967)</td>
</tr>
<tr>
<td>Northern Pashto [pbu]</td>
<td>Iranian</td>
<td>(Robson and Tegey, 2009)</td>
</tr>
<tr>
<td>Wakhi [wbl]</td>
<td>Iranian</td>
<td>(Bashir, 2003)</td>
</tr>
<tr>
<td>Purik [prx]</td>
<td>Tibeto-Burman</td>
<td>(Zemp, 2013)</td>
</tr>
</tbody>
</table>

Table 1. Sample languages and primary data sources.

Table 1 presents an overview of the languages in this sample, 12 in total, their ISO 639-3 code and the main data sources available for each language. Each language will be presented in more detail below.

### 3.1.1 Burushaski [bsk]

Burushaski is a language isolate with about 100,000 speakers. Burushaski consists of the Hunza, Nagar and Yasin varieties (Willson, 1996: 1). The Hunza dialect is spoken between groups of Shina and Wakhi speakers and among the Wakhi villages in the north Hunza valley; the Nagar variety is spoken in the Nagar valley and on the south side of the Hunza River and the Yasin variety throughout the Yasin valley. There are some villages and settlements in the Ishkoman and Gilgit valleys where Burushaski is spoken as well (Yoshioka, 2012: 2-4).

The neighbouring languages are Khowar in the west, Shina in the south, Wakhi in the north, Balti to the east, and the few remaining Domaaki speakers in Hunza and Nagar (Willson, 1996: 1). The Hunza and Nagar varieties have most contact with Shina speakers, while the Yasin dialect is influenced by Khowar as all Burusho in the Yasin valley speak Khowar as their second language. Urdu is used as the primary language for trading, traveling and education. Due to this, Urdu is now influencing Burushaski (Yoshioka, 2012: 8).

3.1.2 Dameli [dml]

Dameli is an Indo-Aryan language spoken by 5-6,000 people in about 11 villages in the Damel valley in the southwest Chitral district of northern Pakistan (Perder, 2013: 3). Neighbouring languages include Kati to the west, Gawarbati and Kati to the southwest, Gawri to the east and Kalasha to the north. Dameli is also in contact with Palula, Khowar and Pashto. Most of the men speak Pashto and Urdu as second languages; some of them also speak Khowar or Kati. The second languages are mostly used for trading and work.

The doctoral thesis ‘A Grammatical Description of Dameli’ by Emil Perder (2013) is the primary source of data, as well as Morgenstierne’s ‘Notes on Dameli’ from 1942.

3.1.3 Đōmaakí [dmk]

Đōmaakí belongs to the Central Indo-Aryan group. Its isolation from related languages in combination with intense language contact with languages of other groups, as well as extensive bilingualism, have caused Đōmaakí to lose some of its Central Indo-Aryan features. Instead, it shares many features with the Hindukush Indo-Aryan languages. It is spoken in the northern areas of Pakistan exclusively by speakers that belong to the ethnic Đōóm community. Most of them are found in the villages of Đomyaal and Mominabad in the Nagar and Hunza valleys as well as in and around Gilgit. Many ethnic Đōóm have switched to Shina and/or Burushaski. In 2002 there were less than 350 remaining Đōmaakí speakers. There are two mutually comprehensible dialects, one spoken in Nagar by approximately 40 people, and the other in Hunza by 300 speakers (Weinreich, 2008: 299-300). The Đōoma people are a socially marginalized minority and most speakers considers their mother tongue as an obstacle towards full integration into their host communities. This negative attitude result in a situation where parents choose not to teach Đōmaakí to their children and the language is expected to be extinct within a couple of generations (Weinreich, 2011: 1-2).

Matthias Weinreich is one of the few who have written about the Đōmaaki language, and some of his works have been used here; ‘Two varieties of Đōmaaki language’ an article written in 2008, in English translation (from German), and ‘Đōmaaki Grammar Outline’ in English translation (from Russian), from 2011.

3.1.4 Gawri [gwc]

Gawri is the Indo-Aryan language spoken in the Kalam area in the northernmost part of the district Swat and in a few villages inhabited by Gawri speakers in the Panjkora valley of Dir Kohistan. The number of speakers is 60-70,000 according to Baart’s (1997) estimate. Gawri has three main dialect clusters; Kalam proper, Utrot and Ushu. Alternate names for the language are Bashkarik, Garwi or Kalam Kohistani. Various other languages are spoken in the Kalam area, such as Pashto, Gujari and Khowar. Most male speakers also know Pashto which is the medium of education, and a few speakers with some higher level education also know Urdu to some extent. Most women do not know any Urdu and their ability to speak Pashto is limited.

Primary data sources are the works of Joan Baart; A sketch of Kalam Kohistani grammar (1999) and The sounds and tones of Kalam Kohistani (1997).

3.1.5 Kalasha [kls]

In the Chitral district of northwest Pakistan, in the Birir, Bumburet and Rumbur valleys on the west side of the Chitral River, Kalasha is spoken by 5-6,000 people most of which have Khowar as their
second language (Bashir, 1988: 31-34). Kalasha is a HKIA language, and it is closely related to Khowar, with which they share some unique similarities (Bashir, 1988: 31). It has two main varieties, a northern and a southern. The northern variety is in daily contact with Khowar and Kati, and the southern is surrounded by speakers of Pashto, Khowar and Kati, and to some extent Gujar. Kalasha is not just a people and their language, but also a religious group. It is the last animistic religion in central Asia, and the ethnic Kalasha that have converted to Islam instead use Khowar or Palula. Because of this, the language is threatened and possibly endangered (Heegård Petersen, 2006: 7ff).

For this thesis I will use the doctoral thesis of Elena Bashir; Topics in Kalasha syntax, an areal and typological perspective (1988) and the work Local case-marking in Kalasha by Jan Heegård Petersen (2006).

3.1.6 Kashmiri [kas]

Kashmiri is spoken by approximately 5.5 million people in the disputed territory of the Kashmir valley. The Kashmir valley is divided between Azad Kashmir, controlled by Pakistan, and the Jammu and Kashmir province, controlled by India. Kashmiri is also spoken in other parts of India as well as in other countries. The language is closely related to Shina and some other north-western Indo-Aryan languages but it also has some unique features that cannot be found in other Indo-Aryan languages. Kashmiri has three dialects inside the valley of Kashmir; Maraz, Kamraz and Sringar as well as two main regional dialects spoken outside of the valley; Poguli in the Pogul and Pakistan valleys and Kashtawari in the Kashtawar valley, which can be seen as separate languages. In the area where Kashmiri is spoken, most inhabitants speak Tibeto-Burman languages such as Ladakh (Masica, 1991: 19).

The primary source is Omkar N. Koul’s work on Kashmiri (2003), and Modern Kashmiri Grammar (Koul and Wali, 2006).

3.1.7 Kati [bsh]

Kati is a Nuristani language spoken in the Kt’ivi valley, the Ūāmg’al and Kul’em valleys of upper Laghmān, in the Landai S’in valley as well as in some tributary valleys of the Kunar River in Afghanistan and along the border to Pakistan. There are about 40-60,000 Kati speakers in total, divided into two major dialect clusters; Katavari and Kamviri. In the area several other Nuristani languages are spoken; neighboring languages include Kalaṣ-ala, Pashai, Munji, Khowar, Kalasha, Palula, Dameli, Farsi, Gawar-Bati and Vasi-vari, even though the languages are not necessarily in daily contact or influence each other. Kati is also called Bashgali or Shekhani and Strand (2011) refers to it as Kāmk’ata-vari.

