A Grammar and Dictionary of Indus Kohistani
Volume 1: Dictionary

Trends in Linguistics
Documentation 21-1

Editors

Walter Bisang
Hans Henrich Hock
Werner Winter
(main editor for this volume)

Mouton de Gruyter
Berlin · New York

A Grammar and Dictionary of Indus Kohistani

Volume 1: Dictionary

by

Claus Peter Zoller

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter & Co., Berlin.

Library of Congress Cataloging-in-Publication Data

Zoller, Claus Peter.
A grammar and dictionary of Indus Kohistani / by Claus Peter Zoller.
p. cm. − (Trends in linguistics. Documentation ; 21-1)
In English and Indus Kohistani.
Includes bibliographical references and indexes.
ISBN 3-11-017947-4 (hardcover : alk. paper)
1. Indus Kohistani language − Grammar. 2. Indus Kohistani language − Dictionaries. I. Title. II. Series.
PK7045.M3Z65 2005
491.4'99−dc22
2005012901

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

ISBN 3-11-017947-4

Contents

Acknowledgements

This grammar of Indus Kohistani is the result of a project conducted within the framework of the joint Pakistani-German research project Cultural Area Karakorum. The research project was financed by the Deutsche Forschungsgemeinschaft (DFG, German Research Council) and by the Gesellschaft für Technische Zusammenarbeit (GTZ, Society for Technological Cooperation). My thanks go to both, especially to the DFG for its generous funding of my research over an extended period of four years.

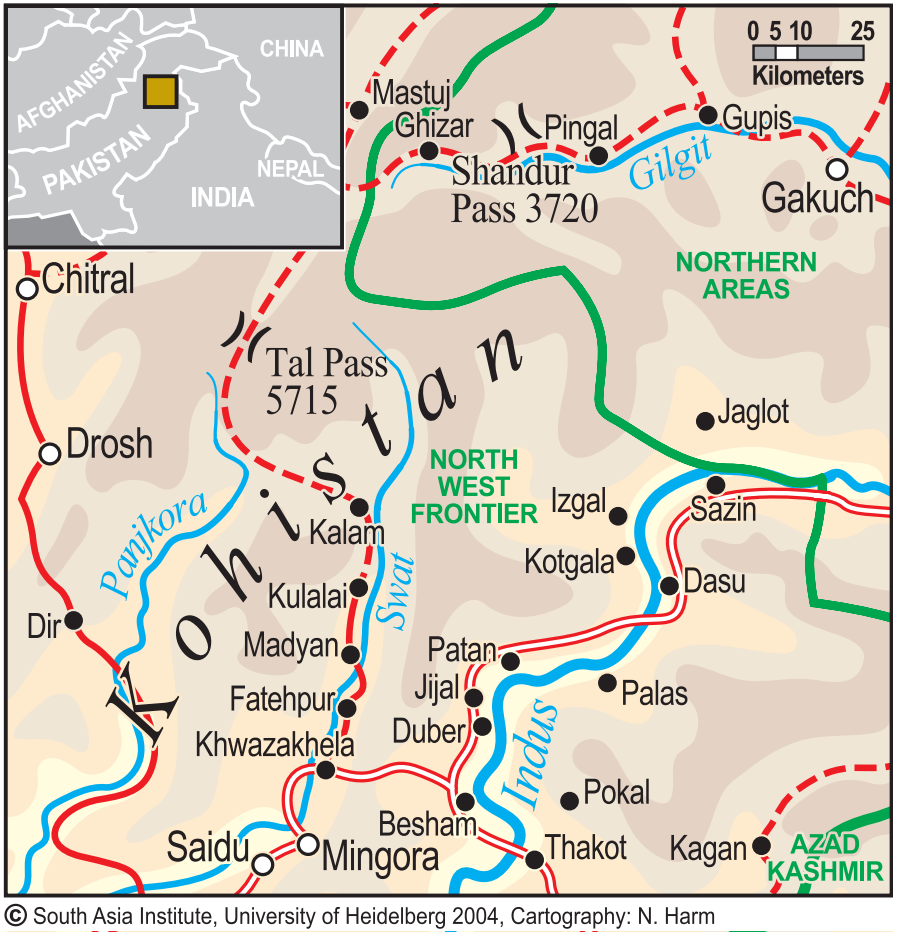
Being a guest of the people of Indus Kohistan turned out to be an exciting and challenging experience. In the first place, this was due to the circumspection and dependability of my main language consultant Khangir Khan from the village of Jijāl. Our work on the grammar was, however, significantly supported by his brother Zaman Khan and by Juma Khan from Jijāl. We received additional help from Sher Mohammad Khan and Sher Afzal from the village of Basīn near Gilgit, and from Yar Mohammad and Mashruf Khan from the village of Pari Bangla below Gilgit. My thanks also go to Akhtar Khan, main language consultant for Bhaṭīse (Baṭeṛa), to Muhammad Zaman from the village of Hil above Batgram, language consultant for Gabār (Gowro), and to Sarwar Khan and Muhammad Ghulam, language consultants for Shāṭōṭī. Muhammed Manzar Zarin from Rawalpindi was the good spirit in the background who secured the smooth continuation of our work.

Working with the musicians and storytellers of Indus Kohistan was always especially pleasurable for me. I am thankful to the singer Hazrat Ali and the Dom Sidak, both from the village of Jijāl.

Georg Buddruss – the head of this project, mentor and friend – introduced me to the Dardic languages. He was my consultant whenever I faced a special problem, and his unerring judgement saved me from various wrong interpretations of the data collected. Any remaining errors are, of course, exclusively my responsibility. I also received assistance, important suggestions and technical support for the speech analyzer from Joan L. G. Baart and Carla Radloff in Islamabad. During decisive periods of the project I received help from Irmtraud Stellrecht and from members of the German Embassy in Islamabad.

It is thanks to Werner Winter that this grammar can be published at Mouton de Gruyter. I am especially grateful to him for undertaking the painstaking task of proofreading a dictionary. Over many months Christoph Eyrich helped

me with much expertise and patience in bringing the LaTeX computer files of this grammar into the present shape. My thanks go also to Birgit Sievert and Wolfgang Konwitschny from Mouton de Gruyter. Nils Harms prepared the map of this book. Despite this broad support, the project could hardly have been completed without my wife and children, who decided to accompany me to Pakistan. It was not an easy decision for them, but they did not regret it.



Indus Kohistani language area to the west of the river Indus

Chapter 1
Introduction

1. Indus Kohistani

Indus Kohistani is a major language of the Dardic group of Indo-Aryan languages. It is spoken in District Kohistan of the North-West Frontier Province along the west bank of the Indus and in the side-valleys leading from the west into the valley of the Indus. The language is spoken by around 220.000 people (Hallberg 1992). Aside from large communities of Indus Kohistani speakers in the big cities of Pakistan, there are two language enclaves outside District Kohistan: one in the Tangir valley and called *Kanyawali* by Buddruss (Buddruss 1959) and the other, discovered by this author, was formerly spoken in the village Šāṭōṭ in the Rondu gorge. The speakers of *Šāṭōṭī* are now, however, dispersed over several villages below Gilgit as a result of religious clashes in 1988.

To the south and south-west of Indus Kohistani, Pashto is spoken. Pashto is also the old lingua franca all over northern Pakistan but is losing now its importance due to the increasing impact of Urdu. There are several villages with Pashto speakers within Indus Kohistan in the Bankhaṛ settlement area, and one in the valley of Dubēr.

The east bank of the Indus forms the western boundary of the Shina speaking area. It is also the home of three small languages, the number of which varies between a few hundred and a few thousand. They are Bhaṭīse (or Baṭeṛa in Pashto) in the south opposite Besham, Gabār (or Gowro) opposite Jijāl, and Chilisso in the Jalkōṭ village area. All three are heavily influenced by Indus Kohistani and Kohistani Shina. They are possibly language enclaves of other Dardic languages yet to be identified. In fact, the southern mountains between the Indus in the east and the Kunar valley in the west are so to speak sprinkled with numerous small language enclaves (see below 1.7.).

All over District Kohistan, in fact over the greater part of northern Pakistan one finds the community of the Gujars. The Gujars speak an Indo-Aryan language akin to Punjabi. Traditionally they have been following a semi-nomadic lifestyle, but many of them have meanwhile settled down in villages in the upper reaches of the valleys.

1.1. Language name

The speakers of Indus Kohistani have actually no name for their language. I follow here Hallberg and Hallberg (1999) in calling the language *Indus Kohistani*. This designation has the advantage of indicating a certain linguistic closeness to Kalam Kohistani (i.e. Gawri, spoken in the upper reaches of the Swat valley) and Dir Kohistani (spoken in the upper reaches of the Panjkoṛa valley). Moreover, speakers of Indus Kohistani accept this designation in an English-language context. Whereas the dialect forms spoken in the upper Dir valley are linguistically very close to Kalam Kohistani, the differences between the latter and Indus Kohistani are considerable, and it has yet to be investigated whether the two derive historically from a common Proto-Kohistani or whether the similarities are due to diffusion.

When asked for the name of their language, the Kohistanis sometimes say that it is called *kōstʌ̀ỹ* (lit.: '(language) of the mountaineous region'). This, however, sounds artificial and is virtually the same as *kostyő* used by the speakers of Shina on the east bank of the Indus when asked the same question (Schmidt and Kohistani 1998) or *Kohistani* for Gawar-Bati (Cacopardo and Cacopardo 2001: 231). Biddulph's claim (1880: 12) that the Indus Kohistani-speaking people "call themselves Mayon" (in later literature the word appears in the form of Grierson's designation *Maiyã̄* [1919]) could not be verified by later researchers (see Hallberg and Hallberg 1999: 2). The same holds true for Leitner's *Shúthun* (see Hallberg and Hallberg loc. cit.). When the Kohistanis speak in an informal way about their language, they use expressions like "our language", "the language of this side of the river", etc. It is also common practice to name one's language according to one's settlement area. Thus, Jijālī is the variety of Indus Kohistani spoken in the Jijāl settlement area, and Seois is the variety of Indus Kohistani spoken in the Seo settlement area. The same seems to hold true for many other dialects of the Kohistani sub-group. Perhaps it must be seen in connection with the fact that hardly any proper names of rivers and mountains are in use in Kohistan.

1.2. The term Dardic

The first person of modern times to make reference to the Dards and Dardic language was Izzet Ullah in the *Quarterly Oriental Magazine* of Calcutta in 1825 (Clark 1977: 329f.). Ullah was an assistant of William Moorcroft, an

explorer and veterinary surgeon of the East India Company, and it seems that for him the Dards were a people living in the Astor or Chilas area. Some decades later, however, the use of the terms Dard and Dardic was greatly expanded by G. W. Leitner to include peoples and languages of a much larger area. This was done in scholarly reverence to Sanskrit, Greek and Roman sources (Jettmar 1975: 19) by assuming that Latin *dardæ* and Sanskrit *darada* both referred to the same people(s) living in the mountains to the west of Kashmir. There have been claims that the word Dard was still used in the 19th century by Kashmiris to designate their north-western neighbors (Jettmar loc. cit., see also Biddulph 1880: 157 and Drew 1875: 393–461). Bellew (1880: 90) writes about the Dadikai of Herodot, "Others, again, have considered them to be represented by the hill people located north of the Gandarians, and formerly called darada, a name which is still known to, but not in common use amongst, that people, though it is still the patronymic of the natives of Chilas, on the other side of the Indus, who style themselves Dárd." Moreover, Biddulph (1880: 12) claims that the Indus Kohistani-speaking people on the right bank of the Indus "apply the name of Dard to the people living on the left bank of the river." Although I asked many times many people in different areas of 'Dardistan' whether they knew the word, the answer was always negative.

The scholarly background of this term is confused by van Driem (2001: 1079) who first correctly points out Grierson's term Paiśācī (which Grierson used as a cover term both for Nuristani and Dardic) but then draws the surprising conclusion that "Grierson's term Dardic still survives" (loc. cit.). Equally irritating is van Driem's statement that "The Dards had fled into the mountains of Kashmir in face of the Kuṣāṇa incursions in the first century, were exposed to Buddhism during the Kuṣāṇa period until the coming of the Huns in the Vth century" (ibid.). The syntax of his sentence is wrong, the whole statement is made up, and Buddhism (and Hinduism) flourished in the area even after the coming of the Huns.

