Michael Witzel

## Early Loan Words in Western Central Asia:
## Substrates, Migrations and Trade.

Summary

Recently discovered evidence suggests that there is a body of loan-words preserved independently from each other in the oldest Indian and Iranian texts that reflects the pre-Indo-Iranian language(s) spoken in the areas bordering N. Iran and N. Afghanistan, i.e. the *Bactria-Margiana Complex.* These loans include words from agriculture, village and town life, flora and fauna, ritual and religion. They were taken over and then exported to Iran and N. India by the speakers of the various Old Iranian and Old Indo-Aryan (Vedic) languages, as well as a western off-shoot, the Mitanni Indo-Aryan of Syria/Iraq and the language of related tribes indicated by some Indo-Iranian words in Kassite. All these represent series of Indo-Iranian intrusions into the world of the great Mesopotamian, Bactro-Margiana and Indus civilizations.

§ 1.1.  Introduction

Over the past few decades archaeologists have discovered an increasing number of sites of the great Oxus Civilization, perhaps better known nowadays as the Bactria-Margiana Archaeological Complex, as well as its Neolithic and chalcolithic predecessors.[1] While they have filled in a large gap between the great civilizations of Mesopotamia and the Indus, so far no written documents have been found, with the exception of the seal from Anau reported by F. Hiebert.[2] However, little to nothing is known about the language(s) spoken in the areas east of Mesopotamia (Hurrite, Akkadian, Sumerian, Elamite)[3], and those west of the Indus area. The language(s) of the Indus civilization, of course, also are by and large unknown, that is if we neglect the materials that can be distilled from the materials contained in the earliest texts in Indo-Aryan, the Vedas[4] but that have been neglected for that purpose. Nevertheless, these serve as a guide of what language(s) may have been present in the subcontinent in c. 2000 BCE.[5]

The picture can be enlarged and projected back in time by using the oldest Iranian counterparts of the Vedas, the Avestan texts of the Zoroastrians, as well as the Old Persian inscriptions. Since Old Iranian and Old Indo-Aryan (Vedic) are so closely related, items common to both languages can be used to reconstruct the common-Proto-language, Indo-Iranian.

---

[1] Sarianidi 1992, 1998, Dani 1992, Francfort 1989, 1990, 1998, 1999, Hiebert 1988, 1992, 2001.

[2] Elsewhere in this Volume. A few Elamian seals have been found in S. Turkmenistan.

[3] Languages known from barely more than the names given to their speakers in Mesopotamian sources, such as those of the Guti and Lullubi at 2300 BCE, are neglected here. For the contemporary situation on the Iranian plateau, see Vallat 1980, 1985, Steinkeller 1982, 1989.

[4] For such (loan) words see Witzel 1999 a,b.

[5] For a discussion see Witzel 1999 a, b, 2001b, and forthc,. a.

Otherwise, we have virtually no evidence for the areas between the great civilizations and those north of Greater Iran as they are too distant from the Near Eastern, Indian and Chinese cultures.[6] However, the seal recently discovered at Anau should alert us to the *possibility* that early writing might be found in the area after all.[7] In the meantime all that we can establish for the languages used in the western Central Asian area comes from early Near Eastern and Indian (and also Old Iranian) sources. There are some references in the Sumerian and Akkadian documents of the 3nd to 1st mill. BCE, but they deal just with the border areas of Mesopotamia[8] and furnish some only vague ones such as the reference to Aratta, probably Arachosia.[9] Similarly, we have only a few vague reminiscences in the earliest Indian texts (Ṛgveda) composed in the Greater Panjab (c. 1200 BCE-1000 BCE)[10] which seem to refer back to the area between the Volga (*Rasā*); secondly the people along the the River *Sindes* (Tacitus, the Merv or Tedzhen river): the *Dāsa* or O.P. *Daha* (whom the Greeks call *Da[h]ai*), the *Arii*, and the *Parna* (Ved. *Paṇi*?, cf. Ptolemy, Geogr. 6.10.2 *Parnoi, Daai/Parni, Dacae;* otherwise *Dahae*).

In addition, after the sparse atttestation found in the old Iranian and the much later Middle Iranian sources[11] most of these areas have become Turkish speaking about 1400 years ago; this has obliterated much if not most of the older Iranian and IIr. record, frequently also that of topographical names.[12]

Even in this unfortunate situation, we can retrieve, based on the written records of neighboring Indo-Iranian peoples and on old loan words, an increasing amount of details of the pre-IIr/Iranian languages of the area, notably that of the BMAC (Bactria-Margiana Archaeologcial Complex or Oxus civilization, c. 2200-1700 BCE) and of Greater Afghanistan. However, it is precisely these Indo-Iranian sources that have largely been neglected so far.[13]

For some years (1995-2001) I have drawn attention, mostly in brief and passing fashion, to a common body of words in Old Indian and Old Iranian texts that do not seem to be of Proto-Indo-Iranian (thus, Proto-Indo-European) origin. These words represent the non-IIr languages spoken in Iran and in the northwestern part of the Indian subcontinent at the time these texts were composed, that is late in the second and early on in the first millennium BCE. As such, they are invaluable materials for the study of the language(s) preceding the introduction of Indo-Aryan (Vedic) and Old Iranian (O.Persian, Avestan). More importantly, both hieratic texts share a common substratum that can only be that of S. Central Asia. As will be seen below, it cannot come from elsewhere: as both Vedic and Old Iranian individually imported it into their particular habitat, the Greater Panjab and Iran/Afghanistan.

---

[6] For some such data see, however, §2, where the linguistic boundaries of W. Central Asia are discussed.

[7] See however V. Mair 2001, and note however Proto-Elamian seals close by at Tepe Hissar (as well as at Shahdad, Shahr-i Sokhta etc.)

[8] Such as the Guti, Lullubi; similarly, the texts of the Hittite and Urartu realms for the boundary areas of Anatolia, NW Iran.

[9] *Lapis lazuli* is found in the nearby Chagai Hills and in Badakhshan. Note Steinkeller 1982: 250 with details about a green variety, 'carnelian with green spots', possibly turquoise, from Marḫaši. However, the blue Badakhshan variety is more famous, until today.

[10] The lowest date depends on the date of iron, c. 1000 BCE; see Possehl and Gullapalli 1999.

[11] A few texts in Bactrian, Khorezmian, and Parthian as well as (frequently mythical) data in Pahlavi, and in the Graeco-Roman sources.

[12] See Turk. *-su, -kul, -kum,* etc. Needless to say we do not have etymological dictionaries of Turkmen, Uzbek, or even for the Ir. languages, (with the exception of the old one of P. Horn for Persian, and G. Morgenstierne for Pashto).

[13] For initial suggestions see Witzel 1995, 1999a,b.

Such substrate words are quite common in languages that have occupied the territory of an earlier people speaking a different language.[14] In English, for example, such common words as *sheep (Dutch schaap, German Schaf)* belong to the Neolithic substratum of the North Sea coast of Northern Germany and Denmark, the homeland of Anglo-Saxon.[15]

§ 1.2. Sources

In order to evaluate the scarce materials at our disposal properly, a brief look at our sources is in order.[16] The Vedas were composed (roughly, between 1500-500 BCE) in parts of present day Afghanistan, northern Pakistan and northern India. The oldest text at our disposal is the Ṛgveda (RV); its is composed in archaic Indo-Aryan (Vedic Sanskrit). It is followed by a number of other Vedic texts, usually listed as Saṃhitās, Brāhmaṇas, Āraṇyakas and Upaniṣads. Linguistically, however, we have to distinguish five distinct levels: Ṛgveda, other Saṃhitās (Mantra language), Yajurveda Saṃhitā prose, Brāhmaṇas (incl. Āraṇyakas and Upaniṣads) and the late Vedic Sūtras (Witzel 1987, 1997)[17]

The language of the RV is an archaic form of Indo-European. Its 1028 hymns are addressed to the gods and most of them are used in ritual. They were orally composed and strictly preserved by exact repetition through by rote learning, until today. It must be underlined that the Vedic texts are "tape recordings"[18] of this archaic period. Not one word, not a syllable, not even a tonal accent were allowed to be changed.[19] The texts are therefore better than any manuscript, and as good as any well preserved contemporary inscription. We can therefore rely on the Vedic texts as *contemporary* sources for names of persons, places, rivers (Witzel 1999c) and for loan words[20] from contemporary local languages.[21]

The Ṛgveda was composed in the Greater Panjab[22] and is to be dated before the introduction of iron in the northwestern subcontinent. Later texts cover all of northern India up to Bengal and southwards towards the Vindhya hills.

Some 4% of the words in the Ṛgvedic hymns that are composed in an archaic, poetic, hieratic form of Vedic, clearly are of non-IE, non-Indo-Aryan origin, in other words, they stem from a pre-IA Panjab substrate.[23]

---

[14] Exceptions are the territories of Australia, Polynesia, and the Americas when first settled.

[15] Cf. Huld 1990, Polomé 1986, 1990, Th. Vennemann 1994, 1998.

[16] For the sparse Mesopotamian sources, see below (Steinkeller,Vallat).

[17] For abbreviations of the names of texts, their dates and their geographical location see attached list.

[18] The middle/late Vedic redaction of the texts influenced only a very small, well known number of cases, such as the development : *Cuv > Cv.*

[19] They even preserve very *special* cases of sentence intonation, see Klein 1997, Witzel 2001a.

[20] Summary and discussion for RV words by Kuiper 1991; for post-RV texts, see Witzel 1999a,b.

[21] The Vedas are followed by the ancient Tamil "Sangam" (*Caṅkam*) texts from the beginning of our era, all virtually unexplored for substrates and adstrates. -- On the Iranian side, there are sources such as the Pahlavi and early New Persian texts (Šah Nameh, etc.), all beyond the present scope. For place names, see Eilers 1982, 1987, Savina 1964, Schmitt 1995. Such investigations, however, are largely lacking for Afghanistan (note, however, Gryunberg 1980, Pakhalina 1976, Rozenfel'd 1953 for the northeast).

[22] See Witzel 1997, 2001a: roughly, from Eastern Afghanistan, Gandhara, Panjab up to Delhi and even up to the Ganges (twice mentioned); and from the Pamirs/Himalayas southwards to the Bholan area.