There is no grammar published of the Kati language, and the most important source is the Nuristan.info website, administered by Richard Strand (2011), and a publication on the Encyclopaedia Iranica web site by Strand (2010). Edelman’s Dardic and Nuristani languages’ (1983) contains some information on Kati as well.

3.1.8 Kohistani Shina [plk]

Kohistani Shina is one of several Shina dialects, or rather one variety in a dialect continuum stretching from the Kunar valley in Afghanistan to the Kashmir valley. Following Strand (2001) the continuum is divided into two main clusters, and the Kohistani variety belongs to the Chilasi cluster. It is spoken by approximately 370,000 people in Kohistan and by a number of speakers that have migrated to other
districts. Kohistani Shina is spoken in Kolai, Palas, Jalkot, Shatial and Sazin along the left bank of the Indus River. Bilingualism is low since Shina is the dominant language in most local spheres. The local administration is mostly in Urdu but the Shina speaking community has little or no education, and therefore little knowledge in Urdu (Schmidt and Kohistani, 2008: 12f).

A grammar of the Shina language of Indus Kohistan (Schmidt and Kohistani, 2008) is the primary data source for the language.

3.1.9 Northern Pashto [pbu]

Pashto is an Iranian language and one of the principal languages of Pakistan and Afghanistan. With 7.5 million speakers in Afghanistan it is spoken by over 50% of the population. It is also spoken by approximately 90% of the population of the Northwest Frontier Province of Pakistan, some 14 million speakers, primarily in the south western parts of the region and is used as a lingua franca for most speakers of Pashai, Gawarbi, Dameli, Palula and most of the Kohistani languages (Henrik Liljegren, personal communication). Outside of Pakistan and Afghanistan, Pashto is spoken to some extent in Iran. It is one of the official languages of Afghanistan, but does not have official status in Pakistan (Robson and Tegey, 2009: 721).

Data of the northern variety of Pashto was difficult to find, so Pashto by Barbara Robson and Habibullah Tegey (2009) was used even though their research is focused on what they describe as a central dialect of Pashto.

3.1.10 South-eastern Pashai

Pashai is an Indo-Aryan language spoken to the north of the Kabul River and south of Nuristan in northeast Afghanistan, in a number of valleys from Panjshir to the Kunar River, by 108,000 speakers. It is surrounded by Pashto, Dari, Shumashhti and the Nuristani languages. The language is influenced primarily by Pashto and Persian while Pashai is influencing Shumashhti. Pashai is a dialect continuum with mutual incomprehensibility, both Morgenstierne (1967) and Lehr (2014) divides it into four different languages, this division is followed by the Ethnologue (Paul, Simons and Fennig, 2016). The southeastern variety is spoken in and around Darrai Nur. The incomprehensibility between the varieties has increased because of the intense language contact in the north and northeast with other Hindukush Indo-Aryan languages, and in the south and southeast with Persian and Pashto.


3.1.11 Purik [prx]

Purik is a Tibeto-Burman language spoken by almost 40,000 speakers in the Kargil district of Jammu and Kashmir in northeast India, which is situated between the cities of Leh in Ladakh, Padum in Zangskar, Srinagar in Kashmir and Skardo in the Gilgit-Baltistan area of Pakistan (Zemp, 2013: 3).

An alternate name for the language used in the past is Burig. The speakers of Purik are in regular contact with other Tibetan languages and varieties, such as Shamskat, Central Ladakhi, and Zangskari as well as Central Tibetan languages on the market in Kargil. There is also some language contact with the HKIA languages Shina and Brokskat. Purik has been influenced by Burushaski and Kashmiri in the past, but in recent times more so by Urdu and English (Zemp, 2013: 9-10).
A historical grammar of the Tibetan dialect spoken in Kargil (Purik), a Ph.D. thesis by Marius Zemp (2013) is the main source for the Purik data.

3.1.12 Wakhi [wbl]

The language belongs to the Pamir branch of the Iranian language family. In Pakistan, Wakhi is spoken in the valleys of Hunza, Gojal, Ishkoman, Yasin and Yarkhun with the largest Wakhi population situated in Gojal, a part of the Hunza valley upstream from the Burushaski speaking central Hunza. In Gojal, there is also a small population of Burusho. In 1990-91 there were about 12,000 Wakhi speakers in Pakistan and some 9,500 in the Wakhan region of Afghanistan. There are some differences between the dialects spoken in the different valleys, but not enough to cause problems with mutual intelligibility. Wakhi is also spoken in Tajikistan and China (Bashir, 2009: 825).

In Pakistan, language contact is primarily with speakers of Shina, Ishkoman, Khowar and Burushaski, mainly because of trading and other services in Gilgit. Urdu is used as a trade language. Khowar has influenced the Wakhi language in the past, and Urdu influences it today. This might be accelerated by the use of Urdu as lingua franca as well as the language for education. No language is threatening to replace Wakhi at the moment but the ability to speak Urdu is becoming widespread through education (Bashir, 2009: 826).

The primary data sources are Bashir’s work Wakhi (2009) and the second volume of Morgenstierne’s Indo-Iranian Frontier Languages (1938) that contains his work on Wakhi.

3.2 Analysis

The analysis of every data source for every language will be conducted systematically to establish whether or not there is a grammatical gender system present in the language, on what criteria it divides the nouns and what other parts of a clause agrees with the noun.

If there is a grammatical gender feature, it will have controller genders, i.e. the set of nouns that are assigned to a particular gender. Every noun will be assigned to one gender, and the first step is to establish which rules apply to the gender assignment system.

First of all, the semantic core divides at least some nouns into genders by animacy of some level and the analysis needs to establish where along the animacy hierarchy each system makes the cut, and how which categories are present above and below that line (e.g. human, animate and inanimate categories or masculine and feminine). The analysis must identify the set of rules that each system has for assigning the residual nouns into genders, that is, either phonological or morphological criteria. The markers on the noun can be different in each language; some number and case markers can make the gender visible on the noun, although it may also be an inherent category of the stem.

Next, a gender system has to have target genders, i.e. the gender marked on the elements showing agreement with the noun. The elements showing gender agreement can be adjectives or other attributive modifiers such as demonstratives, verbs, articles, numerals, participles or adverbs. Attributive modifiers show agreement with their head noun, and some may be invariant. Verbs can show agreement in gender with the subject or object noun. In some cases, depending on the structure or layout of the grammar, there is a need to investigate the syntax to clear out which argument of a clause agrees with the verb. They can agree in gender with either the subject of a transitive clause, or its object, or both. The verb may take different agreement depending on the tense and aspect; this is a question of whether the language is accusative or ergative, or split ergative.
4. Results

In the following section the results of the study will be presented, first with a general overview of the findings presented in table (2), a brief summary of the nominal categories in each language, followed by a more detailed explanation of the findings in 4.1 where the gender systems, controller genders and the assignment systems are presented, and in 4.2 focused on the target genders.