1.3. Dialects of Indus Kohistani

There are two main dialects. One variety is spoken in the Dubēr and in the Khandiā valley,[1] and the other along the west bank of the Indus and in the valleys in-between and in the Bankhaṛ settlement area which branches off from the Dubēr valley near its lower end. The main difference between the two di-

alects consists in the different historical development of the Old Indo-Aryan (OIA) affricates. Whereas the variety of the Dubēr and Khandiā valley has preserved them as palatal affricates, there is a strong tendency to pronounce them as dental affricates in the other variety. For instance, Vedic *catúraḥ* 'four' is pronounced as *čōr* in the Dubēr and Khandiā valley but as *ċλur* in the Indus valley. According to Leitner (1893), the former variety is called Bunzári and the latter Mani, and according to Jettmar (1983) and Fussman (1989a: 49) they are called Manzari and Mani. These designations could not be confirmed by later researchers, but Hallberg and I were told independently from each other by some speakers that the names actually refer to two clans. However, a survey of the clan names of Indus Kohistan conducted by one of my language consultants (see appendices) does not contain any of these names. If one recalls what I said above about language designations, it appears very unlikely that speakers of Indus Kohistani use dialect designations based on tendencies of historical phonology. The issue is again confused by van Driem who formulates, "The *Maiyā̃ are* also known as the 'East Kohistani', the Mani *or* the Manzari" (2001: 1095, italics by this author). He uses Mani and Manzari as synonyms, and his use of *Maiyā̃* in the plural was apparently motivated by his unconscious equation of the final vowel of this word with the Hindi plural of feminine nouns ending in *-ī*.

The variety of Indus Kohistani spoken in the Indus valley contains also a substantial number of words with affricates which have not undergone the aforementioned sound shift. As a rule, they are not borrowings from neighboring Shina (where this sound shift occurs quite rarely). The appropriate context in which this sound shift has to be seen is, indeed, not the comparatively small area of Indus Kohistani but the whole of northern Pakistan and eastern Afghanistan. The historical implications of this depalatalization tendency are discussed in detail in the grammar, but the issue will be taken up once more below (see 1.9.).

There are only small dialectal variations within the two main dialects (see the word list percentages matrix in Hallberg and Hallberg 1999: 7). It is, however, a fact that despite the apparent minimal differences, speakers of one settlement area are able within minutes to locate the home of speakers unknown to them of the same dialect variety but from a different settlement area.

Both the varieties of Indus Kohistani spoken in Tangir (Kanyawali) and south of Gilgit (Šāṭōṭī) closely agree with that spoken along the Indus. There are some indications that the ancestors of the speakers in Tangir hail from the

Bankhaṛ settlement area in the south of Indus Kohistan. There is, for instance, the name Bankhaṛī for the Kohistani village in Tangir which is practically the same as that of the Bankhaṛ settlement area. According to their own tradition, they migrated to Tangir in the 19th century. On the name Kanyawali see below the section 'language setting'. The ancestors of Šāṭōṭī left their ancient home probably earlier than those of Tangir. This is suggested by two facts: Šāṭōṭī has abandoned the Indus Kohistani system of pitch accents and taken on the Shina system (that is why the speakers of Indus Kohistani say that Šāṭōṭī sounds "funny") which Kanyawali has not done;[2] Šāṭōṭī lacks the 'new' phonemes *q* and *ɣ* (typically found in loans from Persian and Arabic) although they were extant in Indus Kohistani already in the 19th century (Leitner 1893).

1.4. History of research

The first word list containing Indus Kohistani was published in 1893 by G. W. Leitner. Although it doesn't correspond to modern linguistic standards it is remarkable in showing that the language, at least the pronunciation of its vocabulary, has hardly changed since then. The book on "The tribes of the Hindoo Koosh" published already in 1880 by John Biddulph, the first political agent of the British crown in Gilgit, contains short lists from two of the three aforementioned language enclaves in Indus Kohistan, namely on Chilisso and Gowro. Biddulph had not visited the area and gained his knowledge from informants in Gilgit. The same holds true for the data published in Grierson (1919, VIII/2, p. 522ff.). The first westerner to enter the Indus Kohistani language area was Aurel Stein in 1941 who, however, did not collect linguistic data. This was again done by Fredrik Barth who published together with Morgenstierne a word list (1958) which, however, contains many mistakes. The first modern in-depth study is Buddruss' monography on Kanyawali (1959). I had opportunity for a very moving encounter with his then language consultant Mohammed Hussain in September 1998 in the village Bankhaṛī. I used the opportunity to recheck the data collected by Buddruss. Although fourty years had passed since then, I could not detect any significant changes. Since Buddruss could spend just one week in the field, his inventory of phonemes for Kanyawali is not complete (see Hallberg and Hallberg 1999: 19, fn. 26). Word lists and some text specimens were subsequently published by Hallberg (1992). These were followed by a monography on the Indus Kohistani dialect of Seo (Hallberg and Hallberg 1999).

1.5. Own research

I began working on Indus Kohistani in May 1997, and my field research ended in March 2001. Through the good offices of Ruth L. Schmidt from Norway I came in contact with Manzar Mohammed Zarin in Rawalpindi. Zarin, a highly respected expert on Kohistani Shina and the regional traditions of District Kohistan, hails from the Jalkōṭ settlement area (where Shina is spoken), but has close contacts also with speakers of Indus Kohistani. He accompanied me on my first trip to Indus Kohistan. These were tense days as I had heard many deterrent stories about 'Yaghestan' ('land of the independent ones') and its bloodfeuding inhabitants, and how other researchers had to flee from the area. Colleagues also warned me that my chances for a successful completion of my research were not too good. However, with the help of Zarin's diplomatic skills I was able to build up good and lasting contacts with many Kohistanis. Quite a number of them have become my close friends.

Indus Kohistan has been called the most difficult corner in the world for travel (Jettmar 1975: 22). Before the construction of the Karakorum Highway (1966–1978), people travelling along the Indus from one settlement area to the other had to follow paths at dizzy heights which were too dangerous to be used by animals. There were sections where the travellers had to insert rods into crevices as kind of ladders. With the Highway travelling has become easier, but field research is still not without hardship. Entering the side-valleys with the jeep or on foot is still a risky enterprise, but an enterprise leading through overwhelming landscape.

On my first journey I visited all the side-valleys of Indus Kohistan and introduced myself to the local chiefs, and I met Khangir Khan from the Bhadar clan of Jijāl who became my principal language consultant till the end of my field research. Khangir Khan turned out to be a prudent and conscientious consultant who, moreover, was able to sit almost endlessly together with me when we worked with the speech analyzer over long months on the many tricky problems of the Indus Kohistani pitch accent system.

I began my research on Jijālī, Gabār (also called Gowro) and Bhaṭīse (also called Baṭera). My main language consultant for Gabār was Muhammad Zaman and for Bhaṭīse it was Akhtar Khan. Unfortunately, work on these two language had to be put to the background after some time, as work on Jijālī, especially on its pitch accent system, turned out to be very demanding and time consuming. Nevertheless, all words collected in the two languages have been incorporated into this dictionary. It was not possible to analyze their ac-

cent systems with the speech analyzer, but each word was checked whether or not it had a parallel in Jijālī.

The speakers of Indus Kohistani belong all to the Sunni community of Muslims. During the past decades their culture has been strongly influenced by the lifestyle of the Pashtuns. For instance, pardah and gender segregation have been tightened up in the recent past. With the single exception of an interview of a few hours with an elderly lady from the village of Maharīn (where Gabār is spoken) all data presented in this book are from male speakers. I am nevertheless sure that the language of female and male Kohistanis does not display significant differences. Although I could not work with female speakers, I had opportunity to observe their conversations. I owe this to the very unusual honouring that my Kohistani friend Said Faqir allowed me as a non-family member to stay within his family precincts.

But my field research was also influenced in other ways by the cultural conditions of Kohistan District. Since many clans of the area are engaged in bloodfeuds with other clans, I had to adapt to the specific situation of the Bhadar clan with whom I am associated. Several times it so happened that when I told my Kohistani friends of my intention to visit a certain place, I got the answer that this was not possible because "there are not our own people" and that my security could not be guaranteed. Of course, I always accepted this. In 2000 an old bloodfeud involving the Bhadar clan flared up again. This made it impossible for me to come again to Khangir Khan's village. For a good description of the social systems, the traditional lifestyle of transhumance, and the backgrounds for the bloodfeuds see Zarin and Schmidt 1984.

Aside from studying the Indus Kohistani language I also worked on the oral traditions. This again was a delicate task. Many of the traditional musicians, called Ḍōm, had been expelled from the area in the 1970s by zealous guardians of the religion. In fact, nothing was known about their oral traditions. After one year, during which I occasionally alluded to my interest in oral traditions, I felt sufficiently integrated and started to contact Ḍōms. Later on I was able to record many songs and stories as well as the popular epic of Prince Bahrām. It was, however, not possible to record sung performances accompanied by musical instruments within Indus Kohistan. Performances of this type are officially branded as sin. So I had to take the musicians either up to Gilgit or down to Islamabad. I organized twice public performances in Islamabad, once at Lok Virsa (the Pakistan National Institute of Folk Heritage), and once at the German embassy. This had a stimulating effect on the

subdued love of the Kohistanis for their old traditions. I hope one day a more tolerant atmosphere will allow those traditions to be cultivated again in a way they deserve.

The Ḍōms of Kohistan District do not entertain close connections with the Ḍōms of Hunza and Nager. Still both might belong to the same immigration wave. The Ḍōms of Hunza and Nager have preserved their inherited language (which is, however, acutely threatened), whereas the Ḍōms of Kohistan District speak usually Shina (sometimes also Indus Kohistani).

Less threatened than the oral traditions of the Ḍōms of Kohistan District are the traditions of story telling, especially of fairy tales. The fairy tales, which resemble the Persian *dāstān*s, are very popular in Indus Kohistan. It is not difficult to find good story tellers. A selection from the collection I made will be soon published.

1.6. Language setting

To this day, the Dardic languages are classified by combining historical with geographical features. This is not surprising because, as I will point out below in the next section, the linguistic situation of northern Pakistan is best described with the punctated equilibrium model of Dixon as a case of language splitting while maintaining geographical contiguity. A combined classification is, for instance, provided by Richard Strand (see his webpage). The Kohistani subgroup is divided by him into a western and an eastern branch with Indus Kohistani belonging to the latter, but his further subdivisions within this sub-group are not always correct. He also mentions Damiā-bāṣā or Damēlī, which has both Nuristani and Kohistani features (see also Morgenstierne 1942b). This gives rise to speculation on a former direct geographical contact between Nuristani and Kohistani. Whether the language of Woṭapūr and Kaṭārqalā (see Buddruss 1960) is also related to Kohistani is not clear.

Recent additional information has helped to develop a more exact picture. I owe most part of the information of this and the next paragraph to Joan Baart and his language consultant Muhammad Zaman. In the upper Swat valley Gawri i.e. Kalam Kohistani is spoken (Baart 1997, 1999). There are two slightly different dialects in two valleys with the main settlements Utrōṛ and Ūṣū. But there are also other languages. Baart (1997: 1f.) mentions besides Pashto and Gujari also Khowar. Khowar is locally called *'gōkhā* 'language of Gok, i.e. Chitral' or in Pashto *qaš'qārī*. South of Kalam in the village

of Ariani there is another dialect called Dāçvā. Further south, from around Bahrain till Madyan, Torwali is spoken. In Madyan the long Chail valley, which stretches from east to west, leads into the Swat valley. Chail Kohistani is a dialect of Torwali. Among the dialects of the Swat valley area it is, to my knowledge, the one closest to Indus Kohistani (although mutual comprehension is not possible). Higher up comes the Bishigram settlement area. Ushojo is spoken there. Ushojo is a Kohistani language with Shina elements. More cannot be said at the moment about this language. According to local tradition, the ancestors of the speakers of Ushojo had left Kolai (in District Kohistan), where a form of Shina is spoken, and migrated to Ūṣū where they married speakers of Khowar. From there they came to their present settlements. In Bishigram live also the so-called Gurnewals who speak a dialect of Torwali, as well as speakers of *qaš' qārī* (Khowar) and Pashto. In a nearby side-valley live the speakers of Khilliwal, a form of Indus Kohistani. Fredrik Barth (1956: 15) is not right in claiming that the word is not a proper name but merely means 'villager'. In fact, either this term or Khilōš or Khaniawāl is used in Swat Kohistan to refer to a speaker of Indus Kohistani, usually to a speaker of Indus Kohistani who has settled in Swat Kohistan. The expression Khaniawāl (as well as the language name Kanyawali) is derived from the Khandiā valley. Also Khilōš is probably related with this valley. Indus Kohistani *khíl* means 'fallow land' and there is a village of the same name in the Khandiā valley. Finally, there are also speakers of Iranian Badeshi or Badakhshani who claim to have come from Badakhshan in Afghanistan.