[23] This is remarkable: if one were to apply this to a Near Eastern context, it would mean that an ancient Jerusalem temple ritual would contain Philistine, Lebanese, Akkadian, Egyptian or other "heathen" words. The Indian

The situation is similar but not quite as beneficial as far as the Old Iranian texts are concerned. Only about a quarter of the original Avesta has been preserved. The oldest parts are Zarathustra's RV-like poems, his 5 long Gāϑās (Yasna 28-53), and his(?) contemporaneous ritual text embedded among the Gāϑās, the Yasna Haptaŋhāiti, a collection of Mantras used for fire worship. The rest of the Avestan texts is post-Zoroastrian and composed in Young Avestan language. However, the initial oral tradition of the Avesta has been converted in Sasanide times (c. 400 CE) into a written tradition whose surviving earliest manuscripts are not older than a thousand years and have been corrupted by centuries of decline during the early Islamic period of Iran.[24] Nevertheless, the philologically restored Avestan texts offer some data from Greater Afghanistan as Zarathustra's homeland was probably situated in northwestern Afghanistan (near the Kashaf River)[25] and much of the later Avesta was composed or redacted in southern Afghanistan (Sistan, Arachosia). However, in spite of being geographically closer to the Mesopotamian cultures with datable historical information, the Avestan texts are as elusive to absolute dating as the Vedic ones. Mesopotamia (or early China) simply do not figure in all these texts.

The older Avestan texts (Gāϑās/Yasna Haptaŋhāiti) point to a copper/bronze age culture quite similar to that of the RV. The younger texts might overlap with the expansion eastwards of the Median realm (c. 700-550 BCE).[26] The few Old Persian inscriptions that have survived date from 519 BCE onwards. However, other than is the case with old Indian tetxs,[27] the 'foreign' words in the Old Iranian texts have not been evaluated so far. Researchers apparently were of the opinion that only a few could be found; the matter simply has been neglected.

### § 1.3. Loan words and substrate languages

At this stage, a few words about linguistic substrates are in order. 'Words from substrate languages' are defined here as all those words in early Vedic that do not conform to Indo-European/Indo-Iranian word structure (including sounds, root structure and word formation) and have no clear IE/IIr etymology.[28]

We have to distinguish various types of loans (Anttila 1989: 154 sqq). Some are due to cultural and economic contacts, such as the modern *guru* or *karma* (from India), or the slightly older *coffee* (from Arabia), *cocoa, chocolate* (from Meso-America), or *tea* (French *thé, etc.*) whose origin can be traced to S. Chinese (Amoy *t'e*), while the Russian, Indian and Japanese *chai/cha*,

---

sitiuation also differs remarkably from that of the Hittite empire, where the preceding non-IE language, Hattic, was actually used as the ritual language.

[24] Modern recitation depends on these written texts and cannot be used in the same way as Vedic recitation.

[25] Humbach et al., 1991.

[26] Discussion by Skjaervø 1995. However, the YAvest. local name of Bactria (*Bāxδī*) is attested earlier, in the Atharvaveda, see Witzel 1980. -- Current estimates for Zoroaster range from the 14th to the 7th c. BCE. However, an early date is indicated by the name of Ahuramazdā: O.Avest. *mazdā ahura* (or *ahura mazdā*), Y.Avest. *ahura mazdā*, and Old Persian (519 BCE) already one word, $A^{[h]}uramazdā$. For the transfer of Zoroastrianism into the Persis, see K. Hoffmann 1992.

[27] Note the ongoing debate, since mid-19th century, especially S. Lévy, Przylusky, Kuiper, and the relevant summaries in Mayrhofer, KEWA and EWA; last update in Witzel 1999a,b.

[28] Lubotsky (forthc.) adds also some less indicative features: limited geographical distribution, specific semantics, i.e. a category which is particularly liable to borrowing.

Nep. *chiya* are from a N. Chin. dialect. The Indian word has thus come overland and not by sea. The example would also be instructive if we did *not* know the history of transmission: linguists would be able to pinpoint the origin of the loan in two areas of E.Asia. Similar examples will be found below for Central Asian words. This kind of introduction of loan words is from an 'adjoining' language, an *adstrate*. Examples abound in multi-lingual societies (India) or of societies in close contact (ancient and modern W. Europe, with cases such as *street* < Latin *(via) strata, Kaiser < Caesar, castle < castellum, cellar < cellarium, cella;* etc.).

Loans stemming from previously existing languages, upon introduction of a new, dominant language, are different (Anttila 1989: 171 sq). The new language may function as *superstrate*, properly used and understood only by a minority at first (such as Latin in Celtic France), but it then spreads by assimilating an often large number of local words from the previous language, the *substrate*, (note the Celtic place names in England, below).

Sometimes the superstrate does not become dominant (as Norman French failed to do in England); in this case we may still expect a large number of words from the superstrate in the persisting local language (French *beauté > beauty, ancêtre ~ ancestor,* where the English form reveals the older French one, with *-st-*).

Even if the source of the loan remains unknown many loan words from 'foreign' (substrate/adstrate) languages can be easily detected by linguistic means, and even if it belongs to a long disappeared language. The reason is that all languages follow certain patterns, allow only certain sounds or groups of sounds while others are that difficult to pronounce must be substituted by local ones. A typical example from English is that, until fairly recently, German and Yiddish words beginning with the sound *sh-* (*schnitzel, strudel*, to *shlep*) would have been impossible as English allowed only *s-*, as in *snit, strut, slip*). By now, these sounds have been accepted and are pronounced correctly. Similarly, even today words beginning in *ng-, mf-* etc. are not allowed (though by now a few African names have been locally adopted, such as *Mfume*).[29]

Words with such uncharacteristic sounds or sound clusters therefore indicate a certain cultural influence, even if the native speaker (or a latter day scholar) may not know where these words had come from originally.

This is especially true when we have to deal with toponyms and hydronyms that have come down to us from prehistory. It is well known that place names, especially names of (larger) rivers, are very conservative. Even today they may reflect languages spoken many thousand years ago. For example, we have the Rhine (Lat. loan word Rhenus < Celtic *\*Rēnos* < IE *\*reinos*), Danube (Lat. *Danubius* < N. Iran. *Dāna-*), Don (Gr. *Tanais* adapted from *\*Dāna-*), Tigris (Latin, Greek < O.P. *Tigrā,* cf. O.P. *tigra* 'quick'; Arab. *[Nahr al] Dijlat,* both < Akkad. (I)*di-iq-lat* / Sumerian *Idigna*, all from a pre-Sumer. substrate!), Euphrates (cf. Arab. *[Nahr al] Furāt)* which has been taken over from Greek < O.P. *[h]Ufratu* (close to *[h]u-frātar* 'good brother)*',* Sum. *Buranuna* / Akkad. *Purattum* / Elam. *ú-ip-ra-tu-iš,* all from a pre-Sum. substrate more than 5000 years ago.

The early river names of most of Europe belong to one and the same old system.[30] A different prehistoric system is found in Greece and the Aegean area, with the typical pe-Greek

---

[29] Not all loans are as easily discernible as the Amerindian loan words *tipi, squaw, papoose, Manitou,* etc.; note however, the more difficult words *moose < moosu*, chipmunk < *ösitomu,* or woodchuck (Marmota monax) from Algonkian *otchek, ochig, odjik* 'fisher, weasel'; nevertheless, the English folk etymology, however, gives the word away.

[30] Explained, since H. Krahe, as an 'old European' layer of IE (summary by W. P. Schmid 1995); *pace* Schmid, this layer of river names has several elements that seem to differ from, and to predate PIE; note also that many Germanic words or names in the North Sea/Baltic area belong to a pre-IE substrate, see Polomé 1990, Huld 1990.

-s(s)-, -nt-, -mn- suffixes. It has been extensively discussed for much of the 20th century.[31] The detailed investigation of both regions mentioned just now can serve as a guide -- and as warning post -- for the following deliberations.

It is important to keep in mind that names taken from a previous language (or from an adstrate) have more often than not lost their original meaning. If the source language is little known or unknown we can only analyze and compare the outward form of the names involved. This includes the sound system as well as typical suffixes and prefixes that frequently indicate the type of name, such as 'river, place, mountain, plain' or that describe the item in question, such as 'quick/slow, white/black' (river), 'high/low' (place). However, these names have often been adjusted or re-interpreted by later languages, frequently by popular etymology (see above, *Tigris* as *tigra* 'quick'). As may be seen in the discussion of the pre-Hellenic and Old European place/river names, these conditions may lead to many pitfalls. Some may appear in this exploratory paper as well.

The particular situation of Central Asia may be approached by a comparison with that of place names in England. We know that the early form of English, an Old Saxon dialect (a part of the Germanic branch of IE) has overlaid, in the middle of the first mill. CE, the Celtic (and Latin) languages of Britain. Both Celtic and Latin have left a number of loan words in Old English as substrate words, such as London < Celtic *Lugdunum* 'town of the god Lug', *-chester* < Latin *-castrum* 'fortified settlement'. Later on, English saw the superimposed (superstrate) influences of the Viking language (N. Germanic, with words such as *egg*, place names in *-vik, -ay*), then of Norman French with a large number of loans (*beauty, ancestor,* etc.), and finally an equally a huge amount of learned, newly formed Graeco-Latin words, as well as various minor adstrate influences from the neighboring languages such as Dutch (words such as *dike, boss, mate,* etc.). Most interestingly for our purpose, Old Saxon and Germanic in general can be shown to have a large percentage of non-IE substrate words (such as *sheep, eel, roe, boar, sour, lentil, land, delve, prick*) derived from a long-lost prehistoric Northern European language).[32]

The situation in the Greater Panjab (the area of the earliest Vedic texts) and in Greater Iran (the area of the Avestan and O.P. texts) is quite similar. A brief, simplified summary would look like this.

-------------------------------------------------------------------------------------------
GREATER PANJAB                              GREATER IRAN
-------------------------------------------------------------------------------------------
English loan words

Urdu/Perso-Arabic superstr. loans
influx of learned (Sanskrit) words

(development to New Indo-Aryan)         (development to modern Iranian)

influx of learned (Sanskrit) words       Arabic superstrate loans
(dev. of various Prākṛts)                (dev. to Middle Iranian)

Old Greek loan words                      some Old Greek loans
Old Persian/Iranian loan words           Old Persian /Later Avestan

---

[31] Summary by Th. Lindner 1995.
[32] Cf. Hamp 1998: 328, Huld 1990, 1990, Vennemann 2001.