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>Semantic base</th>
<th>Number of genders</th>
<th>Assignment system</th>
<th>Primary target genders</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Burushaski [Isolate]</td>
<td>Yes</td>
<td>Biological sex</td>
<td>4</td>
<td>Semantic</td>
</tr>
<tr>
<td>2</td>
<td>Dameli [HKIA]</td>
<td>Yes</td>
<td>a) Biological sex</td>
<td>2</td>
<td>Semantic and formal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b) Animacy</td>
<td>2</td>
<td>Semantic</td>
</tr>
<tr>
<td>3</td>
<td>Domaaki [Central IA]</td>
<td>Yes</td>
<td>Biological sex</td>
<td>2</td>
<td>Semantic and formal</td>
</tr>
<tr>
<td>4</td>
<td>Gawri [HKIA]</td>
<td>Yes</td>
<td>Biological sex</td>
<td>2</td>
<td>Semantic and formal</td>
</tr>
<tr>
<td>5</td>
<td>Kalasha [HKIA]</td>
<td>Yes</td>
<td>Animacy</td>
<td>2</td>
<td>Semantic</td>
</tr>
<tr>
<td>6</td>
<td>Kashmiri [HKIA]</td>
<td>Yes</td>
<td>Biological sex</td>
<td>2</td>
<td>Semantic and formal</td>
</tr>
<tr>
<td>7</td>
<td>Kati [Nuristani]</td>
<td>Yes</td>
<td>Biological sex</td>
<td>2</td>
<td>Semantic and formal</td>
</tr>
<tr>
<td>8</td>
<td>Kohistani Shina [HKIA]</td>
<td>Yes</td>
<td>Biological sex</td>
<td>2</td>
<td>Semantic and formal</td>
</tr>
<tr>
<td>9</td>
<td>Southeast Pashai [HKIA]</td>
<td>Yes</td>
<td>a) Biological sex</td>
<td>2</td>
<td>Semantic and formal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b) Animacy</td>
<td>2</td>
<td>Semantic</td>
</tr>
<tr>
<td>10</td>
<td>Northern Pashto [Iranian]</td>
<td>Yes</td>
<td>Biological sex</td>
<td>2</td>
<td>Semantic and formal</td>
</tr>
<tr>
<td>11</td>
<td>Purik [Tibeto-Burman]</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>Wakhi [Iranian]</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2. General overview of the languages gender assignment systems and agreement targets.

4.2 Assignment systems and controller genders

Every gender system has a semantic base which divides nouns into groups depending on a cut somewhere on the animacy hierarchy. The residue is dealt with in accordance with different strategies, assigning the nouns to genders by a set of phonological or morphological rules. In some languages the gender is overt, at least in part, or visible on the noun by the means of number or case affixes. This section is devoted to the semantic base, the formal criteria for assigning the residue, and, in the languages that have them, affixes as gender markers.

4.2.1 Burushaski

Burushaski nouns are inflected for gender, case, and number (Yoshioka, 2012: 31). The gender system has four classes according to Willson (1996: 8). The nouns are divided into human masculine (HM) or
human feminine (HF) and these merge in the plural to form a non-sex-differentiated human category. The residue is further divided into two categories where the X-class consists of concrete objects such as animals and fruits and the Y-class of objects such as liquids, trees and notions. Yoshioka (2012: 33) divides the Y-class further forming a subclass for time, place and number he calls the z-class. It behaves as the Y-class except in agreement with numerals and in the genitive where it takes the same oblique marker as the HF-class. There is no overt gender marking, that is, it is impossible to tell by the form of the word which gender it belongs to, but the plural markers might give a clue as to which gender a noun is to be assigned to (Yoshioka, 2012: 32). The most frequently used plural types are –caro or –tiŋ for the merged human plural gender, -muc, -nc or –o for the X-gender (which can be used for the H-gender as well) and the –y suffix that is used for the Y-gender. The suffix –őyo can be used for every noun class, and every suffix has a number of allomorphs (Yoshioka, 2012: 41). The case suffixes in Burushaski are not inflected for gender (Yoshioka, 2012: 49); even though some combinations of case and gender require an additional oblique marker to be grammatical (e.g. the HF-class in locational cases and the Z-class nouns in the ablative case cf. Yoshioka 2012: 49-64).

4.2.2 Dameli

Dameli nouns have the inherent categories of gender and animacy, and are inflected for case and number (Perder, 2013: 49). The two-gender system in Dameli is based on biological sex and divides humans and animals into either masculine or feminine gender. The residue is assigned to either gender by phonological criteria; most nouns ending with –a are assigned to the masculine gender and the nouns ending in –i are assigned to the feminine. There are some exceptions to both the semantic and phonological assignment rules; many nouns referring to body parts are feminine even though their form is typically masculine, and some other nouns are assigned to either masculine or feminine by unknown causes (Perder, 2013: 52-53). The forms of the number- and case suffixes are invariant for gender (Perder, 2013: 56-62).

The other system present in Dameli is based on the strictly semantic division between animate and inanimate nouns (Perder, 2013: 50-51). There are no overt marking on the nouns, and the number and case suffixes are invariant for this type of gender as well, even though the ergative case marker applies only to animates, and the instrumental only to inanimates, but this division is not grammatical since the ergative is used for agents, which almost always are animate and the markers are identical in form (Perder, 2013: 56-62).

4.2.3 Đomaakí

Đomaakí nouns are divided into two declensional classes depending on their ending. (Weinreich, 2011: 7-8). They carry the property of gender, and are inflected for number (the Rectus plural suffixes) and case (Weinreich, 2011: 6-7). The Đomaakí gender system has a semantic base of biological sex, and the residue is divided depending on the ending of the noun stem; the masculine nouns ends in –u or –ő and the feminine in –a. Noun gender is visible when the feminine suffix –a is added to the stems of some nouns (Weinreich, 2011: 6-7), on the noun case suffixes in nominative plural and instrumental-ergative singular. The suffix of singleness, which Weinreich (2011: 6) describes as the singulative-indefinite suffix, displays the inherent gender of the noun and is added to a noun in nominative singular. The masculine forms are –ek, –ak or –k depending on the form of the nominative singular suffix, and the feminine forms are –ka, –ika or –aka (Tikkanen, 2011: 209).
4.2.4 Gawri

Gawri nouns distinguish between oblique and nominative case, and singular and plural number (Baart, 1999: 31-32). The language has two genders, masculine and feminine. All animates are assigned to either gender based on their biological sex, even though the masculine is used in a more general way when referring to no particular sex. The inanimate nouns are not easily predicted as to which gender they belong to. Nouns referring to smaller objects are more often feminine, and nouns referring to larger objects are often masculine (Baart, 1999: 31-32). The case and number suffixes are not affected by the gender of the noun they are attached to (Baart, 1999: 35-37).

4.2.5 Kalasha

Kalasha nouns show the categories of gender, person and number. Since Kalasha has lost its inherited gender system based on biological sex, the gender system present today is based on the strictly semantic criteria of animacy, dividing nouns into the animate or inanimate categories (Bashir, 1988: 39-40). The animacy distinction is only visible on the nouns themselves in some cases; animate nouns have separate forms in nominative PL, genitive SG and PL, the other cases are used for inanimate entities, except for the vocative which is exclusively used for animates (Heegård Petersen, 2006: 53).