West of the Swat valley is the Panjkoṛa valley with further variants of Kohistani. The dialect in the uppermost villages is very close to Gawri i.e. Kalam Kohistani. Further down, in the Rājkōṭ settlement area, Rājkōṭī is spoken, apparently a dialect of Dāçvā. It still can be understood by speakers of Gawri. Further down is Kālkōṭ. There are two dialects within one and the same village. One is called Kālkōṭī or Goedijā, the other Darākī or Darāgī. There are different opinions of how closely related the latter is with Gawri, but Kālkōṭī (resp. Goedijā) seems to be related with Phalūṛa (resp. Palūla), a dialect of Shina spoken in Chitral. Interestingly, like Šāṭōṭī (the Indus Kohistani language enclave south of Gilgit), Kālkōṭī has apparently given up its Shina-type pitch accent system and adopted the Gawri tone system of its immediate surroundings.

This amazing patchwork of languages and dialects shows that language enclaves are a regular and typical feature on the language map of the southern mountains of northern Pakistan. Apparently, there has always been a high

degree of mobility among the people. Resettlements were and are frequently caused by bloodfeuds. Allegiance to powerful chiefs in the new settlements created perhaps sometimes language enclaves with several different linguistic roots, as in case of Ushojo. Of course, resettlement is not identical with the emergence of a new language enclave. This is quite unlikely to happen when individual families migrate or individuals, who are involved in a bloodfeud, seek asylum in a distant settlement. I met families from Indus Kohistan who had settled as far north as Hunza. They still speak Indus Kohistani at home, but Indus Kohistani will not survive there for long.

1.7. A diachronical sketch

The Dardic languages are the modern successors of Middle Indo-Aryan (MIA) Gāndhārī and other unknown MIA languages more or less closely related with Gāndhārī. They are certainly not the successors of MIA languages like Māgadhī, Śaurasenī, Mahārāṣṭrī, etc. It is thus inexact when Morgenstierne says that "There is not a single common feature distinguishing Dardic, as a whole, from the rest of the Indo-Aryan languages. . . Dardic is simply a convenient cover term to denote a bundle of aberrant Indo-Aryan hill languages" (1961: 139). A similar view is taken by Fussman (1972: 12) who says that the term "n'implique pas, pour l'instant, que les parlers dardes aient une origine commune, différente de celle des autres parlers I-A" (i.e. indo-aryen). The common feature distinguishing the Dardic languages from the other New Indo-Aryan (NIA) languages is the preservation of the three OIA sibilants *s, ś, ṣ*. Despite some earlier reports to the contrary, all Dardic languages have preserved these three sibilant.[3] It is also not clear in which sense Dardic languages are supposed to be "aberrant" in contrast to, as apparently implicated by Morgenstierne, the Nuristan languages. One cannot, of course, exclude the possibility that speakers of IA languages not deriving from Gāndhārī, or a related language that also preserved the three sibilants, migrated into the north Pakistan mountains in the past. But since the non-Gāndhārī IA languages had their number of sibilants reduced from three to one already in pre-Christian times (with the exception of the West Pahāṛī languages many of which have retained two sibilants), such migrations must be either traceable or they are mere speculation. Clear cases for the former are Gujari and Ḍomaakí. Gujari has one sibilant and Ḍomaakí, like many Gypsy languages, has two in its inherited vocabulary. The third sibilant *ṣ* is found only in loans from neigh-

boring Shina and Burushaski. We thus get a family tree with successive bifurcations: Proto-Nuristani branched off at a pre-OIA (probably Proto Indo-Iranian [PII]) stage from the rest of Iranian and Indic, and Proto-Dardic at a post-OIA stage from the rest of Indic (see below 1.9.). It is possible that there was again a later bifurcation which led to modern West Pahāṛī and Gypsy.

Whereas we thus get a fairly clear picture of the most important steps in the unfolding of Indo-Iranian which is framed by a huge geographical space and a great time-depth, the picture within the much smaller frame of the Dardic languages looks very different. It is at the present stage of our knowledge only possible to outline for them an extremely rough family tree model. The historical relationship among the subgroup of the Kohistani languages, for instance, is still an enigma. Besides the fact that still many Dardic languages are only imperfectly studied, there is another reason for this.

The history of the Dardic languages can be best understood, I think, with Dixon's punctated equilibrium model (1997: 67ff.). This model integrates the family tree and the linguistic area (wave) models. He says (p. 73), "During a period of punctuation new languages will develop at a steady rate. As the period of punctuation comes to an end, it can be modelled by a family tree diagram. As a new period of equilibrium sets in, the original genetic relationships of the family tree diagram will become progressively blurred, due to the diffusion of linguistic features throughout the equilibrium period." Dixon further distinguishes two types of language splitting, namely language splitting under geographical separation and language splitting while maintaining geographical contiguity (op. cit. p. 59ff.). Whereas the former is typically a gradual process linked with the increasing geographical distance between the divided groups, the latter is typically characterized by periods of gradual dialect divergence followed by sudden splits which are again followed by a gradual further divergence of languages. The latter type of language splitting, which seems to reflect the scenario of the Dardic languages, is, according to Dixon, invariably motivated by political reasons (op. cit. p. 62): "In the 'geographically contiguous' situation. . ., each group is fully aware of the other, and the sudden escalation of diverging dialects into distinct languages is primarily a political move, to institutionalise political self-identity and demonstrate antipathy towards the other group." I have pointed out above the high mobility of population groups in northern Pakistan, and I will quote some examples below in section 1.9 which will show that speakers of Indus Kohistani must have had knowledge also in the more distant past about the Northern Areas.

1.8. A diatopical sketch

The initial punctuation creating the Proto-Dardic languages was followed by long equilibrium periods. The equilibrium periods were certainly punctuated time and again leading, for instance, to the different Kohistani languages. However, since these processes happened within a frame of geographic contiguity, they resulted in a continuous diffusion of linguistic features. The history of the Dardic languages thus resembles the situation of the indigenous languages of Australia (see Dixon op. cit., p. 89ff.), even though the temporal and geographic dimensions of the former are much smaller than those of the latter.

The diffusion of linguistic features in the Dardic realm is exemplified by the following traits: (a) language boundaries and boundaries of isoglosses are frequently not identical; (b) there are a number of linguistic features distinguishing within Dardic a central (or progressive) from a peripheral (or conservative) area.

(a) The distribution of the number of aspirated stops from east to west is like this: Shina of Gilgit has the three phoneme classes *t, th, d*; Kohistani Shina, Indus Kohistani and Chail Kohistani have four (*t, th, d, dh*); Kalam Kohistani (Gawri) and upper Dir Kohistani have again three (*t, th, d*); Darāgī has (almost) just two (*t, d*).[4] One observes a gradual increase and decrease of the number of aspirated stops from east to west rather than clear boundaries coinciding with language boundaries. It is probably not surprising that Darāgī closely resembles the Nuristani languages in terms of loss of aspiration when we recall the mixed Kohistani-Nuristani character of Dameli which is spoken to the north-west of Darāgī in southern Chitral. This observation stresses the *diatopical* character of the occurrence of aspirated stops against the previous claim that Nuristani shows here an ancient Iranian feature. Tikkanen (1988: 308) has formulated the same thought even more pointedly by saying that the loss of aspiration in Proto-Dardic *and* Nuristani is due to a *substratum influence*. In other words, the loss of aspiration in Nuristani must not be understood as a development shared with Old Iranian because "even Old Iranian maintained the distinction between voiceless non-aspirates and voiceless aspirates by turning the latter into fricatives" (Tikkanen, loc. cit.).

This tilt from east to west from four towards two aspirated phoneme classes becomes reversed further north-west and west. Kalaṣa and Pashai again have four, though the aspiration of the mediae in Pashai seems to be rapidly

vanishing (Morgenstierne 1926: 88f.). Khowar and Woṭapūrī have three, but the number of aspirated stops in comparison with the non-aspirated counterparts is much less than, for instance, in Indus Kohistani or Punjabi. The diatopical rather than diachronical nature of this distribution is further illustrated by Iranian Parachi which has again four stop classes *t, th, d, dh*, though the number of aspirated stops is quite small (Morgenstierne op. cit., p. 18ff.; Nawata 1983: 3).

An almost identical relief emerges with regard to vowel palatalization. OIA *sūrī*- ‘sun’ appears in the following forms from east to west:

Shina *súuri*; Indus Kohistani of Seo *suri*; Gabār (Gowro) *sū̀r*^{i}; Indus Kohistani of Jijāl *suī̀r*^{i}; Kalam Kohistani *sīr*; Dir Kohistani *sī*; Kalaṣa *sūri*; Pashai *sur*; but Woṭapūrī again *sir*.

(b) More examples would only illustrate the same phonological tendency with a kind of gravitational center in the Dir and Kalam Kohistani language area surrounded by peripheral, that is phonologically more conservative, languages like Kalaṣa, Khowar and Shina. This (admittedly inexact) bipartition is again reflected in the use of two different vigesimal systems. For instance, the figure forty-three is either expressed as ‘three plus two times twenty’ or as ‘two times twenty plus three’. The former pattern is found in all Kohistani languages, the latter in the languages surrounding them, e.g. in Shina, Khowar, Kalaṣa, Pashai and Woṭapūrī.

Since the above examples illustrate a parallelism of phonological and morphological features one can assume a fairly early separation of the Proto-Kohistani languages from the other Proto-Dardic languages.

1.9. Dardic and Nuristani

The results of my field research on Indus Kohistani do not at all question the status of the Nuristani languages as a third branch of Indo-Iranian. On the contrary, it is possible to further support Morgenstierne’s arguments. Several times Morgenstierne has pointed out (see e.g. 1973c: 339) that Nuristani, as Iranian, has preserved the difference between PIE palatalized **ḱs* and non-palatalized **ks*. In Old Iranian (OIr.) this difference appears as a contrast between *š* and *xš*. In Nuristani it is realized as *ċ* versus *č*/*c̣*.[5] Dardic, being Indic, has not preserved the distinction and represents both as *c̣*. See the fol-

lowing correspondences in the first table, all deriving from words containing an original palatalized consonant cluster.[6]

Meaning	OIA	Nuristani	Dardic
bear	*ṛ́kṣa-*	*iċ*	*i̯ç^{h}*
belly; flank (of body)	*kukṣí-*	*kuċ*	*kuíç^{hi}*
carve, cut	*tákṣati*	*taċ-*	*taçh-*
right	*dákṣiṇa-*	*daċü(ẽ)*	*daçhő*

On the other hand, Nuristani coincides with Dardic in the representation of PIE **ks*:

Meaning	OIA	Nuristani	Dardic
knife, razor	*kṣurá-*	*çurī*	*ç^{h}ür*
honey	*mākṣiká-, *mākṣa-*	*maç' i*	*mā́ṣ*

It is very unlikely that the Nuristani words containing **ks* should all have been borrowed from Dardic, especially in the light of words like *wučəw-* 'sneezes' (OIA **vikṣuvati*) for which Dardic parallels are missing.

For Nuristani *iċ* 'bear' one has to reconstruct Proto-Nuristani **ṛ́kša-* which led to **ṛčša-* and **ičа-* until the present form.

This shows that Nuristani has preserved the distinction between PIE **ḱs* and **ks* as a contrast between an Iranian type of RUKI and an Indic type of RUKI. What is even more interesting is the fact that this old contrast, which has been preserved in Old Iranian as a contrast between a consonant cluster and a single consonant, is changed in Nuristani into a contrast between two classes of affricates differentiated by place of articulation. This is neither Old Iranian nor Old Indo-Aryan. Languages with three sibilants and three affricates differentiated by place of articulation are characteristic of a linguistic area which is shared by Nuristani, Dardic, Burushaski and some East Iranian languages.