(development from Vedic to MIA)

Later (Ṛg)Vedic / OIA dialects   Old Avestan
Dravidian adstrate      Old Iranian superstrate
   <---- immigrant Old Indo-Aryan   <-----immigrant OIA in Iran
               (*Sarasvati, Sarayu, etc.)
               Central Asian substrate


Harappan language (see below)   Indo-Iranian in C. Asia, south of Uralic, Ket
unknown local language(s)      <------Indo-European
-------------------------------------------------------------------------------------------


§ 2 . Triangulating the Central Asian Area


As has been indicated earlier, we know even less of C. Asia than about the substrate situation of Iran and Northern India since we do not have any old Central Asian written or other traditional records, such as the orally transmitted Avestan texts. In addition, in most of these areas, people have been speaking Turkic languages for the past 1000-1500 years or so, which has obliterated much of the older Iranian and IIr. and pre- IIr. local record. Yet, even there we can make out, based on the written records of neighboring peoples and on old loan words, some details of the pre-IIr/Iranian languages of the area, notably of the BMAC (Oxus civilization) region.

As western Central Asia and the lands south of it were later on occupied by speakers of the various Indo-Iranian languages such as Saka, Avestan, Median, Old Persian, Nuristani, Vedic etc., many of which have left us texts, it is best to begin with this language family. The original speakers of Proto-Indo-Iranian (PIIr, sometimes also called Aryan) have been located in various areas, such as the southern Urals and northern Kazakhstan, the Ukraine and the Caucasus area, or in recent Indian revisionist writing, in even in Northern India.

However the combined data of the reconstructed PIIr language allow to pinpoint the general area where the still united Indo-Iranian proto-language was spoken. PIIr. data reflecting material culture, when compared with archaeological data, can be used to determine a time frame and a date ad quem. All of this points to a copper/bronze age civilization, using the horse-drawn spoked wheel chariot for war and sport, a mainly cattle-based tribal economy, three social classes and a common ritual and a religion stressing both nature worship as well as deities of social obligation.[33] The large amount of PIIr data allows to find loan-words from IIr languages and accompanying cultural contacts with neighboring languages and language families.

These are those of the Uralic (Finno-Ugrian) and Yenesseian (Ket) languages to the north, the three Caucasian language families to the west, Altaic and Sino-Tibetan as well as early on (though attested only much later) an Indo-European language, Tocharian, to the east, and finally

---

[33] Some of these deities may be reflected in the BMAC, see Witzel 2000, and below cf. n. 150-154 for the same origin of some of the IIr deities.

the various Indo-Iranian languages themselves in their post-immigration homelands, in Greater Iran and in the Greater Panjab, to the South.[34]


§ 2.1  The Northern Border

§ 2.1.1  Finno-Ugrian (Uralic)

Starting in the extreme north, it is very important to note that early IIr loan-words are preserved in the Finno-Ugrian (Uralic, PFU) languages. They establish the existence and spread of P-IIr in the areas bordering the PFU homeland in the East European and Siberian woodlands and, thus, the relative age of the speakers of IIr. This evaluation is based on the exact form of IIr. the various loan words preserved in Proto-Finno-Ugrian. Harmatta (1992) has mistakenly subdivided these IIr loans into 17 stages[35] which must be collapsed into just a few linguistically attested stages.[36]

First of all, there are just a few loans that might be attributed to the Proto-IE period, such as PFU *mete 'honey' < PIE *medhu or perhaps PFU *arwa 'price, worth' < PIE *orgʷha-.[37] However, the bulk of the loans comes from the Proto-IIr. (or pre-Proto-IIr) period during which PIE *kʷ, kʷh, gʷ, gʷh became k, kh, g, gh; thus, FU *warkas(e) 'wolf' < PIIr *vṛka-s < PIE *wḷkʷo-s, etc. Later on, *P-IIr *k', k'h, g', g'h developed to c', c'h, j', j'h: FU, F-Volg. *porc'⟨ porśas 'piglet', FU *śaka, śawa 'goat' (Rédei 1986: 59), and still later, PIIr c' > ś : FU *reśmä < *rac'mí.[38] Finally, some forms in FU point to Iranian developments: FU *pakas(e) 'god' < IIr *bhaga-s, IIr. 'share, God Bhaga' > common Iranian (Median, Scythian, Sogdian) 'god, Lord, Mr.' > PSlav. bogᵘ 'god'.[39]

---

[34] Occasionally the Mesopotamian sources can provide some data such as on Aratta, see Steinkeller (1982). Chinese sources are geographically too distant for most of the the pre-Chang Ch'ien period (2nd cent BCE), though some loans from IE into Sinitic can be discerned, see below, n. 62.

[35] Harmatta's (1992: 360-367) specific levels of IIr. are as erroneous as the dates ascribed to them; however, all of this was taken over, with some modification, by S. S. Misra (1992) to establish that *Vedic Sanskrit* was spoken around 5000 BCE in the neighborhood of the Finno-Ugrians and that the various IE languages are *derived* from this hypothetical Central Asian Skt. and from a still *earlier* Panjab-based Sanskrit; for a discussion see Witzel 2001a, cf. Hock 1999.

[36] For further details see Joki 1973, Rédei 1986, 1988, Katz 1985, 2001.

[37] Rédei 1986: 44; not < IIr *argha, Mayrhofer, EWA 114, rather from *arɣa, Finn. arvo, cf. Sogd. 'rɣ. Rédei has the following oldest loans: *miɣe 'to give, sell', *muśke 'to wash', *nime 'name' (which may rather be Nostratic, cf. Jpn. na- etc.), *sene 'sinew', *toɣe 'to bring', *waśke 'metal', *wete 'water'. Note the preponderance of words relating to exchange/commerce.

[38] See Rédei 1986. There are problematic representations of certain vowels in FU when compared with their successive sources in (P)IIr. or older (pre-IIr): old IE, Pre-PIIr -o- + PIIr -c'- in F-Volg. = IIr *porc'os, or Common/Late PIIr (k > c') in F-Volg. *oraśe < IIr *varāj'ha < PIIr *wæroj'ho, and in later IIr forms (c' > ś, j' > j), as in F-Volg. *waśara < IIr vaj'ra < PIE *h2weg'-ro. The FU representation is due to the nececessities of the FU sound system, FU vowel harmony, and certain substitutions (e > i, o > u), as seen in F-Volg.*reśmä < *ræc'mí. Cf. also Rédei 1986: 33sqq. for similar substitutions of Iran. vowels in Permian (Wotyak) loans, such as a > o, u; ā > u, etc. -- Harmatta's FU and IIr/Ir. forms (1992) are not reliable.

[39] This word, via normal Slavic sound shifts, from N. Iran. (Skythian) baga- 'god', like so many of the Ukrainian/S.Russian river names (Don, Donets, Dnyeper, etc. < dānu 'water'). Only in some areas of Iranian, the IIr word *bhaga "(God) Share " has developed the meaning 'god, lord'. The word thus is derived from Old Iranian, not

§ 2.1.2 Another northern neighbor is the isolated Siberian Ket language that is still spoken on the middle Yenessei river.[40] It belongs to a group of related languages (Arin, Kott, Assan, Yugh, Pumpokol), that have now largely disappeared. Their toponymy shows that they must have covered a much larger territory, from the lower Yenessei to Tuva, and from the Middle (and to a lesser degree west of the Upper) Irtysh almost to the Angara, in short, roughly the territory between the Irtysh and the Yenessei-Angara rivers (Vaijda 1998: 10). Their place names in Ket *se's* 'river', *as* well as in *-ces, -tet, -set, -šet; -ul*; note also Yen. *\*xurĪ* 'water', Kott *kem*, Pump. *tōm* 'river'.

Like PFU, the Ket languages have a number of old loan words: Kott *art`a* 'true, veritable' < IIr *\*ṛta*, Ved. *ṛta*, OAvest. *ərəta*, O.P. [*ərta*], Mede *arta*, YAvest. *arəta*, Mitanni (and Greek historians) with the spelling *arta-*; Kott *c'āk* 'force', Kott *c'aga* 'strong' < IIr *c'ak*> Ved. *śak* 'to able; force,' Ket *ku'š*, Yug *ku's*, Kott *husa*, Arin *kus*, Pumpokol *kut* 'cow' < Pre-PIIr *\*gʷāus* > PIIr *gāus*, Iran. *gauš* and note Tocharian A *ko*, B *keu,* but cf. also Yen. *\*kuqs*[41] (cf. IE *\*gʷōu-s*).

§ 2.1.3. Some words may shed more light on the older, central Asian language(s) of the area that are now lost. IE *\*medhu* 'honey, mead'[42] has already been mentioned. It appears as Ved. *madhu* 'sweet, honey, mead', Avest. *maδu* (cf. Bur. *mel* 'wine, from grapes'), Sogd. *mδw, mδw* 'wine' > Toch B. *mot* 'brandy', Toch. B *mit* 'honey' < *\*m'ət* < IE *\*medhu*, Gr. *méthu* 'wine' etc. and it has been widely borrowed by neighboring languages, both northwards into Uralic *\*mese, mete*: Finn. *mete*, Hung. *méz* 'honey', as well as eastwards into Proto-Turk. mïr, Chin. *mi* < *\*mjit/mit*,[43] Sino-Kor. *mil*, Jpn. *mitsu* < *\*mit(u)*. Its Iranian form, Iran. *\*maδu* > Turk., Mong. *bal* 'honey' (note Arab. *mādī?*); finally also into the IE language, Toch. B., in its specialized meaning *mot* 'intoxicating drink'.

However, there is also another source of the word for 'honey, mead', which must be different from the C. Asian source, *\*\*med(h)-*, discussed so far: Gr. *mélit-*, Hitt. *milit,* Lat. *mel, mell-*, Gothic *miliϑ* point to a more western source, *\*\*melit,* perhaps in the Balkans/Anatolia. Note that the early reconstructions of Nostratic[44] list both forms under *\*majlʌ* > *Ural. majδ'ʌ,* Drav. *maṭṭ, miṭṭ* (DEDR 4662 *maṭṭu*), Altaic */m/ala, bala.*[45]

directly from IE. -- PFU *\*pakas(e)* 'luck' > Mordvin. *pavas* 'luck' shows the older meaning , from IE *\*bhagos* > PIIr *bhagas* 'share', see above.

[40] For a detailed discussion see several articles in the journal Mother Tongue IV, Boston 1998: 4-32.

[41] Contrast 'horse': Yugh *ku's,* Kott *huš*, Arin *kus, qus*, Assan *huš*, see n. 65.

[42] Cf. Lubotsky 1998: 379.

[43] Since Polianov (1916), see Lubotsky 1998: 379

[44] Nostratic is the reconstructed source, at c. 10,000 BCE, of IE, FU (Uralic), Altaic, Dravidian, Kartvelian (SW Caucasian: Georgian etc.), and Afro-Asiatic (formerly, Hamito-Semitic), see Illich-Svitych 1976 : 38 sq.