<table>
<thead>
<tr>
<th>Case</th>
<th>Animate Singular</th>
<th>Plural</th>
<th>Inanimate Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>-Ø</td>
<td>-Ø, -án, -an</td>
<td>-Ø</td>
<td>-Ø</td>
</tr>
<tr>
<td>Genitive-Obl</td>
<td>-as</td>
<td>-an, án</td>
<td>-as</td>
<td>-an</td>
</tr>
<tr>
<td>Instrumental</td>
<td>--</td>
<td>--</td>
<td>-an</td>
<td>--</td>
</tr>
<tr>
<td>Locative</td>
<td>--</td>
<td>--</td>
<td>-a, -una, -ai</td>
<td>--</td>
</tr>
<tr>
<td>Ablative</td>
<td>--</td>
<td>--</td>
<td>-yei, -ani, -aw</td>
<td>--</td>
</tr>
<tr>
<td>Temporal</td>
<td>--</td>
<td>--</td>
<td>-ano</td>
<td>-asa</td>
</tr>
<tr>
<td>Vocative</td>
<td>-ow, -Ø</td>
<td>-Ø (-an?)</td>
<td></td>
<td>--</td>
</tr>
</tbody>
</table>

Table 3. Case marking of common nouns (Heegård Petersen, 2006: 53)

4.2.6 Kashmiri

Kashmiri has a two-gender system based on biological sex and distinguishes singular from plural. The nouns are inflected for case, both the ergative-absolutive type and the accusative-nominative type (Koul and Wali, 2006: 25-32). All animate nouns are either masculine or feminine, assigned by their biological sex. Inanimate nouns are assigned to either gender based on their endings. The phonological rules include palatalization, additional vowel, consonant changes or in some cases suppletion to form feminine nouns of masculine stems, such as šur ‘child.M’ > šuɨr ‘child.F’ or the suppletive counterparts nečuv ‘son’ vs. kūr ‘daughter’ (Koul, 2003: 906-908). Kashmiri plural forms are different depending on the gender of the noun. The rules include suffixation, palatalization and vowel changes, and a few nouns have the same form for both numbers (Koul, 2003: 907). All case suffixes have different forms for masculine and feminine singular forms, except the unmarked nominative, and possessive suffixes vary depending on the gender (and number) of the head noun (Koul and Wali, 2006: 34-35).

4.2.7 Kati

Kati nouns distinguish between feminine and masculine gender, and are inflected for case. Plural forms occur only in the oblique case. The nouns are divided into masculine and feminine gender. Nouns referring to animate entities are assigned depending on the referents sex. Inanimate nouns are
assigned to either masculine or feminine on basis of their ending; Nouns ending in \(-i\) belong to the feminine gender and the rest is masculine. The masculine is used as a general, neutral form and the feminine as the marked one (Strand, 2011). Gender is not expressed morphologically on the singular nominative or direct case, but in the singular oblique, instrumental or genitive case, where the nouns are marked with the masculine suffix \(-e\) and the feminine \(-a\). There is no gender distinction on the case suffixes in the plural. (Edelman, 1983: 60-61). \(-i\) is used as a feminine diminutive suffix (Strand, 2011).

4.2.8 Kohistani Shina

Kohistani Shina has a two-gender system based on biological sex, and two sets of inflectional suffixes, where the first distinguishes between singular and plural, and between nominative and oblique. The second set consists of the non-nominative cases, and most of them attach to the oblique base (Schmidt and Kohistani, 2008: 40). The two gender categories are masculine and feminine, and nouns are divided between them according to semantic and phonological criteria. Most nouns ending with \(-o, -u\) or \(-á\) are masculine, or they may be unmarked, i.e. end with a consonant. Most nouns ending with \(-i\) are feminine, and some feminine nouns are unmarked (Schmidt and Kohistani, 2008: 42-45). Each gender consists of three sub-groups, which decide the forms of the inflectional suffixes, e.g. case and number. The nominative case is unmarked and can appear with or without gender suffix, but other cases have different forms depending on the gender of the noun (Schmidt and Kohistani, 2008: 40-41).

<table>
<thead>
<tr>
<th>Nominative</th>
<th>Oblique</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>PL</td>
</tr>
<tr>
<td>M1</td>
<td>ázo</td>
</tr>
<tr>
<td>M2</td>
<td>batshoó</td>
</tr>
<tr>
<td>M3</td>
<td>zab</td>
</tr>
<tr>
<td>F1</td>
<td>ráati</td>
</tr>
<tr>
<td>F2</td>
<td>bheríí</td>
</tr>
<tr>
<td>F3</td>
<td>jip</td>
</tr>
</tbody>
</table>

Table 4. Layer 1 case suffixes in Kohistani Shina nouns (Schmidt and Kohistani, 2008: 42-47)

4.2.9 South-eastern Pashai

South-eastern Pashai nouns are inflected for number, although there is no regular plural marker (Lehr, 2014: 181), and case. The case assignment is split ergative (Lehr, 2014: 145). The language has two separate gender systems, the first based on biological sex and the second on animacy. The first system is based on biological sex and divides all nouns into two categories, masculine and feminine. Most nouns denoting male animates are masculine, and most nouns denoting female animates are feminine, although many animate nouns fit into each gender both by semantic and morphological criteria. The residue is divided into either category based on the ending. Gender is marked on the noun with \(-á\) suffix for masculine nouns, and \(-ék\) suffix for feminine forms. There is a default gender/number marker \(-i\), used when the nouns are either feminine or plural. Many nouns have corresponding masculine and feminine forms for nouns referring to animals, such as bat-ék ‘female duck’ and bat-á ‘male duck’, with some exceptions. The word pakar: ‘shawl’ is gendered in the sense that the added suffix reflects if it refers to a shawl worn by a man, or by a woman. Nouns ending in consonant have no overt gender marking, and they seem to have no other rule as to which gender they are assigned to (Lehr, 2014: 177-180).

16
The second system is strictly semantic and the nouns are either animate or inanimate. Morgenstierne (1967) argues that it is a case of a two-dimensional system illustrated in 3), while Lehr (2014: 176-177) proposes that the masculine and feminine genders are further divided into an animate and an inanimate sub-group.

<table>
<thead>
<tr>
<th>Animate</th>
<th>Inanimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>goːraː aulaː aː(sta)</td>
</tr>
<tr>
<td>‘The horse is big’</td>
<td>‘The hill is big’</td>
</tr>
<tr>
<td>Feminine</td>
<td>goːriː auli aː(sta)</td>
</tr>
<tr>
<td>‘The mare is big’</td>
<td>‘The River is big’</td>
</tr>
</tbody>
</table>

(Morgenstierne, 1967: 59)

4.2.10 Northern Pashto

Northern Pashto has a two-gender system based on biological sex, a two-way number distinction as well as direct or oblique case (Robson and Tegey, 2009: 726). The Pashto gender system is based on biological sex and has two genders; masculine and feminine and overt gender marking; the masculine suffixes are –u and –ay and the feminine suffixes –a, -o and –øy, which can be in pairs where the masculine is the unmarked, and the feminine is formed by adding a feminine suffix; parast-år ‘male nurse’ and parast-år-a ‘female nurse’ (Robson and Tegey, 2009: 726). There are three classes for each gender, the masculine classes are M1a, mostly for animates that take the –ân plural marker; M1b, for inanimate nouns with the plural marker –un; M2a & M2b, containing nouns ending in stressed or unstressed –ay, and lastly the M3 class, which contains mostly inanimate masculine nouns ending in a stressed vowel. The feminine gender is divided into three classes as well, corresponding to the masculine classes; F1 nouns ends with –a, F2 ends in –øy and F3 with nouns ending in –a, -á, -o or –i (Robson and Tegey, 2009: 727). The singular and plural markers are not affected by the gender of the noun, but by animacy. Most inanimate nouns take the plural marker –un and most animates take – ân. The form of the direct and oblique case suffixes for each noun depends on the gender and declensional class of the noun (Robson and Tegey, 2009: 728).