It is natural to assume that Proto-Nuristani **č* (reflecting PIE **ḱ* and **ḱs*) did not change into a dental affricate already at the time of Old Iranian but at a later time. The **č* which developed from the palatalized **ḱ* was the precursor for Old Iranian *s*, Old Indo-Aryan *ś* and Nuristani *ċ*. Since Nuristani also preserves the occlusion of **ḱs* and **ks* as Old Indo-Aryan and Dardic do, whereas Old Iranian loses the occlusion, there is no reason to separate the Nuristani development of **ḱ* from that of **ḱs* and **ks*, and treat it as

the Proto-Aryan precursor of Iranian (Mayrhofer 1984). The depalatalization change from Proto-Nuristani *č to ċ is, in fact, again an areal phenomenon. This change is very common in Indus Kohistani, but also fairly prominent in Dardic Kalaṣa, Woṭapūrī and Iranian Pashto, and found to a lesser degree in various other Dardic languages and Burushaski (in its IA loans). It is difficult to ascertain its age as one has to assume a gradual spread. It was known in Saka, also Gāndhārī seems to have had a dental affricate in a few words.[7]

The development of PIE *ǵ, ǵh, ǵw, ǵwh also does not show a particular closeness of Nuristani to the Iranian branch. Iranian has preserved the opposition between the palatalized and the velarized mediae as a contrast between dental and palatal fricative/affricate (see Avestan *zyam* 'Winter' versus *jyā-* '(bow)string') but has lost the aspiration opposition. Conversely, Old Indo-Aryan has retained the aspiration opposition (in the form *j* versus *h*) and lost the opposition between palatalized and velarized mediae. Since Nuristani has preserved neither the one nor the other opposition (see Nuristani *jim* 'snow' and *jī* '(bow)string'), it is neither especially close to Iranian nor to Indic. Thus the conclusion suggests itself that Proto-Nuristani continued for some time a third kind of opposition, namely **j* versus **jh*. Pashto follows Nuristani (Pashto *žimai* 'winter' and *žaī* 'bow-string') and thus shares with it a regional characteristic.

Since Nuristani has preserved the feature [+palatal] from PIE *ǵ and *ǵh it is also for this reason unlikely that Nuristani ċ should have lost its palatal feature at a very early stage. The change from a palatal to a dental voiced affricate or fricative is again a regional feature (and parallel to the aforementioned depalatalization process of voiceless affricates) and found in Pashto, and various Nuristani and Dardic languages.

The Iranian loanwords in Nuristani are, according to Buddruss, due to a very early Iranian influence (see Degener 1998: 7). This is not at all surprising when we consider that especially Khowar but also various other Dardic languages and Burushaski contain old borrowings from some Iranian source(s). See also Kuiper 1976 on the early contacts between Old East Iranian and Old Indo-Aryan.

We thus can confirm the theses of Morgenstierne and Buddruss that Nuristani forms a third branch of Indo-Iranian. It differs from both in its treatment of PIE *ḱ and of the PIE *ks, ḱs and related clusters, in the way of how it lost aspiration, and in a probably particular way of treatment of PIE *ǵ, ǵh, ǵw, ǵwh. The fact that Nuristani shares some of these developments with Dardic was perhaps partly caused by some substratum influence.

1.10. Borrowings and substrata

The vocabulary of Indus Kohistani derives basically from Old Indo-Aryan. Very roughly estimated, it shares about two third of its inherited vocabulary with Kalam Kohistani and at least one third with Shina. The bulk of borrowings from Arabic and Persian is recent, although the rate of borrowings has perhaps already increased after the Pashtuns conquered the middle Swat valley in the first half of the 16th century (Jettmar 1975: 203) and started missionizing Kohistan. Presently Indus Kohistani contains around 15% Perso-Arabic vocabulary, the bulk of which it shares with Pashto. There might be an additional 10% of genuine Pashto words in Indus Kohistani or shared words with unknown etymology.

There are a few Iranian words which seem to belong to an older layer of borrowing, but it is presently not possible to trace their exact source(s). One example is Indus Kohistani *hʌbʌ̀ỹ* 'co-wife' (also Dameli *a'bēṇi*, Phalura *abhēṇi*, and Shughni *abīn*). It derives from OIr. *ha-paθnī*, but Pashto *bin* 'co-wife' cannot be the immediate source of borrowing.

Indus Kohistani shares also not much less than 10% of its vocabulary with Burushaski (not including the IA words). Whether this fairly high amount, taking into consideration the present big geographical distance, is an indicator that both languages were in the past geographically closer, is not yet quite clear. But this is the opinion of Tikkanen (1988: 305), and below I will list some arguments in its favor. On the other hand, it appears that at least part of this shared vocabulary reflects a common substratum. Two examples: Indus Kohistani *ʌ̀mtʌm* 'greedy; stingy', Burushaski *amtám*, Shina of Chilas *ámtham*; Indus Kohistani *kʌmúkʰ* 'ice', Burushaski *ġamú*, Shina *gamúk*, Khowar *kabukh*, Kashmiri *khambürᵘ̈* (the latter with the special meaning 'a large piece of ice from the heavens').

There are indications that the country of the Burushaski speakers was not unknown to the speakers of Indus Kohistani also in the more remote past. First, the Indus Kohistanis have a unique Indus Kohistani word for the Burushaski language: *gošvārì̄*. The word is contracted from (originally Iranian) *gušpùr* 'a member of the noble families of the Northern Areas' plus a derivation from OIA **vari-* 'speech' which is not (anymore) used independently in Indus Kohistani. Although some Kohistanis were of the opinion that this word originally was used to designate "Chitrali", i.e. Khowar, the first component of the word seems to be otherwise limited to Burushaski (*guśpúr*) and Shina (*guśpúur*).

Second, the speakers of Indus Kohistani have their own word for the well-known Kargah Buddha figure a few kilometres west of Gilgit: *biċkű balʎy*. The second word means 'witch'. In this it corresponds with the Burushaski designations of the figure as *ḍaçhéni, ḍaçheéni* or *yaçheéni* 'witch' (according to Berger contamination of OIA *yakṣá-* and *ḍākínī-*). In Shina the figure is called *Deonee* 'female demon' (Biddulph 1880: 112), obviously a derivation from OIA *ḍākínī* 'female attendant on Kālī'. The first part of the Kohistani expression is perhaps a compound where the first component *biċ-* might be compared with Northwest Prakrit *būtäsa[tvä]* 'Bodhisattva' (see Emmerick and Skjærvø 1997: 118). The second component of the compound might derive from OIA **ḍākkinī-* (Turner 5542) with same meaning as *ḍākínī*. This would explain the vowel of *biċ-* as a typical case of Indus Kohistani vowel leftshift. The final *-ű* is the common Indus Kohistani diminutive suffix. A basic meaning 'Bodhisattva-witch-witch' is not at all unlikely, as various Dardic languages abound in synonym compounds.

Indus Kohistani contains a fair number of words of unknown origin. Some of them it shares with other languages of the area, to others no parallels are known to me. An example for the latter is Šāṭōṭī and Seois *òkhut* or *'ōkhut* 'tomorrow'. In other cases one can trace parallels covering quite big areas. For instance: Indus Kohistani *ḍúṅg* 'deep (as water)', Kangri dialect of Punjabi *ḍuṅgghá* 'deep', Bangani *ḍuṅgɔ*. Somehow related appear also Sawi *ṭungī* and Kalam Kohistani *ḍugur* 'deep'. Turner mentions also *ḍūghā* 'deep' in the Awāṇkārī dialect of Lahndā as a case of metathesis of a derivation from OIA *gūḍhá-* (4223), but it is unclear whether the above examples, which almost all contain a nasal consonant, belong to the same lemma.

Fussman regards the existence of a substrate in the Gāndhārī-speaking area as quite possible (1989b: 446 fn.), and Tikkanen, who discusses the possible existence of ancient substrata in northern Pakistan, notes that "there are also some indications of one or more ancient unidentified substrata in the Hindukush and the Upper Indus region" (1988: 304).

Discussing the question of a possible Dravidian influence, Tikkanen states (p. 317) that "there are no demonstrable ancient Dravidian loan-words in the (north)western Aryan languages." From the Indus Kohistani point of view I basically agree with him, even though there are a few words of possible Dravidian origin without known parallels outside the Dardic language area. See, for instance, Indus Kohistani *guṭùm* 'deep' and *guṭuìm*^{i} 'deep place in water, river', Šāṭōṭī *guṭùmo* 'deep', Burushaski *ġuṭúm*, Shina *guṭúmo*, and perhaps Khowar *kulum* 'deep'. Cf. Tamil *kuṭṭam* 'depth, pond' (DED 1389).

Note also Bhaṭīse *pāṭ*˺ ‘bark of a tree’, cf. Tamil *paṭṭai* ‘bark of tree’ (DED 3205).

In the introduction to this dictionary I cannot comment on the various phonological, morphological and syntactic patterns discussed by Tikkanen under the heading of substrata (this will be done in the grammar), but have to confine myself to the vocabulary. I therefore can only refer to his opinion of an Austroasiatic influence upon Indo-Aryan and Nuristani (see, for instance, 1988: 319). In comparison with the very slight (if it at all exists) Dravidian influence upon Dardic, the situation with a possible Austroasiatic influence even on the vocabulary looks quite different. Nevertheless, one has to keep in mind the possibility of chance similarities which is especially high as long as there doesn’t exist an etymological dictionary of the Austroasiatic languages of South Asia. It also cannot be ruled out that at least some of the parallels quoted in this dictionary are ultimately either Indo-Aryan words or words from unknown and extinct languages, having survived only in peripheral regions. And it remains to be seen whether certain regular allomorphic tendencies in the formation of “expressive” words reflect Austroasiatic morphological patterns. Cf. for instance Indus Kohistani *khaṛʌ̀č* ‘given way (with the knee)’, Burushaski *khaḍáṅ man-* and *qhuḍáp man-*, Shina *khaṛaṅ*, and Santali *khaṛat-khoṛet* ‘limpingly, haltingly’. Or Indus Kohistani *khúč* (*-khuč*) ‘a shout for calling a dog’, Wakhi *kuč-kuč* ‘interj. pour appeler à soi des chiots, des chiens’, and Santali *khuć khuć* ‘said to encourage a dog’.

Tikkanen mentions the word for ‘cat’ Proto-Austroasiatic **pusi/u* (p. 320, fn. 24) which is indeed widespread in the Northwest and quoted by Turner as **puśśī-* (8298). Further examples: Pashai *draṅg* ‘steep’, Santali *daraṅg* ‘steep, a precipice’; Indus Kohistani *čōpúṛ* ‘cockroach’, Santali *capṛa* ‘cockroach’; Indus Kohistani *ḍhòp*^{h} ‘hill’, Santali *ḍhopo* ‘hillock’.

Indus Kohistani (as well as other Dardic languages) has a distinctive preference for synonym compounds (e.g., ‘woman’ basically meaning ‘daughter-girl’). It is not yet clear whether this indicates influence from a substrate, but a number of etymological interpretations suggested in the dictionary have this observation as a background.

1.11. On place names

It is almost certain that there are layers of place names which must be older than the split of Gāndhārī into the Proto-Dardic languages. One finds, for

instance, village and town names ending in *-grām* (< OIA *grā́ma-* 'village') in a strip extending from south of Indus Kohistan in Hazara through the Swat and Dir valley into Chitral (and further into eastern Afghanistan), and even in the north in the Yarkhun valley (*Miragrām*) and in Nager (*Haréi Girám*, a village in Uyum Nager). Names of towns and villages like *Baṭgrām* in Hazara and *Koṭegrām* in Swat do not reflect the modern pronunciations of the words for 'village'. Etymologically related place names ending in *-gām* and *-grə̃* (which are perhaps younger than those ending in *-grām*) are also found within the same area at least as far north as the Shandur Pass (there is a village *Phargām*).

The word *girám* 'scattered settlement' has even been borrowed into Burushaski, but it is striking to see that the word is found only in the eastern Shina dialects with the Astor valley roughly as the western border. From the Astor valley along the Indus and down through Indus Kohistan to Hazara no village names occur with this ending.

Village and town names ending in *-pur* (< OIA *púra-* 'town') are certainly much rarer than the preceding ones ending in 'village'. Perhaps they are confined to the northern and eastern sections of north Pakistan. Among them are *Jalipur* (between Chilas and Bunji), *Gulápur* (near Punyal), *Laspur* (near Shandur Pass) and *Giddarpur* (north-west of Mānserah).

Village and town names ending in *-(i)āl* 'place, settlement' (< ?) are very common in Indus Kohistan (also in the side-valleys, see appendix) and further up the Indus and Gilgit valley. They are also found in the Swat (*Chupriāl*) and the Nilam valley (*Dudhniāl*).