[45] Note even beyond this area, in Polynesia, Samoan *meli*, Hawaiian *mele, meli*; *mele, melemele* 'yellow', Maori *miere*; Tongan *melie* 'sweetness, sweet, delicious', Rarotongan *meli* 'honey', Mangareva *mere* 'honey'. This must be old and cannot just be an introduction from French *miel* (in New Zealand, Hawaii?) as we have Haw. *mele-mele* 'a star name': *Melemele* and *Polapola*, 'the twin stars', the former male, the latter female; at any rate, this is already Proto-Nuclear Polynesian (Samoan, Futunan etc.) *\*melemele;* cf. also Haw. *Melemele* 'a mythical land'. -- For the spread of the word, note the role of *cire perdue* with its inherent use of bee's wax, see Andrew Sherratt, elsewhere in this volume: Patterns of Growth: Nodes and Networks in the Ancient World.

All of this points to an early northern boundary of IIr, coinciding more or less with the heavily wooded Taiga belt of Russia and Siberia.[46]

The old Central Asian word for 'lion', \*\**sengha/singha* has a similar spread and variation Ved. *siṃha* 'lion' < \**sinj'ha* < \**sing'ha.* However this differs, even within Indo-Iranian, from the Proto-Iran. form \**sarg* that has resulted in Khoresm. *sarɣ*, Parth. *šarg*, Khot. *sarau*, Pahlavi, N.Persian *šēr* (Horn 1893, no. 803).[47] One will have to compare Tib. *seṅge*, Chin. \**suân-ŋei* (Henning),[48] Arch. Chin. \**,ṣi* (Karlgren; mod. Chin. *shi-zi*), Jpn. \**si* > *shi(-shi)*. However, Toch. A *śiśäk*, B *śecake* 'lion' and perhaps also Armen. *inc, inj* seem to go back to a S. and E. Central Asian variety like \*\**si(n)c* which is close to pre-Ved. \**sing'h,* but with unvoiced final consonant. -- Similar to the case of 'honey', the more western IE languages have taken their word for' lion' from a different source, which, in this case is an *entirely* dissimilar Balkan or Mediterranean(?) one: Gr. *līs, leon(t)-*, Lat. *leon-.*

It is imperative that more such data are identified and collected to arrive at a closer picture of the northern parts of W. Central Asia.[49]


§ 2.2. The Eastern Border : Xinjiang


It is much more difficult to specify the linguistic eastern border of W. Central Asia. For the early periods around 2000 BCE and for more than a thousand years onwards, the situation in Eastern Central Asia (East Turkestan) is a lacuneous as that in W. Central Asia. To some extent, place names that can fill in the gap. However, there exist only few studies of these names in western languages, the case of Tocharian excepted.

A first hint may be provided by the names from Sogdia, on the SE border of W. Central Asia. Apart from very scanty mentioning in O.P. inscriptions and Avestan texts, it is Herodotos and the Alexandrian historians who transmit the first notices: a few personal names and relatively more place names. However, most of them (such as *Mara-kanda,*[50] *Zari-aspa*) are already of O. Iranian etymology and hardly go beyond the eastern boudary line that is of interest here.

They can be supplemented by modern place names such those of the sole descendent of Sogdian, Yaghnobi (Zerafshan valley),[51] or by those from the Pamirs.[52] Gryunberg (1980: 168)

---

[46] Note, again, A. Sherratt on the spread of bronze smelting and cire perdue casting along the *taiga* belt.

[47] Reconstructed by Henning (see KEWA, s.v. *siṃha*) as \*\**s1eṅgha*; however, there must have been another form from another Central Asian dialect or language, based on \*\**šer(gh)*; cf. below § 5, on the interchange of *r/n*. Not related, apparently, is Turk. *arslan* 'lion.' -- Note that there existed, in eraly history a sub-variety, the 'Iranian maneless lion' (Masson 1992: 39); another remnant population of the Asian lion is found in India, in the hills of Girnar (W. Gujarat).

[48] See now Pulleyblank 1995; Lubotsky 1998: 379.

[49] Leaving aside the speculations of Harmatta (1992) and worse, Sergent (1997) on Dravidians in C. Asia, including Harmatta's unlikely *Haftyār* from Ir. *haft* < IIr \**sapta* 'seven' and Drav. *(y)ār(u)* 'river'; however, *-yār* is a common N.P. suffix, see Horn 1893: 251. Or, Harmatta's Sumer. *Aratta* (= Shahr-i-Sokhta, Sistan) ~ *Arava* 'Tamil man', *Araviti* 'Tamil woman', *Aratta* 'Tamil settlement' or 'Tamil land', as well as his wide-ranging speculations on *Kaššu,* Kassites, Kashgar, and the Caspian Sea (1992: 370 sqq).

[50]E.Iran. *kanda* 'town', cf. Pāṇini's *kanthika* "inhabitant" of Varṇu, mod. Bannu; see below n. 159, 162 for a list of Sogdian names; cf. Khromov 1960, 1980, E`del'man1980.

[51] Khromov 1960 specifies suffixes such as *-ef, -uf,* (< Sodg. *-uv*), *-ic* (< Sodg. *-yc*), *-kam (-kand)* < *kt* 'town', *-yar* < *ɣr* 'mountain', *-zoi* < *z'yh* 'earth, land', *-rūt* < Sodg. *rwt* 'river' - *rovut* (cf. Tajik *rāvad*) 'meadow'.

[52] See Savina 1980: 140-155; E`del'man 1975.

gives a long list of relevant place names from Afghan Badaxšān[53] and specifies that these 'substrate' names point to a widespread language, or at least to certain toponymical area based on a Pamir language, probably Sogdian. Indeed, it must be noted that some of the names mentioned are clearly formed with typical later, E. Iranian suffixes (*-iw, -mand*, etc.), often the same (see Khromov 1960) as found in the successor to Sogdian, modern Yaghnobi.[54] It remains to be seen which real substrate names remain when the individual etyma used in the toponyms have been etymologized backwards to an early Iranian or otherwise, to an unknown local language.

Data[55] from the surviving E. and N. Iranian (Saka) languages could be added, including the only eastern descendent of Saka, Sariqoli, already on Chinese territory.[56] Further, the contribution of Burushaski and its earlier forms[57] as well as the substrate present in the IA Pamir language Khowar[58] should be compared.

A closer study of the local names in the (Gāndhārī) Niya Prakrit in S. Xinjiang[59] with personal names such as *Şekrase, Bhimase?, Bugosa,* and of the substrates preserved in Tocharian,[60] with place name such as *Yursa, Kuci, Hippuka?,*[61] would yield further information on the eastern neighbors of prehistoric W. Central Asia. As could be expected, Tocharian, which seems to have split off from PIE at an early stage (Hamp 1998), and moved into Xinjiang at perhaps c. 2000 BCE, shows contacts with Altaic languages (Róna-Tas 1974, 1988), especially early Turkic (Clauson 1975, Pinault 1998), and since Shang times with Sinitic (Pulleyblank 1966, Lin 1998)[62] as well as with Tibeto-Burmese.

All of this leads, however, much beyond the frame of the present study.[63] Neverthelesses, it should be pointed out that there are very early loans that can indicate a network of languages connecting western and eastern Central Asia. These include the words for 'lion' and 'honey' that have already been discussed (above § 2. 1). The substrate word **sengha/singha* 'lion' straddles the

---

[53] With "substrate toponyms" such as *Karniw, Marc, Malmunj, Muzung, Remān, Raymānd, Khewurz, Dawang* (river), *Šinj, Ādnyal, Zū, Elk, Šalil* etc. She specifies from the side valleys of the rivers Kufar (Darvaz) and Zardev (Sargulyam): *Roghūd, Wyāj, Rawinj, Ghezw, Rāj, Purzārg, Našer, Imj, Kher; Yāsic, Iwinak, Sucu, Yakhcew, Ezwān, Pijangīw, Bušt, Winj, Afrij, Korkhu, Bahārak,* etc. She concludes that in Yaftal, Rog, Darvaz, and Sargulyam there was a common toponymy; however, she also regards it possible that there were several successive levels that built up to the present substrate.

[54] For the same area, L. Dodykhudoeva (2000) lists a number of interesting Tajik "substratum words" that have been taken from East Iranian languages. Some of them seem to have no clear Ir. etymology. -- Note that the 'suffixes' in West Iranian (Persian) toponymy differ to a large degree, see Savina 1964.

[55] See also the discussion of many linguisticx details of the Greater Pamir area in E`del'man's (1968) discussion of Indo-Iranian linguistic geography.

[56] The western variety of N. Iranian is preserved in the Caucasus as Ossetic, a descendant of Alan.

[57] Old Burushaski names in v. Hinüber 1995; comparison with Basque and Caucasian by Bengtson 1992, 1999, 2001, cf. Harmatta 1992.

[58] See Kuiper 1962: 14, Witzel 1999a,b.

[59] See discussion by von Hinüber 1995: 663.

[60] For the contacts between Tocharian and Sinitic see Pulleyblank 1966, 1995, between Tocharian and Altaic Clauson 1972,1975, Róna-Tas 1974, cf. also Róna-Tas 1980.

[61] O. von Hinüber 1995: 662; cf. the place names in Pinault 1987, 1998.

[62] Candidates for late IE/Toch. loans into Chinese include such words as *lak* 'milk' (note the Himalayan language Bangani with non-IA *lɔktɔ* 'milk, and cf. below § 6 for a possible arrival of the PBangani speakers from C. Asia); *mit* 'honey'; *khwin'* 'dog' -- (cf. TB *kwiy* 'dog/puppy').

[63] Early contacts of Sinitic with Tocharian and/or other IE, and those of Tocharian with Altaic are excluded here; see Pulleyblank 1966, Lin 1998, Róna-Tas 1966, 1974, Clauson 1975.

whole area, from Armenian and Iranian to Tibetan, Tocharian and Chinese, and so does \**med(h)-/ melit-* from Finno-Ugrian and PIE to Chinese, Japanese on the one hand and Hittite to Latin on the other (both forms from Nostratic *maj**l**ʌ*?), In this connection it should be noted that one incentive for the eastwards spead of the word for honey may have been that of the use of bee's wax in the early copper casting technology of *cire perdue* (see A. Sherratt, elsewhere in this volume) that spread along the *taiga* belt of S. Russia and S. Siberia.

Further, we have a great number of local words for the horse, which will originally have meant 'wild horse' that was hunted and consumed by Neolithic people (e.g, at Botai in Kazakhstan, 3300-2700 BCE, Mallory 1998). Though this is not the place to begin a comprehensive study of all Eurasian words for the 'horse' the following may be mentioned.