4.2.11 Purik

Purik has no grammatical gender. The nouns are inflected for case, definiteness and number (Zemp, 2013: 109). Even though it has no grammatical gender, the strategy present in Purik to mark nouns depending on the refersents biological sex seems to have been lost almost completely. Some nouns, adjectives and adverbs ending in either –pa, -ba or –ma have had a sexus distinction where the plosive has been used to indicate masculine and the nasal to indicate feminine gender (Zemp, 2013: 118)

4.2.12 Wakhi

Wakhi does not have a gender system. Wakhi nouns are inflected for number and case, but there are only some traces left in some word forms of an otherwise lost gender system (Bashir, 2009: 828)

4.3 Gender targets

This section is devoted to gender agreement targets in each of the sample languages.
4.3.1 Burushaski

In Burushaski the four different genders of the noun trigger agreement on adjectives. Even though adjectives can modify their head noun without any affixes, their plural suffixes depend on the gender of the noun they modify (Yoshioka, 2012: 83). Demonstrative adjectives, which have a two-way distinction of proximal or distal, and interrogatives, function as adjectives and agree with the noun. When the head noun is human, regardless of sex, the demonstratives and interrogatives agree with a merged human category in both singular and plural; resulting in agreement with either H, X or Y class genders (Yoshioka, 2012: 75). The numerals 1-10 change depending on the gender of their head noun. The copula verb has different forms for human (both the masculine and the feminine human gender merge into one category with the copula) or non-human subjects. In the Nager variety the second group is divided further, and the copula has separate forms for human, concrete and abstract nouns. The copula verb is used to make nouns, adjectives or numerals appear in the predicate (Yoshioka, 2012: 32). The verb agreement rules in Burushaski are complicated, but the simple core of the system is that the verb has an agreement suffix, that agrees with the subject of a transitive clause, and a prefix, that agrees with either the indirect object of a transitive clause, when there is one, or with the direct object or the subject of some intransitive verbs.

4) Gús-e hilés-isho curük ó-t-umo
woman-ERG boy-PL/ABS cut 3P-do-3SF/PAST
‘The woman cut the boys’ (Willson, 1996: 53)

In some constructions, the gender of the noun will determine if there will be an agreement prefix or not, in the following examples the Y-class noun do not trigger object agreement with the auxiliary verb manāas ‘to become’ (Willson, 1996: 25):

5) Sá lālam i-man’-ibim
sun.X/ABS shine 3SX-become-3SX/PSTPRF
‘The sun had shone’ (Willson, 1996: 26)

6) GéniSh lálam man’-ilum
gold.Y/ABS shine become-3SY/PSTPRF
‘The gold had glittered’ (Willson, 1996: 26)

4.3.2 Dameli

Gender agreement with the gender system based on biological sex in Dameli is primarily on adjectives and verbs, but also on possessive pronouns and markers. On adjectives the gender of the head noun is marked overtly with the masculine suffix –a or the feminine suffix –i. Many adjectives are invariable, and show no gender agreement with their head noun (Perder, 2013: 83-84). Compare the following examples, the first with adjectival agreement and the second with an invariable adjective;

7) Muu ta kaaya aaê-i daro māā-i šumaa-i
1SG.OBL from remembrance come-CP is 1SG.POSS-F beautiful-F
daaman ta gurum
Domel(F) of morning
‘I remember the morning of my beautiful Domel’ (Perder, 2013: 83)

8) Yee māā-Ø axiri faisala daru
this 3SG.POSS final decision.M COP.INAN.IMPVF.3
‘This is my final decision’ (Perder, 2013: 84)
The verb shows gender agreement in some tense/aspect forms, but not in the future, potential past or imperative. Possessive pronouns and the possessive marker agree in gender with the possessed noun (Perder, 2013: 53).

The Dameli auxiliary verb ‘be’ and possessive pronouns show agreement with the gender system based on animacy and distinguishes if its subject is animate or inanimate. There are different stems for the copula, depending on the animacy of the noun. They are high in frequency and used both as a copula and as an auxiliary (Perder, 2013: 121). The animate copula has two stems, t- and b- and is inflected for person, number and feminine or masculine gender in the third person singular. The inanimate copula has two forms, the present/imperfective daru and the past/perfective beru.

9) See mač aluuna tʰ−aa
    3SG.DIST man tasteless.M be−IMPFV.3SG.M
   ‘That man is tasteless (not serious)’ (Perder, 2013: 145)

10) baaṭ buum−a daro
    stone ground−LOC be.inan.IMPFV.3
   ‘The stone is on the ground.’ (Perder, 2013: 145)

Possessive markers can distinguish animate and inanimate possessors, although it is uncommon for the possessor to be inanimate. 3rd person pronouns are either animate or inanimate and are used in a variety of ways, including personal pronouns, demonstrative pronouns and articles (Perder, 2013: 77).

4.3.3 Đomaakí

In Đomaakí the gender of the noun agrees with adjectives, some demonstratives, the numeral ‘one’ and the third person singular suffixes on the verb. The suffix of singleness (Tikkanen, 2011: 205) which Weinreich (2011: 17) calls the singulative-indefinite suffix, show gender agreement: SG.M -(e)k and SG.F -(e)ka. The combination of the plural formation (i.e. noun + plural suffix) may depend on the gender of the noun, as well as the ending of the noun stem (Weinreich, 2011: 6).

4.3.4 Gawri

Adjectives and verbs in Gawri show agreement with the noun. The verb shows split ergative agreement, depending on the tense/aspect form. In the simple past, non-specific present/future and future forms the verb agree in gender with the subject, but in the past imperfect it agrees with the object of a transitive verb or the subject of an intransitive one. There is a number distinction in the masculine gender, but the feminine has the same form for both singular and plural. In examples 11) and 12) the nouns trigger agreement on the adjective and verb.

11) Ami dü gän pätîlî mēz râ thu
    these two big.M pots table on are.M
   ‘These two big pots are on the table’

12) Ami dü gän kōrî mēz râ thi
    these two big.F water.pots table on are.F
   ‘These two big water pots are on the table’ (Baart, 1999: 31)

Gawri demonstratives function as adjectives and modify their head noun, agreeing in gender with it (Baart, 1999: 32).
4.3.5 Kalasha

Kalasha adjectives are invariant (Bashir, 2003: 852), verbs agree with the subject in person, number and gender (Bashir, 1988: 41). The auxiliary ‘be’ has separate forms for animate and inanimate nouns. Five out of nine finite verbal forms are formed with a participle and a finite form of ‘be’ (Bashir, 1988: 41).

13) Se miSTeri bo wat kai ayik ghō-‘an
he teaching much time do.CP be.AN.INF say.P/F-3P
‘(I hear that) he has been a teacher for a long time’ (Bashir, 1988: 78)

14) krAkA uk chi šik ghō-‘an
Karakan water cut.CP be.INAN.INF say.P/F.NS-3P
‘(I hear that) the water is cut off in Karakan’ (Bashir, 1988: 78)

4.3.6 Kashmiri

In Kashmiri, some adjectives are inflected for gender, other are invariant. The genitive postposition agrees in gender with the governing noun, as well as in number and case (Koul, 2003: 909). Tense suffixes show agreement in gender with the unmarked subject or object of a clause (Koul and Wali, 2006: 83), and pronominal suffixes inflected for gender and number. The auxiliary chu ‘be’ agrees with the nominative subject in gender (Koul and Wali, 2006: 84). In the perfective aspect the auxiliary and the past participle forms of the verb in –mut ‘be’ agree with the transitive object and the intransitive subjects in gender.