Geographically not structured appear to be the numerous place names with a suffix meaning 'fort', e.g. *-kōṭ* and *-ōṭ* (< OIA *koṭṭa-* 'fort'), and *-gaṛ* (< OIA **gaḍha-* 'fort'). Place names ending with Arabic *-qilā* are of course of more recent origin. Perhaps worth mentioning here is *Rāni-kōṭ* 'fort of the queen' which is the old place name of the Bhaṭīse settlement area opposite Besham, because it has a parallel *Soṅi-kóoṭ* 'fort of the queen', a village name in the Shina language, but located in Chalt in Hunza.

Geographically also unstructured appear to be place names with *-khāṛ* (< OIA **khaḍḍa-*) or *-gā* 'river', *-gali* 'mountain pass' (< OIA **galī-*), and *-d(h)ār* 'mountain range' (< OIA *dhā́rā-*).

Many areas of northern Pakistan have additional suffixes in place names which appear to be regionally limited. This holds true also for the Burushaski language area. I have no explanation why such regionally limited suffixes are not found in Indus Kohistan. Aside from the widespread ending *-(i)āl*

and the aforementioned *Rāni-kōṭ*, the only transparent and recurring place names are names referring to geographical features or to constructions (e.g. 'ridge', 'valley', 'river', 'tree', 'mill', 'bridge', etc.). I have mentioned above the absence of village names with the suffix 'village' in the area between Astor valley and the southern end of Indus Kohistan.[8] It is now interesting to see that one finds roughly in the area covered by the western Shina dialects and by Burushaski several place suffixes which are either limited to this area or especially typical for it.

The suffix *-śal* is especially common in the Burushaski language area (e.g. *Bériśal, Bóośal, Boróśal, Ǵuśóśal*, etc.) but seems to be extant also far down the Indus in the village name Dudīshāl (Indus Kohistani pronunciation *ḍōḍšȁl*) between Sazin and Chilas.

Also the place name and suffix *-bar* 'ravine' seems to be especially common in the Burushaski language area (e.g. *Bar, Tálmuśi Bar, Híspar*, etc.) but is also found further down the Indus in the village name (Indus Kohistani pronunciation) *khʌyṇḍ̊-bʌỳr*^{i} (lit.: 'mountain-ravine') between Dudīshāl and Chilas.

If *-śal* and *-bar* are indeed Burushaski suffixes,[9] then this would be a strong indication for the language's former greater extension to the south-west.

A typical Dardic place suffix is Indus Kohistani *dhȁs*, Shina *dáas*, Khowar *dās* 'wasteland, plain' which also appears in numerous Burushaski place names as the suffix *-das*. It is again strange that, although the Indus Kohistani pronunciation is closest to the original OIA *dhā́sas-*, it is not found in place names inside Indus Kohistan.

Chapter 2
Technical aspects of the dictionary

1. Introductory remarks

The dictionary contains around 8.000 lemmata. The vast majority are lemmata of the Jijālī (J) variety of Indus Kohistani. A minority comprises lemmata from the Šāṭōṭī (Š) variety (plus a few from the Seo (S) variety) of Indus Kohistani and from the adjacent languages Gabār (G) (also called Gowro) and Bhaṭīsē (B) (also called Baṭēṛā). There are cross-references between identical, similar or in other ways related lemmata of these four tongues in case they are alphabetically located in different places. However, they are ordered by a hierarchical principle: a B lemma and a related G and Š lemma (located, however, at a different place) might all contain the statement "same as J xyz." In this case, the etymology of the lemma (if known) is only mentioned under the J lemma. Moreover, the J lemma usually does not contain a reverse cross-reference to one or more of the other three tongues. In case the etymological relationship between the different words or between a lemma and a possible source of origin or borrowing is not certain, usually the expression "Cf. xyz" is employed.

Many lemmata are supplemented with parallels from other dialects of Indus Kohistani, from other Dardic, from Nuristani, Indo-Aryan, Iranian, Dravidian and Munda language. Numerous parallels have also been quoted from Burushaski. This helps the reader to get acquainted with the multi-layered linguistic background of Indus Kohistani. A detailed historical analysis of these data is, however, part of the grammar. Most parallels are from sources quoted in the references. Additional information on parallels in Pashto and Shina came from native speakers and from experts, in the first place from Georg Buddruss. Those cases where language consultants could not provide an exact pronunciation but only say to be sure that a certain word is also used by speakers of Kohistani Shina or of the Pashto variety spoken close to Indus Kohistani, appear in the dictionary formulated as "also Sh." or "also Psht." The lemmata have been, wherever possible, provided with information about their etymology or source of borrowing. In case of IA words decisions, whether a word is inherited or borrowed from another Dardic or other IA language, are of course based on insights into the historical phonological de-

velopments of Indus Kohistani and other Dardic languages. An outline of the developments in Indus Kohistani will be provided in the grammar. However, I have to point out here that in some cases the lines of development of sounds or sound groups are quite clear, whereas in other cases (e.g. in the development of the nasal consonants and of NC clusters) the picture is very inconsistent. As a result, there are always words about which it is very difficult to say whether they are inherited or borrowed.

Parallels to lemmata have, as a rule, only been quoted when an etymology for them has not been suggested in the literature. An exception are all those cases where I regarded the quoted parallels especially interesting due to the following reasons: the words show particular phonological, morphological or semantic correspondences or differences or the words belong to a group with parallels in other languages which have not yet been pointed out.

2. On phonology and tonology

It is the J variety which has been completely phonologized and tonologized. The phonological analysis of the three other languages and dialects (Š, G and B) has been conducted less thoroughly, although a final analysis would generally show only minor differences. However, this does not hold true of the B vowel system. Phonetically it is the most complicated one among the four tongues with a large number of allophonic realizations of the vowel phonemes. Although I have analyzed the basic mechanisms of the allophonic variations, the results could not yet be incorporated into the dictionary. See below 2.7 on the transcription conventions. Among field researchers the difficult nature of vowel systems of many Dardic languages is well-known. In the grammar I will discuss the possible central role of the pharyngeal feature opposition between advanced tongue root (ATR) and retracted tongue root (RTR). This opposition appears not only to bear on certain vowel phonologization processes but perhaps also (together with other pharyngeal and with laryngeal features) on the historical development of Dardic tone and pitch accent systems.

The majority of the lemmata of the Š variety of Indus Kohistani in this dictionary are tonologized with the help of the speech analyzer. The lemmata from B and G have not been analyzed in this way. This means the following: in the majority of cases, accent marking of the G words only indicates the place of the accent. It does not indicate whether the accent is rising or falling.

In case of the B words, many of the have been marked either with the rising or the falling accent. This marking, however, was never checked with the speech analyzer. However, since B and G appear to have similar cases of accent shift under specific conditions (a synopsis of the rules for J is given below in section 2.3.), it is to be expected that the accent systems of B and G are quite similar to those of J. One finds in the dictionary that I have placed in many B words the accent on the consonant preceding a vowel. Though it is naturally the following vowel which is the tone-bearing segment, a major difference on the tonetic level between the dialects and languages of Kohistan District consists in different degrees of peak-delay (see also next paragraph). Although the role of peak-delay in historical tonologization processes in Dardic languages is still an unresolved question, it has become clear that this is one of the features which causes to make the "intonation" of neighboring dialects sound "strange" or even "funny" in the ears of native speakers. Since B appears to have a shorter peak-delay than J and since, in contrast to J, this early onset of the accent contour frequently even has a phonetic effect on the preceding consonant (the exact nature of which has yet to be analyzed), I opted for this way of presenting accent in B.

I had to decide whether to dedicate the available time for tonological analysis in the field to Š, B or G. I chose Š in the hope to learn from this language enclave something about tonogenesis in Indus Kohistani. However, the surprizing result, mentioned already above, was to find that Š had adopted the Shina accent system and given up its old Indus Kohistani system. Meanwhile we know that this is not an isolated case among the language enclaves and minor languages in northern Pakistan. However, I think that on a tonological level the Indus Kohistani and the Shina accent systems are very similar. On the tonetic level the major differences appear to be different degrees of peak-delays and apparently also different tonetic realizations of the rising and the falling accent contours.[10] Consequences of these differences are differing rules for accent shift and the phenomenon of accent spread onto a following word under certain conditions which is found in Shina but not in Indus Kohistani (see Radloff 1999: 73ff.). There are also a number of common words shared by Indus Kohistani and Shina but with different accents. This kind of difference, in case it can be explained at all, appears to be sometimes connected with features like vowel length, aspiration, etc.

The J lemmata have been thoroughly analyzed with the speech analyzer. A speech analyzer is a very useful addition in tonological analysis, though it does not replace a sensitive ear. It helps to solve many problems, but it

does not solve all. The most important feature of a speech analyzer in tonal analysis is its faculty to present pitch graphs by transforming acoustic wave forms into F_0 contours. In order to achieve reliable results it is important to use a constant frame for the words to be analyzed. Thus my language consultants permanently used sentence patterns like "this is X" or "we call this X." Most of the J words were analyzed several times in this way, a large part of them even with several native speakers. Moreover, the tonal analysis of a selection of words has been independently compared by another linguist with his own records made from Indus Kohistani. He ascertained a very high degree of agreement between his and my analysis.

The existence of many minimal pairs proves that Indus Kohistani has two pitch accents, one rising and the other falling. It may have been noted that I use the term pitch accent in case of the languages and dialects of Kohistan District and in case of Shina, but that I speak of tones in connection with Kalam Kohistani. It is useful, on the one hand, to distinguish between tone languages like Kalam Kohistani (which has five tonemes) and pitch accent languages like Indus Kohistani (which has two emically distinctive tones). The difference between the two languages in this regard is that Kalam Kohistani has basically two tonemes from among its inventory of five tonemes associated with a morpheme, whereas Indus Kohistani has basically one pitch accent associated with a morpheme. On the other hand, it has been frequently pointed out that there is no absolute division between tone and accent languages (see Yip 2002: 4), which, however, is quite clear between them and non-tone languages.

Baart estimates (see his internet article "Tonal features in languages of northern Pakistan") that from among the thirty or so languages of northern Pakistan there are thirteen with reported tones (or distinctive accents) and five with suspected tones. This means that probably more than half of these languages have tones. This is a lower percentage than the percentage of tone language worldwide, which is estimated to be between 60 and 70 per cent (Yip 2002: 1), but it is certainly much higher than the average percentage of tone languages within the Indo-Aryan and Iranian languages.

A major problem in the analysis of the Indus Kohistani accent system is to determine whether the Tone-Bearing Unit (TBU) is the syllable or the mora. The answer to this problem will decide whether on a level more abstract than the tonemic level with the two distinctive accents only a monotonic H or a ditonic H and L tone will have to be established. A final treatment of this question will be undertaken in the grammar, but some preliminary decisions

had to be made in order to decide on the graphemic representation. For Shina, whose accent system is close to Indus Kohistani, Radloff and Baart favor a monotonic interpretation which associates the accent with the mora. With that they continue an approach which started with Berger for Burushaski and which was continued by Buddruss for Shina. The basic concept is to say that a word with a long vowel consists of two morae. Depending on whether the accent H is located on the first or on the second mora, the tonetic output will either be a falling or a rising accent contour (Radloff 1999: 63ff.). Thus Shina *tíino* 'sharp' has a falling accent and *toóm* 'own' has a rising accent. In polysyllabic words H is associated with one mora of one syllable, whereas the other syllables are automatically associated with L.

The main reason why I gave up the above writing conventions has to do with the fact that in Indus Kohistani the *tonetic* contrast between rising and falling accent is partially neutralized on syllables with a short vowel under conditions described in the next section. It is, for instance, not neutralized when the syllable ends with an *l*: *kul* 'grain' has a falling accent and *kul* 'people' has a rising one. In graphemic mora representation this had to be written as *kúl* 'grain' and *kuĺ* 'people'. Though there is no reason to say that on an abstract level H cannot be either associated with a vowel or a consonant, the medium of tonetic realization is in both cases the vowel only. Another problem is this: Indus Kohistani has both short and long monophthongs *and* diphthongs. With a moraic interpretation, long diphthongs had to be regarded as trimoric. But since the two accents show the same behavior with long monophthongs as they do with long diphthongs, the long vowel of a monophthong had to be regarded as a complex unit and the long vowel of the long diphthong as a simple unit. This is highly undesirable. And another problem is this: words of the form CV have always the falling accent in Indus Kohistani (see next section). Thus the words *be* 'we' and *na* 'not' both have a falling accent. Both words have under certain conditions allomorphs with long vowels. Then, however, *bē* has still a falling accent, whereas *nā* has a rising accent. There is, at first sight, no way to graphically express this underlying accent difference between *be* and *na* when a monotonic way of notation is used. Although autosegmental phonology does provide solutions to problems of this kind, it should be clear by now that the complicated ways of representation it requires are part of a phonology chapter in the grammar but cannot be used in a dictionary.