IE *h₁ek'wo-s*[64] :: FU *\*lox* :: Yenesseian (Ket) *\*kuqs*[65] :: Turkic OT *\*(x)at* '(riding) horse' (Janhunen 1998: 415, note *yunt* 'horse' Róna-Tas 1980: 378); however, the word is represented in a number of other language families: Caucas. *\*k'otu*,[66] Hittite *kur-ka* 'foal' (EWA 373), Sumerian *anše kur* 'mountain ass' (rather the *kur*-donkey?), Drav. *kutir(ai)*,[67] Munda (Koraput) *kurtag*,[68] Korku *gurgi* (= *kurki*), Sabara (Sora) *kurtā*, Gadaba *krutā* < *\*ghurta*, Tib. *rta,* but Tsangla (in Bhutan) *kurtā*,[69] Meithei (Manipuri) *sa-gol*,[70] Bur. *ha-ɣur* < *\*saɣúr*,[71] Drav. *gur-*[72] Brahui *(h)ullī*, Khasi [*u*] *kulai*, Amwi *kurwa'*, Indones. *kudja;* note also Late Vedic loan word *ghoṭaka*, (modern NIA *ghoṛ-,* etc.). The same word seems to have been introduced to S. and S.E. Asia along various routes (including those across the eastern Himalayas) after the domestication of the horse. As Berger (1959: 32) put it: "das weltweit verbreitetete *\*kur/kul*". It seems to be reflected even in the relatively late German *Gaul* < Middle High German *gūl* 'worthless horse; male animal, (e.g.) boar' which would come from an isolated, hypothetical IE *\*ghūl-*. Taking into account also Yenesseian *\*kuqs* and Turkic *\*(x)at*, one may posit an older Eurasian form *\*\*kuC : *kur/kul :: *kuq / * h₁kʷ-/ *k'ot-u*?

Finally there are further, isolated designations such as Munda *sadom,*[73] Eastern Himalayan (Dhimal) *ōnyhā*, Lepcha *on*; Drav. *\*(h)ivuli*,[74] Tib. *rta*[75].

On the other hand, there exists also another very wide-spread, in fact, common Eurasian, word, for the horse: *\*\*mar-/mor-*. To begin with the Central Asian nucleus of the term, we have

---

[64] IncludingTocharian A *yuk*, B *yakwe*.

[65] Arin *kus, qus*, Assan, Kott *huš,* Yugh *ku's,* cf. Werner 1998: 20.

[66] Tuite 1998: 464, cf. Andij *k'otu* 'horse' (Andic, belonging to the Avar-Andic branch of NE Cauc.).

[67] Drav. (DEDR 1711 b) Tam. *kutirai,* Tel. *kudira, kudaramu,* Kota *kulyr,* Toda *kïϑïr,* Kodagu *kudïre* have been compared with Elam. *kuti* 'to bear', *kutira* 'bearer' (McAlpin 1981:147-8; Southworth 1979: 181). Note also Tirahi (Dardic) *kuzəra* (Harmatta 1992: 375, cf. CDIAL 3219?).

[68] Drav. (DEDR 1711) *kutirai* 'horse' > Koraput Munda *\*kuXrtag*, see Zide & Zide 1976: 1331.

[69] Shafer 1954: 25 *\*ghurta,* contra J. Bloch's *\*ghutr-* (Bagchi 1929: 24).

[70] Berger 1959: 27, n. 36, p. 33 sqq. (Ling. Survey I, pt. II, p. 105).

[71] Berger 1959: 27 n. 36.

[72] DEDR 1711(b): Tel. *gur̠r̠amu,* Kolami *gurram,* Naikri *ghurram, kuramam, kurrmam,* Parji *gurrol,* Gondi *gurram,* Konḍa *gur̠am,* Kuwi *gūrumi, gurromi* 'horse'; (for Burrow on Tel. *gur̠r̠amu* < Skt. *ghoṭaka*? see IJDL 1, 23-24).

[73] Pinnow 1959: 78 § 70: "hardly to be connected with Khmer *seḥ,* Bahnar *y̆šeh*".

[74] Tam. *ivuḷi* and Brahui *(h)ullī* 'horse' (< 'half-ass, hemion', Burrow 1972, DEDR 500); -- other Drav. words include: *DEDR 500* Tam. *ivuḷi,* Brah. *(h)ullī,* 1711 Tam. *kutirai* etc., *DEDR 3963* Tam. *pari* 'runner', 4780 Tam. *mā* 'animal' (horse, elephant), Tel. *māvu* 'horse'.

[75] Tamang, Gurung *ta,* but cf. *kurta(g)* (above).

Modern Mongolian *morin, mörin* < \**morï*, Tunguse *murin* (borrowed into Ghilyak as *mur, murng)*, Korean *mar, mal* < *mål*), Japanese *uma (mume)* / Ryukyu *nman* < \**uman,* \**mVrV*, Chinese *ma* < \**mraq (mra')* < \**mVra,* (or \**mwa,* Benedict 1972: 189), Dun-Huang Tib. *rmang*, Burm. \**mrang-h* > *myin-h,* Kachin *gù-mrà;* Thai *ma,* note Drav. *mā(v)*[76], Nahalī *māv.* Janhunen (1998: 415) considers them to have been "introduced to East Asia form a single source,[77] possibly by a single wave of cultural impact," -- perhaps all from a Proto-Altaic \**mVrV,* a proposition dismissed by Janhunen (1998: 424).

On the other hand we also have, at the other end of Eurasia, the isolated W. IE \**mar-ko-* 'horse' (Pokorny 1959: 700) as in Celtic: Irish *marc*, Cymr. *march*, Gaulish *marko-*; Germanic: O.Norse *marr,* O.Engl. *mearh,* Engl. *mare,* O.High Germ. *marah,* fem. *meriha, Mähre,* cf. *mar-shall.* Due to their isolation, Janhunen (1998: 425sq) does not wants to connect western IE \**mar(ko)-* with the Central Asian \**morï,* and regards any similarity as accidental (likewise, Lubotsky 1998: 385).

However, this view may be too limited, as the Eurasian correspondences of \*\**gul/kur* 'horse' may indicate. If we think, somewhat along the lines of J. Nichols (1997-98), of a widespread pre-proto-Germanic area that extended from N. Europe backwards to Northern Russia, before others (Balts, Slavs) moved in,[78] then the second Eurasian word for 'horse', \*\**mar/mor,* may have come that way to W. Europe. It may have designated the non-domesticated horse, just as \*\**kur/gul,* while IE *h₁ek'wo-* may have referred to the domesticated one, taken from the adjective \**Hōk'u-* 'quick' (EWA I 179). It is well known that the homeland of Equus caballus is in the steppe areas of Eurasia, while pockets of the habitat of the wild horse were also found in the forest zone. From these areas, the horse has been secondarily introduced into the Near East, South and South-East Asia (see Meadow 1998).

An important, perhaps much older cultural loan word is that for the 'hammer.' There are a number of similarities between the words for "stone, hammer, axe" in various neighboring languages. That one and the same word could have these meanings should point to stone age, Neolithic level at the latest. "Altaic" \**paluqa* 'hammer (Mayrhofer, EWA II, 214 s.v. *paraśu)*; Bur. *baluqá* 'big hammer', *bulqá* 'kleiner Hammer (zum Beschlagen der Pferde)', Yasin Bur. *baluqá* (Berger); IE \**pelek'u-* > Greek *pélekus,* Ved. *paraśu-* (< IIr \**parac'u,* cf. CDIAL 7947 *parśu)*, O.P. \**paraϑu* < W.OIr. \**parasu* < IIr. \**parac'u.* However, there also is a variant PIran. \**paratu* > Osset. *færæt.*[79]

The overlap between 'stone' and 'hammer' is also seen in the pan-Eurasian children's game of 'stone, scissors, paper' (or, whatever our "scissors" and "paper" might have been in Neolithic times!). The Bur. children's game of 'hammer, scissors, sword' has *baluqa* 'hammer' which obviously is derived from 'stone' (cf. Shina *bulká*). Such change in meaning between 'stone' and

---

[76] DEDR 4780 Tam. *mā* 'animal' (horse, elephant), Tel. *māvu* 'horse, (cognates mean 'deer' etc. in other Drav. languages!); if derived from \*'horse', Drav. may have transferred the word to other large animals; cf. the (by now IA speaking) Nahali with *māv* 'horse'.

[77] Similarly, Benedict (1972: 189) thinks of a central Asian form (\*\**mərən*) as seen in Mong. *mörin,* which may also have resulted in Tib.-Burm. \**mraṅ, sraṅ,* from a bisyllabic word \**m[]raṅ[].* -- Note that Mong. *-in* in *mörin* is unexplained for the point of view of Altaic.

[78] This particular point is not invalidated by the critique given below of her original locus of IE in Bactria/Sogdia, or by the early criticism of this thesis by Mallory (1998).

[79] Loan word in PPerm., Wotyak etc. *purt,* Tunguse *purta* 'knife', see Rédei 1986: 76. -- Note that the often compared Akkadian *pilaqqu, pilakku* does not mean 'axe' but 'spindle, dagger.' (EWA, s.v. *paraśu)*.

'hammer/weapon' can sometimes still be seen in older myths. For example, the Old Norse *Thor* still kills his enemies with a stone hammer, as does his mythological relative, the Vedic *Indra*.

All these sporadically available terms and names indicate close contacts between the regions east and west of the Pamir-Tien Shan-Altai ranges, but do not allow, at this stage, to draw up a history of their settlements.


§ 2.3 The Western Border : Caucasus and NW. Iran


While we do not have old texts from the western borders of C. Asia, the older languages at its western rim that have not succumbed to the successive spread of IE, Iranian and Turkish[80] can supply some of the desired data.

First of all, IE or rather IIr influence is seen in some of the Caucasus languages,[81] some of which may have retreated from the plains into the mountains, just as the N. Iranian (Saka) language Ossetic did more than a thousand years ago. Such influence can be seen in Udi *ek`* 'horse' < PIE or, pre-PIIr ($h_1$)*ek'wo-*. From common PIIr stem *ac'wa* > Abkhas-Circassian *čhə*, Abkhas *a-*c'ə 'the horse', Andic (Akhvakian) *icwa* 'horse', Dagi *urci*, Andian *ica*, Avar and Lak *cu*, Xinalug *pši*, Proto-Daghestanian *Wci (W = labial prosody)*; Kartvelian: Georgian *acu·ca (acu* to urge on horses)*, acua* 'horse' (nursery word); from a later, but still common PIIr source *aśʷa* comes Circassian *śə*, Kabardian *šə* 'horse'.[82] A clearly late loan from Old Iranian is Khinalug *spa* 'colt' < N. Ir. (Scythian, Proto-Ossetic) or from NW. Ir. (Median) *aspa*, cf. also Avestan *aspa*. -- To an early level may belong PIE *gʷou* 'cow' > Nakh *govr* 'horse', Kabardian *gwaw* 'bull'.[83]

However, the case is entirely different when it comes to the designation of the old domesticated animal, the goat, PIE *Haig-*. Domestication has begun in mountains of the fertile crescent (Zagros Mts.), from where it spread at the end of the Younger Dryas (c. 9,600 BCE). It is often assumed that IIr *aj'a* was loaned into Circassian *ac'a*, or that P-Iran. *aza* > Kabardian *aza* 'goat for breeding'. However, the etymology of IIr. *aj'a* is not clear at all. Usually it is compared to Greek *haik'-s, haig'-*, itself without clear etymology.[84] It is better to compare it to the range of old agricultural and pastoral words extending east and west from the Caucasus, as found in J.