15) mohnɨ voth nendri subihan šeyi baji
Mohan got_up.M.SG sleep.ABL morning.LOC six.ABL o’clock
‘Mohan got up at six o’clock in the morning’

16) tomis cha va:riya:h kɔ:m
he.DAT has.F.SG lot work.F.SG
‘He has a lot of work’ (Koul and Wali, 2006: 86)

4.3.7 Kati

In Kati, the evidence of noun gender is in agreement on adjectives, verbs and participles. Most adjectives take gender agreement, -i is the feminine marker and –a is the masculine, and some adjectives are invariant. Demonstratives distinguish gender (m/f) in the oblique case singular (Edelman, 1983: 62). Finite verbs consist of a verbal stem plus a pronominal suffix that agrees with the subject noun in gender. A participial stem must be followed by the feminine marker -i when the subject is feminine singular (Strand, 2011).

4.3.8 Kohistani Shina

In Kohistani Shina, adjectives ending with a vowel agree with their head noun in gender, while adjectives ending with a consonant are invariant (Schmidt and Kohistani, 2008: 100).

17) Ek bár-o muşāa
one big-M.SG man
‘an important man’ (Schmidt and Kohistani, 2008: 101)

18) l-ái razāal-i xoráak
very-F.SG sweet-F food
‘Very delicious food’ (Schmidt and Kohistani, 2008: 102)
Adjectival demonstratives, which distinguish between near, not-near visible, distant visible and remote referents are inflected for feminine and masculine gender in the singular but not in the plural forms. The present, past imperfective, present perfect and past perfect tense/aspect forms are formed with the auxiliary as- ‘be’, which agrees with the subject in person and gender with a singular referent. There are no gender distinctions on possessives (Schmidt and Kohistani, 2008: 58). The indefinite article –ek and the postpositions are invariant for gender, with the exception of the instrumental postposition, that makes distinctions for inanimate or human referents (Schmidt and Kohistani, 2008: 70-75).

4.3.9 South-eastern Pashai
In Southeast Pashai agreement within the gender system based on biological sex is primarily on adjectives; attributive adjectives agree with the noun they modify in gender and number. Some adjectives are invariant as in Error! Reference source not found. (Lehr, 2014: 184-185).

19) sos-av-ā  gōr-ā
   dry-STV.PCP-M  horse-M
   ‘thin horse’ (Lehr, 2014: 184)

20) paman  vasket
dirty  vest
   ‘dirty vest’ (Lehr, 2014: 185)

Nouns trigger verb agreement in some tenses, such as the perfective tenses proximate, distal and remote perfect. The agreement follows an accusative/ergative pattern as shown in 21) (Lehr, 2014: 226). Of the non-finite verb forms, the stative and imperfective adjectival participles are inflected for gender (Lehr, 2014: 240).

21) pari-y  ašpār  pač-al-e-k-en
   pari.F-OBL  rice.M  cook-CS-PXPF-M-3M.O
   ‘Pari cooked rice (just now)” (Lehr, 2014: 284)

South-eastern Pashai animacy agrees primarily with the auxiliary verb. Some postpositions can occur only with animate nouns: o:da ‘near, to, in presence of’, lamba ‘after, behind’ and puruna ‘before, in front of’ (Morgenstierne, 1967: 74). There is no animacy distinction with the possessed noun in a possessive construction (Lehr, 2014: 169).

4.3.10 Northern Pashto
Pashto adjectives agree with their head noun. The adjectives are divided into three classes depending on number and oblique or direct (nominative) case (Robson and Tegey, 2009: 731). The numeral ‘one’ distinguishes gender by adding a final –a for feminine, and the ordinal numbers are inflected for gender in the same way as adjectives in class 1b (Robson and Tegey, 2009: 737). Pashto has a three-way demonstrative system, and when used as attributive adjectives, demonstratives agree in gender with the head noun. In singular direct form there is no gender distinction, but in the singular oblique form and the direct plural form the end vowel changes depending on the gender. The verbs agree in number and gender in third person with the subject or object depending on the tense/aspect form. Three copula verbs do not show gender agreement; the existential sta, the subjunctive wi and the conjectural and optative way (Robson and Tegey, 2009: 738-739).
5. Discussion

As we saw in the previous section, grammatical gender is present in ten out of twelve languages in the sample and the assignment systems and target genders have been reviewed. This section will focus on the patterns that emerged, and the results will be compared and some explanations suggested. An overview of the general patterns concerning genders in the sample languages is provided in 5.1, in light of what is known about their genealogical relationships. Section 5.2 is devoted to the discussion of other possibilities for the languages that do not fit into the genealogical expectations, and reason about the emerged patterns in light of some of the areal typological work by Koptjevskaja-Tamm (2010) and Nichols (2003). In 5.3 and 5.4 two special cases are discussed in more detail, both from a genealogical and areal-typological perspective. In 5.5 contains a review of the method of analysis used for this thesis, and some suggested improvements. In 5.6 a few proposed topics for further research will be presented.

![Figure 4. Summary of the distribution of grammatical gender in the language sample (Copyright © 2016 by Henrik Liljegren)](image)

5.1 Genealogical distribution

All Indo-Aryan languages in this sample, except for Kalasha, have a gender system based on biological sex. They all have a masculine/feminine distinction and somewhat different strategies for assigning the residue into either gender. The most common strategy for these languages to assign their residual nouns is due to the ending of the nouns. Nouns ending in –i in Dameli and Kohistani Shina are feminine, and in South-eastern Pashai it functions as the default feminine/plural marker. In Đomaaki the feminine suffix is –a and the marked feminine in South-eastern Pashai is –ek. Most of the residual nouns that are assigned to the masculine gender end in –u or –d in Đomaaki, –a in Dameli, –o, –u or –d in Kohistani Shina and in –ā in south-eastern Pashai. At last, both Kohistani Shina and South-eastern Pashai have a number of unmarked nouns ending in a consonant that have no general rule for assigning them to either gender. Gawri and Kashmiri have no general phonological or morphological rules to assign the residue into either gender. A number of complicated assignment rules can be
observed in Kashmiri; including palatalization, additional vowels, consonant changes and sometimes suppletion to form feminine nouns of default masculine stems. Overall, it is most common for the feminine gender to be the marked gender and that nouns are assigned depending on the final vowel. Formal assignment rule seem to favour the ending –i for feminine nouns.

In the Indo-Aryan languages (including Domaaki) most adjectives agree with the head noun, even though some adjectives seem to be invariant in all languages, primarily those ending with a consonant. Verb agreement is also a shared feature within the HKIA, Pashto and Kati, while it is present only in agreement with third person suffix on the verb in Đomaakí. The agreement is largely dependent on the tense/aspect form of the verb. The noun can agree with the verb in some tenses but not in others, or the tense/aspect form affects which argument (i.e. the subject or object) the verb will agree with. Due to split ergative agreement languages can differ in which argument triggers agreement on the verb. This may also vary between the tense/aspect forms within each language.