3. Some pitch accent rules

In order to understand the many changes and shifts of the two accents found in the dictionary, it is necessary to get acquainted with the following selection of pitch accent-rules.

1. The rising accent is graphemically presented as *v́*, the falling accent as *v̀* (but see my above comments on the usual graphemic location of the accent on the preceding consonant in B).
2. Each morpheme contains one syllable either with a rising or a falling accent.
3. The *tonetic* contrast between rising and falling accent is neutralized:
   (a) on short monophthongs and diphthongs in polysyllabic words
   (b) on short monophthongs and diphthongs in monosyllabic words if they do not end in *n, m, r, l, ṛ*
   (c) on short monophthongs and diphthongs in monosyllabic words in word-final position
4. The tonetic contrast between rising and falling accent is sustained:
   (a) on long monophthongs and diphthongs
   (b) on short monophthongs and diphthongs in monosyllabic words if they end in *n, m, r, l, ṛ*
5. In case of tonetic neutralization neither the rising nor the falling accent appears exclusively. Instead, the appearance of the rising or the falling accent is regulated by the immediate phonological environment, i.e. generally by the following consonant. For instance, a free morpheme with an accent-carrying syllable of the form *-vp* has always the rising accent *-v́p*, whereas a free morpheme with an accent-carrying syllable of the form *-vs* has always the falling accent *-v̀s*.

   Note: although the kind of tonetic neutralization is unambiguous in many cases, there are some cases where the pitch graphs yielded no clear results (e.g. very slightly rising contours in one pitch graph record alternating with very slightly falling contours in the next record) despite much effort. Thus the dictionary occasionally contains "contradicting" forms. They reflect, however, the tonetic facts.

6. The falling accent of a lexeme is always static, but the rising accent of a lexeme in ultima position shifts to the right in case a declinational suffix is

added. The suffix, however, has to be part of the word (it has to be a layer I suffix, cf. Masica 1991: 232ff. on the notion of layer I suffixes). Example: *azmḗx* 'an examination' → *azmēxà*^{h} 'examinations'.

7. A consequence of 6. is (a) that an underlying rising tone may tonetically be a falling tone in certain cases of neutralization. In this case he still behaves like a rising tone and shifts in declension to the right in case a layer I suffix is added. Inversely, (b) an underlying falling tone may phonetically be a rising tone in certain cases of neutralization. In this case it still behaves like a falling tone and does not shift to the right in case a layer I suffix is added.

   An example for (a): [*bƛs*] 'a bus' → [*bʌsà*^{h}] 'busses'. According to the above rules, the underlying process has the form */bás/* → */basá/*. In the singular form the underlying rising accent is tonetically changed into the falling one because on short vowels before *-s* always only the tonetically falling accent appears. Underlyingly, however, it is the rising accent, otherwise no shift could occur. The shifted accent then has again tonetically a falling contour on the suffix, because on short final vowels only the tonetically falling accent is allowed to appear (see above). An example for (b): [*dhʎp*^{h}] 'a blow' → [*dhʎpa*^{h}] 'blows'. Here the underlying process has the form */dhàp/* → */dhàpa/*.

8. All lemmata showing an accent behavior as described under 7. are marked in the dictionary with the words "note accent." This indicates that the behavior of the accent (shift or non-shift) is due to an inverse relationship between the underlying form of the accent (rising or falling) and its surface form (falling or rising).
9. The words "note accent" or "note unusual accent" are also employed in those (not too many) cases where the accent behaves in a way which cannot be explained by any rule.
10. In compounding, reduplication and derivation accents usually change their value. This is independent of the fact whether the accent of the free morpheme is tonetically neutralized or not. Change of accent is also an important means to realize different verbal moods. Examples: *dt̄s* 'a day' → *dt́s* 'during the day'; *ʎk*^{h} 'one' plus *ʌỹc̣à* 'having an eye' → *ƛk*^{h} *ʌỹc̣à* '(a) one-eyed (person)'; *uiávṽ* 'to place' → *sù*^{h} *uiàt*^{h} 'he places (s.th.)' → *sù*^{h} *uiát*^{h} 'he lets (s.o.) place (s.th.)'.
11. In case a layer II suffix, e.g. a postposition (cf. Masica op. cit.), is added

to a lexeme, then the accent of the suffix is completely suppressed. Example: *zaṅgʌ́l* 'a forest' plus *mʌ̀z* 'in' → *zaṅglà-mʌz* 'in the forest'.

12. In the other compounds there is not complete suppression of one of the accents, but one observes a hierarchy: one accent dominates (it is called nucleus), whereas the other accent(s) has (have) a lower energy level. This is independent from whether the involved accents are rising or falling. Wherever possible, this hierarchy has been marked in the dictionary. An abbreviation "(n1)" means that the nucleus is on the first word, "(n2)" means that the nucleus is on the second word, etc. In very many cases the values and the places of the accents change in the process of compounding: *čhìk pʌtī́* (n2) 'a card game with six players' ← *čhikà*^{h} and *pʌ̀tī*. This is to be read: the compound *čhìk pʌtī́* has the main accent on the second word. It is built of *čhikà*^{h} and *pʌ̀tī*, in both words this leads to a shift of the place of the accent, and in case of the second word also the value of the accent changes from falling to rising.
13. Oscillation: in case of short and long diphthongs the onset of the accent is usually on the first component of the diphthong. The onset of the diphthong shifts, however, onto the second component in the following cases:
    (a) when the coda consists of a consonant plus an ultrashort vowel
    (b) when the second component of the diphthong is an accent-bearing grammeme: *guìl̥*^{i} 'a flat bread' but *ùī* 'udder'; *sù*^{h} *ìã́ bé́t*^{h} 'he comes continuously' but *sù*^{h} *uiã̄̀ bé́t*^{h} 'he wraps continuously (a shawl around himself)'.

4. Further technical details

1. Whenever a lemma is unmarked with regard to language affiliation it automatically means that it is a lemma from Jijālī.
2. The vowels *ʌ* and *a* are allophones in all languages. In Š there is an additional tendency to pronounce this phoneme as *ɜ* and in B sometimes as *ə*. There exists at least a certain tendency to regulate the appearance of these allophones within one and the same word according to certain euphonic rules. For instance, if a word contains two *a* phonemes, then all four languages and dialects (J, G, B, Š) show a certain tendency to pronounce one phoneme slightly closer than the other. The same tendency also exists to a lesser degree with regard to the other vowels. The languages J and G, for instance, differ however with regard to which of the allophones they

usually prefer. J prefers the closer allophone, whereas G prefers the more open allophone.

3. Whenever a nominal form can add a direct case plural suffix, this has been listed in the dictionary: the plural ending is separated from the singular lexeme by a comma. Thus: *bə̀ɣ, -a^{h}* 'garden' is to be read as *bə̀ɣ* 'a garden' and *bə̀ɣa^{h}* 'gardens'; *bāɣvə́n, -à^{h}* 'gardener' is to be read as *bāɣvə́n* 'a gardener' and *bāɣvānà^{h}* '(several) gardeners'.
   (a) Many words have more than one way of building a direct plural. This has usually been noted in the dictionary. It appears, however, that there exists a very high degree of fluctuation among the native speakers with regard to the use of different plural forms. Example (i): *aṅgʌ́ṛ, -à^{h}* or *-ī́* 'courtyard' is to be read as *aṅgʌ́ṛ* 'a courtyard' and *aṅgʌṛà^{h}* '(several) courtyards' and *aṅgʌṛī́* '(several) courtyards'. Example (ii): *ghuír$^{i}_{̣}$ gə́l, . . . galà^{h}* or *. . . galṵ́, -ṓ* (n2) n.f. 'a cuckoo'. This is to be read: only the second component of the compound inflects (therefore the dots). It forms the plural in two ways: (a) shortening of the long vowel and adding of *-à^{h}* in the plural; (b) instead of singular *gə́l* also singular *galṵ́* is possible which changes into *galṓ* in the plural. There is the additional information that the main accent (nucleus) is located on the second component.
   (b) Sometimes both components of a compound change from singular to plural: *ṣhirʌ̀ỹ kōrṭṵ́* 'a milk bucket' but *ṣherʌ̀ṽ kōrṭṓ* '(several) milk buckets'.
   (c) Sometimes a second plural ending is redundantly added: *ʌỹ́ṣ^{hi}* 'an eye' → *ə̃́ṣì* or *ə̃́ṣìa^{h}* 'eyes'. In this example the noun stem takes on its non-palatalized (original) form in the plural.
   (d) It appears that at least some of those relatively few words ending in the singular in *-à^{h}* are actually old plural forms. If they add the plural ending *-ī́* the *-à^{h}* is deleted: *kuċà^{h}, -ī́* 'narrow village lane' is to be read *kuċà^{h}* 'a narrow village lane' and *kuċī́* '(several) narrow village lanes'.
   (e) In the other cases no deletion takes place: *kṹ̃ċì, -a^{h}* 'plait curled into a bun' means that the plural is *kṹ̃ċìa^{h}* 'plaits curled into buns'.
   (f) Other forms of constructing the plural are listed in the dictionary. They are also separated by a comma from the direct singular. Thus: *dȉs, dìs* 'day' is to be read as *dȉs* 'a day' and *dìs* '(several) days'.
   (g) Note that when a word ends with an ultrashort vowel in the nominative singular, then it loses that vowel in the direct plural and in the oblique

forms. Thus: *bašʌỳlʲ, -aʰ* ‘flute’ is to be read as *bašʌỳlʲ* ‘a flute’ and *bašʌlyaʰ* ‘(several) flutes’.

(h) Other forms of syncope are either listed in the dictionary or they are explained in the grammar.

(i) A partly devoiced (or lax) consonant in word-final position in singular is pronounced fully voiced in plural: *ə̀ğ* ‘a sheep’ → *ə̀gaʰ* ‘(many) sheep’.

(j) The oblique singular and plural endings are usually regular and not listed in the dictionary. Only irregular forms have been given. They are separated from the direct singular or plural ending by a semicolon. Thus: *patʌṅgɪ̃; -ʌṽ* ‘moth’ is to be read as *patʌṅgɪ̃* ‘a moth; (several) moths’ and *patʌṅgɪ̃ʌṽ* ‘of a moth; of moths’. This is irregular, because a right shift of the accent was to be expected.

4. Many adjectives have different masculine and feminine forms. They are either listed at their respective places in the dictionary (with cross-reference) or they are grouped together in the following way: *bhiyʌ́ṭʰ* adj.m. – adj.f. *bhiyʌýṭʰⁱ* ‘fearful, anxious, timid; terrible’. This is to be read as *bhiyʌ́ṭʰ* ‘(a) fearful (man)’ and *bhiyʌýṭʰⁱ* ‘(a) fearful (woman)’. Another example: *guṭùm* (G, J) adj.m. – adj.f. *guṭuìmⁱ* (J), *guṭúmi* (G); Š *guṭùmo* ‘deep’. Sh. and Bur. *guṭúmo*, Ḍ. *ɣoṭuʻm* ‘deep’. Perh. Dravidian. Cf. Tamil *kuṭṭam* ‘depth, pond’ (DED 1389) and semantics of next entry. This is to be read: Gabār (G) and Jijālī (J) have practically identical adj.m. forms, but their adj.f. forms differ from each other. Šāṭōṭi (Š) has a slightly different adj.m. form and does not seem to have a separate adj.f. form. This information is supplemented with parallels from Shina, Burushaski and Ḍomaakí. Then an etymological background is suggested, and finally a cross-reference is provided.