---

[80] Cf. Nichols 1997, 1998.

[81] Note the materials supplied by J. Nichols 1997: 125-129, 143.

[82] Cf. the materials in Harmatta 1992: 369.

[83] However, Chechen *gowr* and Ingush *gawr* 'horse' < IIr *gāura*, Ved. *gaura* 'whitish-yellowish-reddish, wild ass, wild buffalo'; Iran. *gaura* > N.Pers. *gōr* 'wild ass'; cf. also Pashto *γγara* 'wild ass' (Morgenstierne 1927, no. 337), which must be kept seperately. A similar development might have occurred with IIr., Ved. *khara* 'donkey', Avest. *xara*, Pashto *xar* 'donkey', NIA *khar* 'donkey', *khōr* (etc.) 'onager, wild half-ass', for which note Morgenstierne 1927: 97 with Pashto *xaṛ* 'muddy, turbide, dirty brown', etc. Cf. also Rau 1980/81.

[84] See EWA s.v. *aja*, but cf. Macro-Cauc. words Bur. *ac'ás* 'goat, sheep, small cattle', Cauc.: Adyge *ā⁻a* 'he-goat', Dargwa (Akushi) *'eža* ~ (Chirag) *'a⁻:a* 'goat', etc. < PNC *ĕjʒ'wē* (Bengtson 2001). -- These forms are to be contrasted with cf. Nakh *gāza* 'goat' (loans from IE into Nakh have *d* > *z*), Lak and Dragi *gada* 'kid' (Nichols 1997: 129) and with dial. PIE *ghaid-* in Germanic (Engl. *goat*), Italic (Latin *haedus*); -- EWA also notes Slav. *koza* < IE *kag'*, next to IE *Hag'* > Skt. *aja*, Avest. *aza*, Lith. *ožȳs*, Latvian *âzis*, etc.; cf. Gr. *aik-s, aig-* and the clearly related Macro-Cauc. words quoted above. -- Cf. also Ved. *chāga* 'ram', Osset. *sæg(æ)* 'goat', and FU correlatives (Mordwin. *śava, śeja*), see EWA I 558; to be connected are the following Macro-Cauc. words: Bur. *c.igír* (Y), ~ *c.higír* (N.) ~ *c'hiír* (H), Cauc.: Karata *c':ik'er* 'kid', Lak *c'uku* 'goat', etc. < PNC *z̲·kV / *k̲·ZV, Basque *zikiro ~zikhiro* 'castrated goat'.

Bengtson's Macro-Caucasian (Basque; PNC *ʾ ējzʹwē :* NE Caucasian = Nakh-Daghestanian;[85] NW Caucasian = Abkhas-Adygian; Burushaski *ac'ás*).

   To the same substrate may belong Engl. *buck*, attested in Celtic and Germanic 'male goat, buck', Avestan *būza* 'male goat' and Cauc.[86] (PNC *\*b[a]c'V*), to which we can now add the Indian substrate word *\*bokko* (Turner CDIAL 9312), such as Nep. *boko;* cf. also Burushaski *buc.*

   A similar pattern can be discerned for the spread of the word for 'wheat', first developed in the western Fertile Crescent about 10,000 BCE, from where it quickly spread north and west by population expansion, for example as seen in Macro-Caucasian (Basque *gari* 'wheat', Bur. *gur*).[87] Its southwards and eastward move was slower. It took 2,000 years to reach the Nile valley (O.Egypt. *xnd*). It is found in the Caucasus area as *\*ghond* and on the Iranian plateau[88] as a later *\*gant-um-,* and it is first seen in S. Asia in E. Baluchistan (Mehrgarh) at c. 6500 BCE (discussion below, § 5).[89]

   It is likely that several of the river and place names in S. Central Asia once belonged to the Macro-Caucasian language family, such as the *Sindes*, attested north of the Caucasus (R. Kuban) , in Turkmenistan (R. Tedzhen), E. Iran (Avest. *həndu*) and in the Indus valley: Bur. *Sinda* / Skt. *Sindhu* (see below, § 4).[90] Tuite (1998: 449) has indeed given some ethnological and linguistic reasons for a possible gradual movement of speakers of Proto-Burushaski eastwards towards the Pamirs.[91] The NEC *\*šᵀor-* 'water' is reminiscent of (Gr.) *Silis* 'Syr Darya', and *\*ɬᵀenc'o-* 'river' of (Gr.) *Sind-es*, IIr *\*Sindhu,* and Bur. *sende.* I leave the decision to specialists of Macro-Caucasian.

   Whether this scenario of an widespread Macro-Caucasian presence between the Caucasus and Pamir mountains in pre-IIr. times can be substantiated or not, the evidence presented so far (and that in §3 ) does not allow to state how far the BMAC language once spread into C. Asia.

---

[85] Subdivided into Avar-Andic, Tsezic, Lak-Dargwa, Lezgian.

[86] Lakh *buxca < \*buc-xa?* 'he-goat (one year old)', Rutul *bac'i* 'small sheep', Khinalug *bac'iz* 'kid', etc. < PEC *\*b[a]c'V* (Bengtson 2001); -- cf. also Nakh *\*b'ok'* 'male goat' (Nichols 1997: 128).

[87] According to Bengtson (2001): Bur. (H, N,Y) *gur, gurgán* (H, N) 'autumn wheat', Cauc.: Tindi *q':eru*, Archi *qoqol*, etc. < PEC *\*Gōl'e*, Basque *gari* 'wheat' (combinatory form *gal-*).-- Here belong also J. Nichols' (1997-8) words of culture such as *\*woino* 'wine', etc.

[88] For the beginnings of agriculture in Southern Turkmenistan (Kopet Dagh, Jeitun Culture, already with mud bricks, c. 6000 BCE), see Harris and Gosden 1996. They assume import of domesticated wheat, barley(?), sheep and probably goats from further west. -- However, the dates of Ak-Kupruk (N. Afghanistan, of seasonal pastoralists?) are very much under discussion (between 8000 and 1800 BCE!), see Dani 1992: 124-126, Harris 1996: 384.

[89] See Meadow in Harris 1996: 390-412.  Note the import of domesticated wheat to Mehrgarh, c. 4500 BCE, as opposed to a supposed local domestication (see Meadow 1996: 395), which fits the linguistic pattern.

[90] The following list of NE Caucasian designations (Colarusso in Mair 1998, compared with Bengtson 1999) should be counterchecked against C. Asian Iranian and NW-Indian place names: *\*ɬᵀem- / \*ɬᵀu-, \*šᵀor-* 'water'; *\*ɬᵀenc'o-* 'river'; *\*xᵀul- (\*Hwi\*lV,̵ ɬwiri* Bengtson 1999: 49) *\*rélo-* 'sea'; *\*'ōr, 'ūr* 'lake' (*\*ħwiri* Bengtson 1999: 49); *\*mɣērō- / muɣrō-* 'mountain'; *\*qᵀunt'u- / qᵀunt'o-* 'hill'; *\*'arcᵀi, \*naq'o-* 'land'; (especially elements for 'water' such as *\*ɬᵀeN, \*'ōr, \*'ūr, r-).* ---- Colarusso(1998) mentions as criteria for substrate words: ethnonyms, non-native vocabulary, oronymy and hydronomy are the most conservative toponymy, and surviving languages, gives toponyms and hydronyms from three Caucasian families (with no *apparent* C.Asian connections except for NEC, see previous note); cf. also Murzaev 1980.

[91] In fact, it has been suggested that the population of Gorgan and S. Turkmenistan (Kara-Kamar) "reflects the spread of Mesolithic people from the Zagros mountains to the northern foothills of the Hindu Kush via the Caspian coast." (Sarianidi 1992: 124).

This is in spite of, and converse to, the efforts of J. Nichols (1997, 1998) who assumed a Bactrian/Sogdian homeland ('locus') of PIE (1997: 135, 137, 1998: 233)[92] and a still earlier one of pre-PIE south of the Caspian.[93] This location is not born out, and in fact contradicted by the host of words discussed here.[94] Her locus of PIE is, indeed, located precisely in those areas that represent an older, *non-IE* layer of words; these appear as a substrate in IIr. In other words, they belong to the language of the BMAC area (e.g., place name such as *Xnəṇta, Baxδī?, Suɣδa?,* or etyma as *uštr, anc'u,* etc.) Further, agriculture in this area is not as early as she seems to assume: the BMAC to Bactria only at c. 2000 BCE, much too late to have influenced PIE. That Bactria/Sogdia could be the locus of PIE therefore is very doubtful, if not simply impossible. If her localization were correct, *all* IE languages should have received the same substrate words that make O. Iranian and O. Indo-Aryan so typical. (The same argument destroys the revisionist and 'autochthonous' Indian fanatsy, the so-called "Out of India theory" of PIE, see Witzel 2001a).

As the non-IIr language groups of IE do *not* have the loans discussed here the supposed locus of earliest (pre-)PIE must be at least to the north of the BMAC area, north of Sogdia, if not on the very boundary between steppe and woodland (taiga) in N. Kazakhstan, where also the oldest correspondences between PFU and PIE are located (PIE *\*wed+r/n-*, PFU *\*wete* 'water', etc. (cf. Nichols, 1997: 146).[95]

[92] "The locus of the IE spread was therefore somewhere in the vicinity of ancient Bactria-Sogdiana," Nichols 1979: 137, and "a spread beginning at the frontier of ancient near Eastern civilization" ... "in the vicinity of Bactria-Sogdiana ... included the ... urbanized oases of Southern Turkmenistan and Bactria-Sogdiana" (Nichols 1998: 233).

[93] Viz., north of the Black Sea. Certainly not in the clearly non-IIr. Mazenderan, as her map seems to indicate.