Burushaski is a language isolate, and so far there has been no evidence for any genealogical relationships with any language that is still spoken today. The inherent categories of human masculine, human feminine, animals and fruits (X-class) and abstract nouns (Y-class) trigger agreement on adjectives, demonstratives, interrogatives and numerals, as well as verbs. A transitive verb agrees with the subject by a suffix and the object by a prefix. The distinction between human and non-human entities seems to be the primary division in the gender system. The HF and HM class often merge into a single human category and differ from the non-human classes in agreement depending on the agreement target and the case and number of the noun. Burushaski has a complicated system of agreement. Some adjectives agree with their head noun in gender, others are invariant. The plural suffix attached to the adjective is inflected for gender, as well as demonstratives, the numerals 1-10, interrogatives and verbs. Verb agreement depends on several factors and in a transitive clause both the subject and object can be marked on the verb, with several exceptions depending on the semantics of the verb, the nature of the arguments and other factors (for a detailed description see Willson, 1996).

In Tibeto-Burman languages there is no inherited gender, and Purik has none. Although Purik has borrowed features from Indo-Aryan languages, for instance the adjective-noun word order, which has replaced the inherited Tibeto-Burman noun-adjective word order (Dryer, 2008), gender is not easily borrowed and has not been implemented in the Purik grammatical system.

5.2 Gender and areality

The geographical distribution of gender features in the language sample points towards a rather stable situation in the south and east, while the languages that behaves differently from what are expected in terms of their genealogical relationships are located in the northern and western parts of the GHK area.

Burushaski was once spoken in a larger area than it is today and evidence of that assumption are the substratal effects in the Shina variety of Gilgit, and perhaps further. Tikkanen (1988: 305) will not go as far as to claim that there is an early proto-Burushaski substratum in every language in the Hindukush and Pamir. There are some innovations on several levels in the Indo-Iranian languages that suggest both early and later substratum effects from Burushaski, or a language related to Burushaski earlier spoken in the Hindukush area.

Kalasha has an inherent semantic division of nouns into either animate or inanimate gender. Dameli and South-eastern Pashai nouns are also divided into animate or inanimate categories solely based on
their semantics, parallel with the gender system based on biological sex. The three languages have one gender target in common, the auxiliary.

5.3 Loss of gender

Wakhi, which belongs to the Pamir branch of the Iranian language family, has lost its gender system completely, and there are only traces left in some word forms. This is consistent with the development in other Pamir languages that either have lost their gender completely, or preserved it in a limited number of targets, such as smaller groups of adjectives or some tense/aspect forms of the verb (Edelman and Dodykhudoeva, 2009: 780-781). Edelman and Dodykhudoeva (2009: 775-777) explain it in terms of substratum effects, due to the fact that the languages spread into the Pamir area. Older Pamir languages were already spoken there at the time, resulting in innovations in the new languages and shared features on every level, including grammatical categories.

Kalasha, and its closest neighbour and relative, Khowar have both lost their inherited gender (Bashir, 1988: 40-41). Two reasons for this development suggested by Bashir (2007) are the long-lived influence of Iranian languages and language contact with Burushaski. Today Kalasha is in contact with Wakhi, the Iranian Pamir language mentioned above. It is possible that the language contact with the Pamir languages that have lost their gender systems have influenced Kalasha and Khowar, and that they influence Dameli in the same way. The substratal effects that caused the Pamir languages to lose their gender might also be one of the reasons as to why Kalasha and Khowar, and now Dameli, are losing their inherited gender systems. In Dameli the gender system based on biological sex seems to be weakened, since the speakers appear not to be certain of the assignment rules for inanimate nouns, especially those lacking overt gender marking (Perder, 2013: 52-55). Perder (2013:55) suggests that this might be one step in a process from the Old Indo-Aryan three gender system towards a system based solely on an animacy distinction, such as the one in Kalasha and Khowar.

5.4 Acquired gender

In this part the languages proposed to have a grammatical gender system based on animacy will be discussed. The possibility of subgenders and over-differentiated targets proposed by Corbett (1991) will be taken into consideration, with the aim to fit these both features into the same gender system. The languages that have the animacy distinction in combination with a gender system based on biological sex will be compared to other languages in the world that is claimed to have the same double or (partly) parallel gender systems; Paumarí (Aikhenvald, 2010) and Mayali (Evans, Brown and Corbett, 2002).

Central to this discussion is the auxiliary and the copula. The auxiliary is a special type of verb that is used to help other verbs and it adds extra grammatical information to a clause, such as tense or aspect. The verbs often used as auxiliaries are ‘be’ as in the English example was driving, or forms of ‘become’. A copula verb connects parts of a clause without any additional information the car is red or I am tired. It is common for the same verbs to be used both in copula and auxiliary constructions.

In a couple of languages in the area there is a grammatical gender system based on animacy. This phenomena is present in three different languages in the sample, in Kalasha (as well as in Khowar, which is not in the sample, cf. Bashir, 2007), in Dameli, and in Pashai which provides the most interesting examples because of its ergative agreement.
In Kalasha the traditional Indo-Aryan gender system based on biological sex has been lost. Instead, a gender system based on animacy has developed. This system is different from the ones present in many other languages because it does not have adjectives as one of the target genders, but the auxiliary verb ‘be’, which distinguishes between animate and inanimate subjects of the clause. In 22) the subject triggers animate agreement, and in 23) animate agreement is ungrammatical since the subject dāur ‘house’ is inanimate.

22) ta tat kūra asuír/*šeér?
   your father where be.AN/*be.INAN
   ‘Where is your father?’ (Bashir, 2007: 213)

23) ta dāur kūra *asuír/šeér?
   your house where *be.AN/be.INAN
   ‘Where is your house?’ (Bashir, 2007: 213)

In some cases this opposition between animate and inanimate is visible within the same clause, as in 24) where the brain is inanimate and ‘he’ is animate.

24) ása mhásta ne shi-álLa khalí ásta
   3S.OBL.DIST brain not be.INAN-PST.1.SG empty be.AN.PRS.1.SG
   ‘He didn’t have a brain, he was empty’ (Heegård Petersen, 2006: 25)

A similar system is present in Dameli, where the gender targets are the auxiliary verb ‘be’, demonstratives and possessive pronouns and markers. The system distinguishes between animate or inanimate subjects. In Dameli, the traditional Indo-Aryan gender system is active parallel to the animacy distinction, and nouns are divided into genders depending on biological sex, and the residue due to phonological criteria as well. The possessive markers distinguish animate and inanimate referents; sāā for animate possessor and ta for inanimate possessor, although it is rarely used.

Animacy is distinguished in third person pronouns (Perder, 2013: 73), and they function as personal pronouns and demonstrative pronouns, and can be used together with nouns as modifiers as well. With nouns they are used as determiners or articles and add extra information to the clause as they agree in animacy, number and distance with their head noun, as in 25) and 26) (Perder, 2013: 77).

25) yee tʰaan-a gram beru tara
   3SG.INAN place-LOC village COP.INAN.PST.3 there
   bai-tʰun
   COP.AN-INDIRPST.3PL
   ‘In this place was a village; there they were’ (Perder, 2013: 124)

26) mane manee dac-i-na tee see tasāā-Ø
   Like this see-IMPFV.3SG.M that 3SG.AN.DIST.NOM 3SG.POSS-M
   žami tʰui
   wife COP.AN.IMPfV.3SG.F
   ‘Thus he sees that it is his wife’ (Perder, 2013: 53)

In Pashai, there is agreement with the inherent masculine/feminine gender of the noun, and with the animate/inanimate categories. The traditional Indo-Aryan gender system on the one hand, and an animacy based gender system on the other. Lehr (2014: 176-177) argues that the masculine and feminine genders can be further divided into feminine animate, feminine inanimate, masculine animate and masculine inanimate. Since the masculine/feminine distinction has in part different agreement targets than the animacy distinction, it may be two different systems. Morgenstierne (1967) proposed a two-dimensional gender system, illustrated as follows:
The auxiliary occurs in either non-past or past tense. In the present (non-past) tense, the masculine and feminine distinction is only displayed on the auxiliary in third person singular referents, and the animate and inanimate distinction with all referents. In the past tense, the auxiliary agrees in gender in first, second and third person singular, with no animacy distinction.