5. The verbal paradigms

The conjugation of Indus Kohistani verbs is characterized by many irregular and unpredictable forms. Moreover, since it is not expected that the general reader has a complete knowledge of the grammar of the Indus Kohistani verb, the verbs in the dictionary are listed with many finite and participial forms. However, in order to avoid an unnecessary inflation it was decided to provide only the verbs of the first third of the dictionary with comprehensive lists of the finite and participial forms. Paradigms of the verbs of the remaining part

of the dictionary mainly contain the irregular forms. In most cases it is not predictable whether a verb (be it intransitive or transitive) is associated with the rising or the falling accent. It is thus necessary to quote a present tense form. Although there is a strong tendency that when the present tense (be it intransitive or transitive) has the rising accent, the aorist has the suffix *-á́*, and when it has the falling accent, the aorist has the suffix *-é́*. Nevertheless there are a number of exceptions. Therefore the aorist forms are regularly quoted. Similar considerations with regard to continuous aktionsart and other verbal categories caused me to quote them in all paradigms.

Almost all the finite and participial forms are from Jijālī. A verbal paradigm is easily located, because the infinitive appears in boldface. The form of the infinitive is followed by a figure indicating the number of moods the verb can realize. This is followed by the "default mood", i.e. the mood the native speakers of Indus Kohistani regard as the basic one. Then the various finite and participial forms are listed. This is frequently supplemented by additional examples and by parallels from other languages. Finally, if possible, the origin of the verb is given. Example of a paradigm from the dictionary:

biṣhấ̃ (4m) v.t. 'to weave (cloth); to pull (newly sheared or old) wool flocks apart'. The paradigm contains some hab. forms which are no longer semantically distinguished from the tr. forms.
Pres. hab. m. *sù^{h} biṣhṓ̃t^{h}* 'he weaves (s.th.)'.
Pres. hab. f. *sù^{h} biṣhvʌ́ýt^{hi}* 'she weaves (s.th.)'.
Pres. tr. *sù^{h} biṣhə̀̃t^{h}* 'he weaves (s.th.)'.
Pres. caus. *sù^{h} biṣhiá̃t^{h}* 'he causes (s.o) to weave (s.th.)'.
Pres. pass. *sù^{h} biẓzə̀̃t^{h}* 'it gets woven'.
Fut. pass. *sù^{h} biẓzàṣʌt^{h}* or *biṣhúizṣʌt^{h}* 'it will be woven'.
Aor. hab. m. *sù^{h} biṣhṓ̃l* or *biṣhấ̃* 'he wove (s.th.)'.
Aor. hab. f. *sù^{h} biṣhoḗ̃l^{i}* 'she wove (s.th.)'.
Aor. tr. *sấ̃ biṣhḗ̃* or *biṣhìl^{i}* 'he wove (s.th.)'.
Bur. *buyéeẹo* 'Weber', Sh. *buyḗ̃ẹo* (do.).
The hab. forms suggest verbal derivation from an original nominal form 'weaving, weaver', thus < **vayitraka-* (11307). The verb can also be used nominally with *karấ̃*.

This is to be read thus: the verb can realize four moods "(4m)" (habilitative [with different forms for masculine and feminine, and own forms in various

tenses], transitive, causative, passive), and the "default mood" is transitive. After the listing of the finite forms, some parallels from other languages are quoted and a probable origin of the verb is suggested.

Note (a): Habilitative verbs or verbs containing habilitative forms are contracted forms of old conjunct verbs consisting of an adjective and the auxiliary *ho-* 'be' (Indus Kohistani auxiliaries do not have an infinitive). They basically express a typical habit of the subject, although this meaning tends to fade away now.
Note (b): Indus Kohistani does not have a "progressive" form corresponding to Urdu *rahā* (Eng. *-ing*). One of its relatively few compound verb constructions uses *biyáṽ* 'to go' to realize a continuous aktionsart. As the forms of the main verb of this construction are frequently irregular, they are usually quoted in the dictionary.
Note (c): There are a few Indus Kohistani verbs with an anticausative mood. This is discussed in detail in the grammar. Here it suffices to say that on a formal level they are part of a fourfold accent contrast: when a transitive verb has also an intransitive form, then the transitive verb *must* have the rising accent and the intransitive form the falling accent. But when a transitive verb has also an anticausative form, then it is the other way round: the transitive verb *must* have the falling accent and the anticausative form the rising accent. Semantically the difference between intransitive and anticausative can be illustrated with the verb 'to cross a river'. In an intransitive sense it means that a person crosses a river, for instance, by walking over a bridge. In an anticausative sense it means that the person crosses the river, for instance, sitting in a ferry or on the shoulder of someone who crosses the river. As there are only very few anticausative verbs in Indus Kohistani, my impression is that this category is on the verge of extinction.
Note (d): Many verbs distinguish between two types of past participles. A very similar difference is known also from Urdu and can be illustrated with the Urdu verb *karnā* 'to do'. What is quoted in the paradigms with the abbreviation part. perf.$_1$ corresponds to Urdu *kiyā* 'done', and what is quoted with the abbreviation part. perf.$_2$ corresponds to Urdu *kiyā huā* 'already done'. In a few cases, however, my language consultants insisted that the morphological difference does not denote a difference corresponding to Urdu *kiyā* and *kiyā huā*, but denotes a difference between masculine and feminine subject. This has always been noted in the dictionary, even though I have yet no explanation for this.

Reading the forms of the verbal paradigms, two more points ought to be noted: (a) a strong tendency of the habilitative (hab.) forms to be used like indicative forms (in a number of cases I suspect that the old hab. paradigm is on the way of being "regularized" with a normal itr. or tr. paradigm), (b) a partial collapse of finite forms in non-present tenses. For instance, it is fairly common that a present-tense distinction between a transitive and a causative form collapses in the future tense: without distinction both forms are used to express future transitive.

Unless the adhortative (adh. meaning: "please do s.th.") is not further specified regarding mood, its "default mood" is transitive.

The data have been presented in a concise, but structured manner. For instance, when two or three nouns from varieties of Indus Kohistani have (partially) the same gender, then the gender is given only once after the last noun. For instance: *ʌ̀ṛ-ċə̃́n* (J) (n2) n.f.; *aṛ-ċān* (G); *àṛ-ċə̃̀n* (B) n.m. 'half-moon' is to be read: the compound-word for 'half-moon' is n.f. in J, but n.m. in G and B. The example contains the additional information that the accent nucleus is on the second word of the compound in J.

Note: In many cases the gender of the G words could not be determined because the language consultant, an elderly man, could not deal with this type of question.

6. Remarks on different alphabetical orders

It was decided to adapt the alphabetic ordering generally used in Indo-Aryan languages which is based on the Sanskrit alphabet. This is not in accordance with the scholarly tradition started by Morgenstierne, who used an extended Roman alphabet for Iranian, Nuristan, and Dardic languages. My decision is mainly based on two arguments. First, the Roman alphabet does not follow systematic principles based on the natural sound classes, which the Sanskrit alphabet does. Thus I do not see why I should apply a non-scientific principle in a linguistic dictionary of Indus Kohistani when even the ordinary dictionaries of neighboring Punjabi use a scientifically structured alphabet. Second, it seems that every scholar who followed Morgenstierne's decision, felt free to extend the basic Roman alphabetic ordering according to his or her personal predilections. Morgenstierne's word lists start with the vowels,

not ordered according to the succession of the vowels in the Roman alphabet but according to the subsequent consonants. This was not continued by later scholars. They, however, differ from each other in the alphabetical order of the affricates and fricatives. Some have *č, ċ, ẓ*, others have *ċ, č, ẓ*, and again others have *č, ẓ, ts* with the last letter ordered after *t*. Some have *ǰ* in one place and *z* together with *ż* in another, whereas others have *ǰ* and *j̇* in one place, and *z* in another, etc. All this does not help examining the parallels and differences between the phonemic systems of the languages of north-western South Asia. I therefore want to plead for the future use of the Sanskrit alphabet at least for the Dardic and Nuristan languages.

7. Remarks on the transcription

There is actually a phonetic reason for the apparent confusion regarding the alphabetic order of affricates and fricatives, a reason common to most of the languages of northern Pakistan. I devoted some time of my field research to a comparison of the pronunciation of these sounds in the Kohistani languages, in Shina, Burushaski, Khowar and the Pashto of the Swat valley. The results, which I discuss in detail in the grammar, are that there is a strong tendency in word-initial position to pronounce *z* and *ż* as voiced fricatives and to pronounce *ǰ* as an affricate. The phoneme *ǰ* is, however, usually pronounced as a fricative word-finally. There thus exists a common asymmetry between voiced and voiceless affricates/fricatives and within the class of voiced affricates/fricatives in many languages of northern Pakistan.[11] The asymmetry is due to an lenition process which is more advanced in case of the voiced phonemes than in case of the voiceless phonemes. In this dictionary I have ignored this largely phonetic phenomenon in the sense that I only use the graphemes *z, ǰ, ż* without graphically expressing that only *ǰ* is always pronounced as an affricate in word-initial position in Indus Kohistani. The reader is requested to keep this in mind.

In the graphic representation of the lemmata I have tried to follow a middle way between the level of phonetic transcription and the abstract systematic levels. With "middle way" I mean a representation which does not lose itself in a multitude of phonetic details but still stays close to the actual pronunciations. This also means that the representations are redundant to a certain degree. Thus I have regularly written word-final aspiration with an exponent *h*, as for instance in *sùʰ* 'he', even though its appearance or non-appearance is

described by a few simple rules. It must be noted, however, that there is fluctuation in the automatic word-final aspiration of voiceless affricates. Similar redundant representations are the word-final ultrashort vowels (also written with exponent). They too are usually predictable, however, not always. And basically the same holds true for word-final devoicing of voiced segments, nasalization of vowels between nasal consonants, and other predictable phenomena.

8. Alphabetic order

The graphemes employed in the dictionary are those commonly used in South Asian linguistics and largely correspond with the conventions found in the works of R. L. Turner.

(a) The vowels

a (with its allophones ʌ, ɜ and ə) *i, u, e, ɛ, o*. Note that Gabār has also the vowel ɔ which is alphabetically ordered after *o*.

All vowels also have lengthened counterparts marked by macron, e.g. *ī*. Nasalized vowels are marked with a tilde, e.g. *ī, ĩ* . Nasalized vowels follow the oral vowels in the alphabet.

As there are short and long monophthongs there are short and long diphthongs in Indus Kohistani, e.g.: *au, aū, āu*. Note, however, that the two short diphthongs *au* and *ai* are, in accordance with their pronunciation, graphically represented as *ʌv* and *ʌy*.

In most cases nasalized diphthongs have phonetically only the second part nasalized. This is also graphically shown: *ghuĩĺ*^{i} *ċuìṭ*^{hi}.

Most languages and dialects of Indus Kohistan have ultrashort unvoiced vowels. Their occurrence is mostly predictable (see grammar), although there are some very few cases where they realize phonemic contrasts. Graphically they are represented as exponents: *ghuṣíl*^{i}.

Umlaut vowels are common in B and occasionally heard in G. They are represented by superscribed dots: *ö*.

In B also centralized allophones of the vowel phonemes occur. They are represented by a sub- or superscribed dot: *u̯, ï*.

(b) The consonants

k	unvoiced velar stop
kh	unvoiced aspirated velar stop
x	unvoiced velar fricative
q	unvoiced uvular stop
g	voiced velar stop
gh	voiced aspirated velar stop
ɣ	voiced velar fricative
ċ	unvoiced dental affricate
ċh	unvoiced aspirated dental affricate
č	unvoiced palatal affricate
čh	unvoiced aspirated palatal affricate
ç	unvoiced retroflex affricate
çh	unvoiced aspirated retroflex affricate
z	voiced dental fricative
zh	voiced aspirated dental fricative
ž	voiced palatal affricate (word-initially) or fricative (word-finally)
žh	voiced aspirated palatal affricate (word-initially) or fricative (word-finally)
ẓ	voiced retroflex fricative
ẓh	voiced aspirated retroflex fricative
ṭ	unvoiced retroflex stop
ṭh	unvoiced aspirated retroflex stop
ḍ	voiced retroflex stop
ḍh	voiced aspirated retroflex stop
ṛ	voiced flap (word-medially and -finally)
t	unvoiced dental stop
th	unvoiced aspirated dental stop
d	voiced dental stop
dh	voiced aspirated dental stop
n	dental nasal consonant
p	unvoiced bilabial stop
ph	unvoiced aspirated bilabial stop
f	unvoiced bilabial fricative
b	voiced bilabial stop
bh	voiced aspirated bilabial stop
m	bilabial nasal consonant
mh	bilabial aspirated nasal consonant
y	voiced palatal approximant
r	voiced dental trill
rh	voiced dental aspirated trill
ṛ	voiced flap (word-initially)
l	voiced dental lateral approximant
lh	voiced aspirated dental lateral approximant
v	voiced bilabial approximant
s	dental sibilant
š	palatal sibilant
ṣ	retroflex sibilant
h	unvoiced velar fricative

Note: Indus Kohistani has a phoneme *ṇ* which is, however, mostly realized as *ṛ* plus nasalization of the preceding vowel. The dictionary entries represent the actual pronunciation. Thus one finds *khãṛ* and not *khāṇ*.