[94] Note also the early criticism, in part based on prepublication materials, by Mallory 1998.

[95] If not going back to common Nostratic, see Witzel 1992.

§ 3 |The Greater Bactria-Margiana Area


§ 3.1. Delineating the BMAC area

All of the data mentioned so far need to be studied in greater detail, especially the reconstruction of an early IIr. presence in Central Asia and on the Iranian plateau as seen in their linguistic, religious, social and material culture-related data (Witzel 1999a,b). For the BMAC area itself, we have no written sources at all, except for the loans quoted above (for further details see below).

First of all, it has to be established why one should think of Central Asian origins for the loans found both in Indo-Aryan and Iranian. In theory, such loans could also have originated in the Panjab and have traveled to Iran (as the words for 'rice' did, indeed); or, conversely, from an Iranian area to the Panjab (as is the case with 'wheat').

However, the words that can be reasonably well plotted both in time and place, that is, *uštr 'camel', *khar 'donkey', and *išt 'brick', point to the areas along the northern rims of Greater Iran (BMAC, for short). This is the first Central Asian area with a highly-developed agriculture and town civilization that the speakers of Indo-Iranian could have come into contact with, south of their original contact zone with the Uralic and Yenneseian (Ket) speaking peoples. We know that, in this civilization, the domesticated camel was used,[96] that it continued the large scale use of unburnt bricks, and that the donkey was introduced from the Near East at the time. These three *leitfossils* also provide a time frame: The speakers of IIr cannot have moved into this complex earlier than the introduction of donkeys.[97]

However, as will see seen, some of the IIr loan words have been taken over twice or thrice, independently of each other, into various IIr languages in different areas of Greater Iran, but they still look very similar to each other. This kind of difference is a clear indication of dialect variations in an underlying substrate language. A selection of the most typical loans tends to center on the Bactria-Margiana area (cf. Witzel 1995, 1999a,b, 2000), -- perhaps, for some words, with the inclusion of Sistan/Arachosia where the delineation of the boundary of the 'BMAC' language is unclear so far .

However, for other parts of the Greater Iran (Iran proper, Baluchistan, Afghanistan), the southern boundary of the "BMAC" language(s) is fairly clear. We can easily exclude the southern belt of Iran as well as Baluchistan. East of Akkadian and Sumerian in Mesopotamia, the isolated[98] language Elamite was spoken, not just in the Susiana (Khuzistan) and *Anšan* (Fars and

---

[96] At 2500 BCE there are camels (figurines) drawing vehicles, in S. Turkmenistan.

[97] They could, however, have learnt about camels from the Central Asian hunters, as represented by the Kelteminar culture.

[98] A genetic link between Elamite and Dravidian has not been established, see the extensive discussion in *Current Anthropology* (McAlpin et al., 1975).

surroundings) but also in Southern and Central Iran,[99] in the areas of Tepe Yahya (*Simaški*),[100] and Shahdad (*Tukriš*),[101] but apparently not further east than Bampur (*Marhaši*).[102]

According to F. Vallat (1985: 52) the language of Marhaši[103] differs from that of Simaški, and is only very partially Elamite related. Here and in Meluḫḫa (Baluchistan, Sindh)[104] the language changes, though Elamite cultural influence extends even to the borders of Sindh. However, eastern names are neither Elamian nor Dravidian.[105] This indicates that there was a language boundary, somewhere to the west of the present Iran-Pakistan border, probably in a southwards prolongation of the Iran-Afghanistan border. That the Meluhhan language was different altogether from Elamite or Sumerian is obvious: the Mesopotamians need a 'translator from Meluḫḫa' (Possehl 1996: no. 2), whose name is reported as *Šu-ilišu* (Parpola 1994: 132).[106]

Baluchistan is characterized by an overlap between the influences from the Elamite and Indus cultures. Baluchi sites such as Kulli are dated about 2000 BCE.[107] The area which is later on called (O.Pers.) *Maka* (Mesopot. *Makkan*) or Makrān, has its center at Marhaši (Bampur), an important place of exchange between the Indus areas east of it, the Elamian areas west of it, and the Arachosian site of Mundigak, and Shahr-i-Sokhta (Aratta, Sistan), Simaški, as well as Turkmenia north of it. Only during the last period of Mehrgarh (level VII) we find a cemetery with BMAC-derived items, and the rituals and metallurgy recall those of the BMAC as well. The spread of late BMAC influences into the Indus valley, Baluchistan, Susiana, etc., is now well documented.

While the southern belt of Greater Iran thus is excluded, he exact alignment of the "BMAC" language boundary in the more northern parts of Greater Iran, however, is not so clear. We have some notices about a series of peoples who entered Mesopotamia from the Zagros

---

[99] The insurrection of Nāram-Sīn (text from the end of 3rd mill. BCE) mentions the rulers allied against Akkade: "the man of Meluḫḫa, the man of Aratta, the king of Marhaši, ... (another country : gap in tablet), the king of all of Elam" (Vallat 1995: 53); note also he list by the Sargonic king Rimus who conquered Paraḫšum, Zahar, Elam, [Ba]sin? and Meluḫḫa (see Gelb 1997: 594).

[100] Perhaps one of the Elamite capitals.

[101] Later taken over by the *Tukriš*; Steinkeller (1982: 265), however, locates *Tukriš* in the Elburz mountains north of Tehran. The word *Tukriš* has been compared with *Tuɣrān, Tuɣrastān, Tokharoi*, etc. (Henning), see EWA I 651 s.v. *tugra*.

[102] For various Elamite sources see Vallat 1985. However, Steinkeller (1982: 255, map p. 265) tentatively locates it a little further west, "in the perimeter of Kerman and Eastern Fars." The original name of *Maḫarši* seems to have been *[m]paraḫši*, represented in Akkadian as *Paraḫšum* (Steinkeller 1982: 237-8). Cf. the modern *Maškai* and other rivers in Baluchistan beginning with *Maš-??*

[103] Steinkeller 1982, 1989 (cf. also Hiebert 1998: 147). Steinkeller connects the name of Marḫaši /Paraḫšum/Paraši with the name of the Persians; for this, however, note the *Parśu* reference in RV, BŚS (Witzel 1999c) and a possible conection with mod. Pashto, see below, n. 199. -- For further E. Iranian/Vedic correspondences see Witzel 1989, ch.10, and for some Saka-like characteristics of O.P. see n. 199.

[104] See Possehl 1996.

[105] Which speaks against Drav. as the language of the Indus Civilization, see further Witzel 1999a,b, 2001b for other possibilities.

[106] Further, the loan word links between Sumerian and Drav. are in need of investigation (see Blažek & Boisson 1992 and Blažek, forthc.). As for the connection between the two areas, a *direct* southern route, from Sumeria via Bampur to Meluhha (provided it was Dravidian speaking!) is *not* likely as it is blocked by a large block of Elamite (from Susa up to Tepe Yahya) and by the separate linguistic area of Simaški (Bampur). Following the suggestions made in Witzel 1999a,b, I suggest that that early Dravidian may have received its agricultural terminology via Aratta (Sistan) or even directly by maritime contacts.

[107] Possehl 1997, with a date of 2500-1900 BCE for the Kulli/Quetta phases.

mountains since end of the 3rd millennium. First came the Lullubi and Guti, since Narām Sīn of Akkade (2213-2176 BCE). On the Iranian highlands, the Simaški state of the 3rd mill. is replaced by the Tukriš state in 2nd mill., which is connected, in some texts, with the Guti.[108] The Hurrites (with a Caucasian language related to the later Urartian) appeared in the same epoch but more to the north; the Kassites entered under the successor of Hammurabi, Šamšu-Iluna, at c. 1740 BCE and soon took over Babylon for centuries. Both the Kassite (Balkan 1954) and Hurrite languages (Mitanni, around 1400 BCE) have some OIA linguistic elements.[109] However, we know much less about Media (the area around Tehran), Herat, Sistan and Arachosia.

While the language of the substrate studied below included the greater BMAC area, roughly from Anau to Balkh, we cannot firmly exclude, at this stage, the Media/Areia/Arachosia areas as belonging to that of the "BMAC" language. Some arguments, however, will be presented below(§ 3-5).[110]

Indeed, *anc'u 'Soma plant' (probably Ephedra, discussion below), which is pressed out to prepare the sacred drink of the Indo-Iranian peoples, points to the high mountains of Central Asia (incl. the Hindukush, Pamir, and the Himalayas, see Staal, forthc.), where according to both the Avesta and the Ṛgveda the best Soma grows. There are, indeed, some indications of non-IIr speakers in the high mountains of Afghanistan even at the time of the earliest IIr texts. Such 'foreigners' indeed still survive[111] in the Pamirs with Burushaski speakers and are visible in the substrate in Khowar (Witzel 1999a,b, 2000).

---

[108] Hammurabi inscription, see Henning 1978, 220.

[109] Needless to say, this fragmentary information will have to be compared with the materials presented here. At first sight, I have not detected correspondences.

[110] For Margiana etc., cf. also Erdosy 1998.

[111] Note that the central Afghan highlands have been occupied by Mongolians (Hazara) after the expansion of the Mongol empire -- wiping out any possible trace of older population; for some details on the Central Highlands see Witzel 2000; cf. the similar situation in the Tien Shan/Pamirs, (Turk.: on the Chin./Tajik border, Mt. *Muzh [Tagh Ata]* < Ved. *Mūja-*, Avest. *Muža*!) with the spread of Kirgiz.

Languages and Peoples of W.Central Asia in the 1st to 3rd mill. BCE

\*\*\*

## § 3.2 . An older Central Asian level: The Oxus/BMAC language

At this stage, therefore, it is more profitable to take a close look at the actual linguistic evidence preserved by the loans into the IIr languages. Such IIr words do not have Indo-European shape and etymologies (Witzel 1995, 1999a,b, Lubotsky, forthc.).[112]

---

[112] I cannot enter here a detailed discussion of Lubotsky's forthcoming appear. His new observations, based on the Indo-Iran. words listed in EWA, include:

(1) unusual IIr suffixes, including -ka, -pa, -sa which are directly attached to the root: *stuka/stūpa 'tuft', *pāpa 'evil', *kac'yapa 'tortoise', *jharm(i)ya 'firm structure, house'; (2) the unusual structure of trisyllabic words with long middle syllable (CəCəCə), such as *yavīyā 'irrigation canal', *kapāra 'vessel, dish', *piyūša 'biestings, *warāj'ha 'wild boar' (but note the early loan into FU: F-Volg. *oraśe, Finn. oras, Rédei 1986: 54!). He assumes that some of these words could have only been taken over *inside* the subcontinent and that the language north of the Hindukush and in the Panjab was the *same* as that in the BMAC. This is unlikely and does not account for loan words with the suffix -əša, -əc'a, -əsa that also occur with short medial vowel (of 'Para-Munda' type, Witzel 1999a,b). Proof for a CəCəCə type Panjab language could only come from some words typical for that climate, flora, fauna or culture. Indeed, there are at least two such words which may be indicative, *mayūra* 'peacock' and *śārdūla* 'tiger', however they have complicated etymologies (Witzel 1999a,b). Only a study of Lubotsky's new words from IIr, as found in the *various strata* of Vedic and Iranian texts, would allow to determine what is securely reconstructable for Indo-Iranian, or just for the subsequent levels. The *earliest* attestation of a particular 'foreign' word, such as in the RV or the older parts of

Shibboleths for time and place are, as mentioned, the local words *uštr 'camel', *khar 'donkey', *išt 'brick'.