Regardless, animacy agreement appears to be a rather limited phenomenon, since it occurs only with the auxiliary, in present tense (Lehr, 2014: 177).

<table>
<thead>
<tr>
<th>Masculine</th>
<th>Animate</th>
<th>Feminine</th>
<th>Inanimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>go:ra aula a:(sta)</td>
<td>dha:ri aula ši:k</td>
<td>go:ri aula a:(sta)</td>
<td>nandi aula ši:k</td>
</tr>
<tr>
<td>‘The horse is big’</td>
<td>‘The hill is big’</td>
<td>‘The mare is big’</td>
<td>‘The River is big’</td>
</tr>
</tbody>
</table>

(Morgenstierne, 1967: 59)

In the examples above, the object of the transitive clause triggers agreement with both the main verb and the auxiliary, in 27) the masculine noun ketāb ‘book’ triggers agreement on the main verb, and inanimate agreement on the auxiliary. In 28) the noun pelek ‘cup’ triggers feminine agreement on the main verb, and inanimate on the auxiliary. In the third example, 29), the object noun kelā ‘boy’ is animate and masculine, which triggers masculine agreement with both the main verb and the auxiliary, as well as animate agreement only on the auxiliary.

### 5.4.1 Subgenders and over-differentiated targets

Corbett (1991: 161-175) suggests some alternate explanations, with the aim to reduce the number of genders in an analysis of a language. Two proposed explanations are subgender systems and overdifferentiated target genders.

Subgenders are different categories within a gender that behave differently in some environments, e.g. in a few case forms, or in the plural. Corbett uses the example of Serbo-Croat for the analysis of situations with subgenders. In Serbo-Croat the genders are feminine, neuter and masculine. The latter is divided into two subgenders; masculine animate and masculine inanimate, because of the behavior in the accusative singular case attributive agreement markers (Corbett, 1991: 161-162). Because this distinction is only in the singular, not in the plural, as well as in the accusative and no other case, this is what Corbett explains as “… agreement classes which control minimally different sets of agreement, that is, agreements differing for at most a small proportion of the morphosyntactic forms of any of the agreement targets” (Corbett, 1991: 163). In my sample, several languages have groups of nouns that behave differently in plural, or in some case forms. One example is the Burushaski Y-class...
and Z-class, where the Z-class behaves in the same way as the Y-class in every instance except in agreement with numerals and in the genitive case (Yoshioka, 2012: 33). For this reason Burushaski has been treated as a four-gender language with one subclass.

In Kalasha the animate and inanimate nouns have the same nominative and genitive/oblique singular forms, but distinct forms in all other cases. In Dameli the accusative agreement makes the inanimate agreement more unusual, because of the semantic order that mostly animate entities are agents, and therefore more often subjects in a clause. Even though the inanimate agreement is not as frequent as the animate one, this distinction is not affected by number or case forms. The same can be said about south-eastern Pashai, the animacy distinction is not limited to any particular case or number.

The analysis of overdifferentiated targets (Corbett 1991: 168-170) aims to reduce the number of genders, as did the analysis of subgenders. An overdifferentiated target is restricted to one word-class and within the word-class there should be further limitations, such as the agreement only applying to certain words. It can be argued that this description fits the situation in Kalasha, Dameli and south-eastern Pashai since their primary target is an auxiliary verb. In south-eastern Pashai the copula(auxiliary verbs are at least two, where ‘be’ distinguishes animate and inanimate and ‘become’ is invariant (Lehr, 2014: 255-270). The other restriction for overdifferentiated targets is that the genders are not clearly distinguished on other targets. In Dameli other targets distinguish between animate and inanimate nouns, such as possessives and third person pronouns.

The claim that a language has two, at least partially, parallel gender system in use is not unique to this area. Among the more well-known examples are Mayali and Paumarí. Mayali is sometimes mentioned in the discussion of languages that have two parallel gender systems, one which divides nouns into four categories; masculine, feminine, vegetable and neuter. These categories overlap almost completely with four morphological classes that display themselves on the nouns as prefixes. There is a fifth morphological class, which is Ø-marked and contains some residual nouns and English loan words that complicate the correspondence between gender and class. All and all, the four genders trigger agreement on modifiers, although masculine is used as a default marker. The morphological classes do not trigger any agreement on any other part of a clause, which by definition is the main criterion for a gender system (Evans, Brown and Corbett, 2002: 134-136). Mayali has one gender system, and one of nominal classifiers. Paumarí on the other hand, has two parallel systems, which Aikhenvald (2010) refers to as ‘noun class’ and ‘gender’. Both divide nouns into categories depending on the rules within the system, and both systems agree with the verb but in two different ways. Even though Aikhenvald (2010: 250-251) argues that the gender system is more stable than the noun class system, both fulfil the criteria for a gender system set up by Corbett (1991). One of these two languages has a similar (and more advanced) parallel gender system as the ones in HKIA. Even though this is not a unique phenomenon, it might be uniquely clustered.

The animacy distinction is used to some extent in Kalasha, Dameli and Pashai. The most interesting cases are when both the animacy distinction and the biological sex system show within the same clause, or even within the same word. In Pashai, due to its ergative agreement, this distinction is more pervasive than in Dameli and Kalasha. What could be rather uncommon is the potential areal factor, and that this phenomenon seems to be somewhat clustered in the western and northern parts of the Hindukush area. It is also quite fascinating that it is not solely a case of one gender system replacing another (at least not at the moment, one could only speculate about the future), as in Kalasha and Khowar (Bashir, 2007: 208-210).
5.5 Method and analysis discussion

The data collected for each of the languages were sufficient, except for Kati, where there were no grammar or language description but a website which was hard to navigate and did not provide all the details that I needed. Regarding Northern Pashto, the description available was not specifically of the northern variety, but in the case of such a stable feature, the differences between the dialects can be expected to be rather small.

5.6 Areas for further research

I propose a more detailed research of the languages of the Greater Hindukush, regarding the category of gender, and especially cases of suspected parallel systems. Animacy distinctions are not uncommon, but the existence of two parallel, partly overlapping systems in one and the same language might be. It is an interesting aspect to add to the already fairly well-described noun gender research and typology that exists. And in particular, it would be interesting to see if more languages in the area have them, making the findings here more certain and proving the property more clustered than we may suspect today.
6. Conclusions

The purpose of this study has been to investigate the feature of grammatical gender in the languages of the Hindukush area. Ten out of the twelve languages in the sample had some type of grammatical gender(s), and Wakhi, one of the two languages without gender systems turned out to be valuable in the discussion on loss of gender, since it is possible that it influenced Kalasha in the process of losing the inherited Old Indo-Aryan gender system. Instead Kalasha grammaticalized animacy distinction, a system that in terms of Corbett’s criteria for gender is a true gender system. This animacy distinction was found in two other languages in the sample as well, Dameli and south-eastern Pashai. The inherited old Indo-Aryan gender system in Dameli might be on decline but at the moment two parallel systems are active in the language. in south-eastern Pashai the two gender systems are active parallel to each other, and the ergative agreement makes the animacy distinction more central than in Kalasha and Dameli.
Bibliography