It has already been mentioned that (in isolated speech) J words ending with unvoiced stops or certain unvoiced affricates are mostly pronounced with a slight aspiration at the end. Graphically this is shown thus: *ʎk*^{h}. On the other hand, the voiced consonants at the end of J words are pronounced in isolation with a certain tendency for devoicing. The detailed rules are described in the grammar. In the dictionary it is indicated by a sub- or superscribed circle: *guìl̥*$^{l̥}$, *ə̀g̊*. In case of word-final voiced stops, e.g. in *adʌ̀d̥*, the subscribed circle does not only mean a tendency for devoicing but also slight delay in the release of the plosion. Phonetically closely related with this phenomenon is the pronunciation of voiced and unvoiced stops in B, where there is (in isolated pronunciation) a very distinct delay in the release of the plosion. This is graphically shown in the following way: *ànḍˀ*. The same grapheme is used to indicate the creaky voice in B words heard once in a while: *phā́ˀl*$_1$. Normally nasalization of a vowel is indicated by a tilde above the vowel. In some B words the tilde had to be placed to the right of the vowel due to technical reasons, e.g. *cə̃ə̃*$^{\sim}$ 'to give'.

9. Transliteration of Pashto words

A large part of the Pashto alphabet coincides with the above alphabet and needs no further explanation. But note that the Pashto consonant *Dze* is transliterated as *dz* and the consonant *Xe, Xin* as *ş̌*.

10. Abbreviations and symbols

10.1. General

Several abbreviations of grammatical terms are illustrated with examples from Urdu. This does not, however, mean that there is an exact equivalence of the corresponding Indus Kohistani and Urdu categories. In fact, frequently it is not so.

acaus.	anticausative
acc.	according
adh.	adhortative mood
adj.	adjective
adj.f.	adjective with feminine inflection
adj.m.	adjective with masculine inflection
adv.	adverb
aor.	aorist tense. Ur. *usne kiyā, vah gayā*
Atlas	see Fussman 1972
ATR	advanced tongue root
attr.	attributive
aux.	auxiliary
C	a consonant
C^i or C^u	a word ending in a consonant followed by an ultrashort voiceless vowel
C^h	a word-final consonant followed by an automatic slight aspiration
C̥ or C̊	lax articulation with devoicing tendency and, in case of unvoiced stops, with delayed release
caus.	causative
comp.	compound
cond.	conditional
conj.	conjunction
cont.	continuous aktionsart
contr.	contrafactive mood
conv.	converb (absolutive)
corr.	corresponding
dat.	dative case
DED	*A Dravidian etymological dictionary* by Thomas Burrow and Murray B. Emeneau
def.pron.	definite pronoun
dimin.	diminutive
do.	ditto (usually in the sense of "same meaning")
erg.	ergative
ex.	example
excl.	exclamation
ext.	extension (usually by a [redundant] suffix)
f.	feminine
fem.	feminine
fut.	future tense
fut. II	futurum exactum
gen.	genitive
ger.	gerundive
H	a high or rising tone or accent in Kalam.
hab.	habilitative aktionsart
HL	a high to low tone in Kalam.
H(L)	a delayed high to low tone in Kalam.
Hy	a high tone with vowel lengthening in Kalam.
i.	intransitive
imp.	imperative
impf.	imperfect tense
int.	interrogative
int.part.	interrogative particle (at end of sentence)
interj.	interjection
intens.	intensive
itr.	intransitive

L	a low or falling tone or accent in Kalam.
LH	a low to high tone in Kalam.
lit.	literally
lw.	loanword
Ly	a low tone with vowel lengthening in Kalam.
m.	masculine
N	a nasal consonant
n.	a noun that can be constructed either with male or female concord
(n1), (n2)...	a syntactic group with the main accent on the first word, on the second, etc.
n.f.	feminine noun
nom.	nominative
n.m.	masculine noun
obl.	oblique case
onom.	onomatopoetic
part.	particle
part. perf.	perfect participle (Ur. *kiyā* or *kiyā huā*)
part. perf.$_1$	perfect participle$_1$ (Ur. *kiyā*)
part. perf.$_2$	perfect participle$_2$ (Ur. *kiyā huā*)
part. pres.	present participle (Ur. *karte hue*)
pass.	passive
perf.	perfect tense
perh.	perhaps
pers.pron.	personal pronoun
pl.	plural
pl.tant.	plurale tantum
plup.	pluperfect tense
postp.	postposition
PP, pp.	past participle
prec.	preceding (entry)
pref.	prefix
prep.	preposition
pres.	present tense
pret.	preterite tense
prob.	probably
pron.adv. refl.	reflexive pronominal adverb
pron.int.	interrogative pronoun
pron.poss.	possessive pronoun
pron.rel.	relative pronoun
refl.	reflexive
refl.pron.	reflexive pronoun
(rf)	"rising-falling": a word with a short vowel (usually between two voiceless consonants) with no clear accent contour on the pitch graph
Rs.	Rupees
RTR	retracted tongue root
subj.	subjunctive mood
TBU	Tone-Bearing Unit
tr.	transitive
V	a short vowel (monophthong)
V^h	a word-final vowel followed by an automatic slight aspiration
V́	a vowel with a rising accent
V̀	a vowel with a falling accent
v.acaus.	anticausative mood

v.aux.	an auxiliary verb		(compounded verbal
v.hab.	a habilitative verb		expressions are not
v.i.	an intransitive verb		marked in this way)
v.imp.	a verb in imperative mood	VV	a long vowel (monophthong)
v.t.	a transitive verb		

<	historically deriving from OIA (it is usually followed by a number referring to the lemma in Turner's etymological dictionary)
←	synchronically deriving from; designates also loanwords with a certain degree of difference in meaning and/or form between source language and Indus Kohistani
*	appears either before a *diachronically* reconstructed form (representing an historically older stage) or before a *synchronically* reconstructed form (i.e. an underlying form which is reconstructed on the basis of synchronic morphonological alternations)
†*	denotes a new reconstructed head-word in Turner's *Addenda and corrigenda*
**	denotes in the Old Indo-Aryan–Indus Kohistani index a reconstruction by this author
$word_1$, $word_2$	enumeration of homonyms (however, only homonyms of the same language are marked this way, and different accents are not considered)
(1m),(2m)	a verb with the faculty to realize one mood, two moods, etc.

10.2. Languages and dialects

In order to graphically express the distinction between the primary data from languages and dialects of District Kohistan from data from other languages, only the former are abbreviated with single capital letters without a period. In order to avoid an inflation of the size of the dictionary the sources of quoted parallels from other languages have only been given in exceptional cases.

A.	Assamese		of the inscriptions of
Ar.	Arabic		Aśoka
Aś.	Aśokan, i.e. the language	Ash.	Ashkun (Nuristani)

Av.	Avestan
B	Bhaṭīse (also called Baṭeṛa) (Dardic)
Bad.	Badaxshānī (Iranian)
Bal.	Balūčī (Iranian)
bhad.	Bhadrawāhī dialect of West Pahāṛī
Bhoj.	Bhojpurī
BHS	Buddhist Hybrid Sanskrit
Bi.	Bihārī
bng.	Baṅgāṇi dialect of West Pahāṛī
Bshk.	Bashkarīk (Dardic)
Bur.	Burushaski
Dm.	Dameli (Nuristani-Dardic)
Ḍ.	Ḍumaakí
dub.	the Dubēr variety of Indus Kohistani
Eng.	English
eur.	European Gypsy
G	Gabār (also called Gowro) (Dardic)
G.	Gujarātī
Gandh.	Gāndhārī
Gaw.	Gawar-Bati (Dardic)
Gy.	Gypsy, Romani
H.	Hindi
IA	Indo-Aryan
II	Indo-Iranian
Ir.	Iranian
Ishk.	Ishkāshmī (Iranian)
J	the Jijālī variety of Indus Kohistani
jij.	the Jijālī variety of Indus Kohistani
Kab.	Kabuli Persian
Kal.	Kalaṣa (Dardic)
Kal. rumb.	Rumbūr dialect of Kalaṣa
Kalam.	Kalam Kohistani (Dardic)
Kam.	Kāmviri (Nuristani)
Kan.	Kanarese
kgr.	Kāṅgrā sub-dialect of the Ḍogrī dialect of Punjabi
Khaś.	Khaśālī dialect of West Pahāṛī
Kho.	Khowar (Dardic)
Ko.	Koṅkaṇī
Kt.	Kati (Nuristani)
ky.	Kanyawālī dialect of Indus Kohistani
L.	Lahndā
lov.	the Lovari dialect of Gypsy
l.rudh.	Low Rudhārī sub-dialect of the Khaśālī dialect of West Pahāṛī
Mai.	Maiy$\tilde{\text{ā}}$ (Dardic)
MIA	Middle Indo-Aryan
Mj.	Munjī (Iranian)
mult.	Multānī dialect of Western Punjabi
NIA	New Indo-Aryan
Niš.	Nishey alā, a Nuristani language of Waigal
ng.	the Nagar dialect of Burushaski
OIA	Old Indo-Aryan
OIr.	Old Iranian
OP.	Old Punjabi
Or.	Oṛiyā
Orm.	Ōrmuṛī (Iranian)
Oss.	Ossetic (Iranian)
P.	Punjabi
pal.	Palestinian dialect of Gypsy of the Nawar

Par.	Parachi (Iranian)
Paš.	Pashai (Dardic)
Paš.ar.	Areti dialect of Pashai
Paš.kuṛ.	Kuṛaṅgali dialect of Pashai
Paš.nir.	Nirlāmī dialect of Pashai
paṭṭ.	the Paṭṭan variety of Indus Kohistani
PD	Proto Dardic
Pers.	Persian
Phal.	Phalūṛa (Dardic)
PIK	Proto Indus Kohistani
Pk.	Prakrit
P.kgr.	the Kāṅgrā sub-dialect of Punjabi
PN	Proto Nuristani
Port.	Portuguese
P.pot.	the Poṭohārī (Poṭhwārī) dialect of Punjabi
Pr.	Prasun (Nuristani)
Pr.Ar.	Proto Aryan
Psht.	Pashto
Rom.	Romani
Rp.	Raṅg pɔ of Garhwal (mixed Tibetan-Garhwali)
rudh.	Rudhārī sub-dialect of the Khaśālī dialect of West Pahāṛī
S	the Seo variety of Indus Kohistani
S.	Sindhi
Sang.	Sanglechi (Iranian)
Sant.	Santālī (Munda)
Š	Šāṭōṭī, a dialect of the central variety of Indus Kohistani formerly spoken in twenty to thirty households in village Šāṭōṭ between the beginning of the Rondu gorge and Rondu town (the households are since 1988 dispersed mainly in the two villages of Paṛi Bangla and Jagloṭ)
Sak.	Saka (Iranian)
Sar.	Sarīkolī (Iranian)
śeu.	Śeuṭī sub-dialect of the Khaśālī dialect of West Pahāṛī
Sh. or Sh.gil.	the Shina dialect of Gilgit (Dardic)
Sh.ast.	the Shina dialect of Astor
Sh.chil.	the Shina dialect of Chilās
Sh.koh. gur.	the Kohistānī and Gurēsī dialects of Shina
Sh.pal.	the Shina dialect of Pālas
Sh.saz.	the Shina dialect of Sazin
Shgh.	Shughnī (Iranian)
Taj.	Tajiki (Iranian)
Tam.	Tamil
Tel.	Telugu
Tir.	Tirāhī (Dardic)
Tor.	Tōrwālī (Dardic)
Turk.	Turkish
Ur.	Urdu
waz.	Waziri dialect of Pashto
weg.	Wegali dialect of Pashai
Wg.	Waigalī (Nuristani)
WKc.	Western Kocī (West Pahāṛī)
Wkh.	Wakhi (Iranian)
Woṭ.	Woṭapūrī (Dardic)
WPah.	West Pahāṛī
WPah. bhad.	the Bhadrawāhī dialect of West Pahāṛī