The Bactrian camel was domesticated in Central Asia in the late 3rd mill. BCE and introduced in the BMAC area late in the 3rd Mill. 2000 BCE BCE (Meadow 1983, Masson 1992: 39 sq, 229, 233). It is also found on a few Indus copper plates. Its Mesopotamian designation, found in middle and new Akkad. udru "Bactrian camel", is a loan from Iran (EWA I 238, KEWA III 652, cf. Diakonoff in *JAOS* 105, 1985, 600). However, the IIr designation, of PIIr *uštra, represented by the identical Ved. uṣṭra, Avest. uštra 'camel', has no plausible IIr or IE etymology.

The case of the donkey is of similar nature, though the source of the word seems to be a more southern one. IIr *khara > Ved. khara, Avest. xara 'donkey' seem to go back to the earlier areas of its domestication, in case Mesopotamia, where an Akkadian (Mari) source has ḫārum, ajarum 'male donkey' (EWA I, 447, 473). Various types of donkeys may be distinguished in Vedic India,[113] in addition to the other wild equid, the hemione (onager, *Equus hemionus khur*). Interestingly Vedic, Tocharian and Dravidian have words that resemble the IIr and Akkadian form, Ved. garda-bha 'donkey' :: *karca-bha? > Toch.B kercapo[114] :: PDrav. *garda(?) > Tamil kaḻutai 'donkey'.[115] They all may go back to a Near Eastern form *xar- that is also reflected by IIr *khara. When and how Dravidian took over the word remains unclear so far (cf., however, Blažek, forthc.)

Finally, the word for 'brick' finds its likely source right in the center of these relationships, in the BMAC area. An IIr or IE etymology is unlikely (cf. EWA I 201) for Ved. iṣṭi, iṣṭakā, iṣṭikā, Avest. ištiia 'brick', zəmō-ištuua 'clay brick', O.Pers. išti, M.P., N.Pers. xišt. One also may compare Toch. iścem 'clay' (sometimes regarded as Iran. loan, see EWA) and, unnoticed so far, Bur. d-iṣ.c.ík 'brick.' All of this leads to a C.Asian source for '(clay) brick', *išt, with an unclear pre-Bur. "prefixed" d-.

Since there are no brick buildings at the northern rim of C. Asia[116] and as both the Indus as well as the Mesopotamian areas are too asymmetrically located to be the source of all these words, the most likely Central Asian source for all these loans is the BMAC area. Instead of some hypothetical Caucasian source, the BMAC area is the first place where the steppe people, the speakers of IIr., coming from somewhere north close to the Uralic and Ket, would first encounter bricks and would need a term for them. As discussed above, the domesticated camel was used

---

Avestan, is the *only* secure basis for analysis as words from the (unknown, complex) popular *local Indian* level constantly emerged in the sacred speech of High Vedic and the educated speech of the Brahmins (cf. Kuiper 1991, Witzel 1989), even at much later periods (cf. the IE *perd, attested only in post-Vedic pardati 'to fart'!) Such words could accidentally have the desired shape, CəCəCə. Note also that Lubotsky so far neglects the typical substrate words found *only* in Iranian such as kaufa 'hump, mountain' (see n. 157) which can serve as a corrective. -- Lubotsky's new data, however, support my initial analysis of Central Asian loans (Witzel 1999a,b,c) though he does not localize this substrate beyond pointing to the towns of the BMAC and the Panjab (echoed by G. van Driem *MT*, Oct. 1999; for a discussion and refutation see Witzel forthc.a).

[113] khara, rāsabha, gardabha, etc. see Rau 1980-81.

[114] Both Vedic and Toch., with the common Indian animal suffix -bha (as in garda-bha, śara-bha, ṛṣa-bha), which would point to a slightly different source.

[115] The overlap of Dravidian kaḻu- and Ved. garda-bha 'donkey' would be one of the few links of the Central Asian (BMAC?) language(s) with Dravidian (cf. McAlpin et al. 1975) and Indo-Aryan. IA words for domesticated animals are entirely different from Drav. However, Ved. garda-bha EWA I 473, Drav. kaḻu-tai DEDR 1364 could, or rather should be independent loans from a language of Greater Iran or Central Asia.

[116] For example in Sintastha there are only reinforced palisades with clay inside; however, the area just north of the Caucasus reportedly has clay bricks, see Witzel 2000.

here, and the donkey had been introduced from the Near East at the time. These three *leitfossils* provide the time frame.[117] It is confirmed by the forms of some words in still undivided IIr, which can be dated at c. 2000 BCE: the common IIr designations for the newly invented horse-drawn chariot (*ratha*, Ved. *ratha*, Ir. *raϑa,* Mitanni *Tuš-ratta,* Kassite *Abi-rattaš*) and related words (*rathin, rathī/raϑī* 'chariot driver' , and the archaic *ratheṣṭhā/raϑaēšta* 'chariot warrior').[118] All of this leads to a successful triangulation of material culture, time, and space in W. Central Asia at c. 2000 BCE or slightly later.

§ 3.3. Wheat agriculture.

To these, the plants cultivated in the area may be added. Wheat probably is the best case. It is a western import, as it originated west of the Zagros and south of the Caucasus[119] in the western Fertile Crescent (between Jericho and Jerf al Ahmar). In S. Asia it is found as early as the 7th millennium BCE, it is first attested linguistically by Ved. *godhūma*, Drav. (Kan.) *gōdi,* Bur. *guriṅ, gureṅ* (pl.), *γárum,* all derived from W.Asia.

Ved. *godhūma*, Avest. *gaṇtuma* must go back, with folk etymology on the Indian side (see below)[120] to a common IIr. < *gant-um,*[121] which in turn, echoes a Near Eastern source, *qend:* Semit. *ḫnṭ* (Arab. *ḫinṭ*^*atum*), Hitt. *kant*, Egypt. *xnd* (EWA II 499). On the other hand, a Caucasian/Bur./Basque (Macro-Caucasian) source is seen in: Bur. *gur* 'wheat'[122] ~ Basque *gari* 'wheat', Proto-East Caucasian (Daghestani, etc.) *Gōl'e*, PKartvelian (Georgian, etc.) *ghomu* (see Witzel 1999a,b,c). Obviously the ultimate Near Eastern source for all these words must have been something like **\*gər // ʕ/gən-d* (ə = uncertain vowel).[123]

The IIr source, *gant+um-a,* differs from its Near Eastern source **\*gənd-* by a suffix *-um* which is attested in Iranian Avest. *gaṇtuma,*[124] Vedic *go-dhūma*[125] but also in Burushaski *γond-um* (and perhaps in P.Kartvelian *ghomu*). One may assume a BMAC word *gant-um* that has entered Iran and India via the northern Iranian trade route (Media-Turkmenistan-

---

[117] Note that the speakers of IIr could not have entered or been in close contact with the BMAC area earlier than the introduction of donkeys (while they could have learned about camels from the Central Asian hunters, as represented by the Kelteminar culture).

[118] Such as 'chariot driver, chariot warrior (bow shooter), spoke, felloe', etc.; (pre-)chariots (*ratha,* not *anas* 'wagon') are first found in the northern steppes at Sintastha and also in Near East about 2000 BCE, see Witzel 2000.

[119] See the recent summary on Ofer Bar Yosef's work, at http://www.harvard-magazine.com/on-line/09016.html.

[120] The unfamiliar *gant-um/gand-um* > Iranian *gantum,* Indian *godum* (OIA and Drav.) was analyzed as IA *go-dhūma* 'cow smoke' (cf. also DEDR 2226 Konda etc. *goyi* 'smoke', thus *godhūma* an original joke form?); similarly, though differently, in Dravidian (see below) .

[121] For the 'suffix,' cf. Neo-Elam. *umi* 'to grind (grain)' and PDrav. *um* 'husk, chaff'; the compound 'Pan-Iranian' *gant + um* (note Berger's Bur. *γund-um*) may therefore have originally meant 'wheat grain'. (Cf. also the link between PDrav. *var* 'seed, grain', Elam. *bar* 'seed', Southworth, 1988: 659-660).

[122] Bur. *guriṅ, gureṅ* (pl.), *γárum < *γor-um < **γund-* (Berger), **γund-um*); cf. also Bur. *gur* 'barley, wheat colored', *bur* 'buck wheat,' Berger 1959: 43.

[123] See below, § 5, on the interchange of *r/n*.

[124] Avest. *gaṇtuma,* M.P., N.P. *gandum,* Shughni *žindam,* Pashto *γanəm < *gandūma*?, Khot. Saka *ganama < *gandama* (Berger 1959: 40f, EWA II 498); however, Brahui *xōlum < IA *γolum* (CDIAL 4287).

[125] As seen in Nur. *gūm,* Hindi *gohū/gehū/gahū,* etc. The northern form, based on Pre-Iranian *gantum* would have resulted in Vedic *gan-dhūma* or perhaps *gandha-dhūma* "perfume smell", cf. CDIAL 4020 Skt. (lex.) *gandhālu* 'fragrant rice', Pashai *gandár* 'a kind of grain'. The actual Ved. form *go-dhūma* must be due to local influence by the Southern (Meluhhan) *godi* (see above on Konda *goyi*) on a northern *gantum/gandum*; it may be due to Dravidian influence on the Panjab in the Middle/Late Ṛgvedic period as *godhūma* appears only in early post-RV texts.

Margiana/Bactria-Aratta/Sistan), while the forms with *go-* (PECauc. *\*Gōl'e*, PKartv. *\*ghomu*) that are reflected in Dravidian, must have come via the Southern route (Elam/Anšan - Simaški/Tepe Yahya - Marhaši/Bampūr), resulting in Drav. *\*gōdi* (Kan. *gōdi*, Tam. *kōti*, cf. DEDR 1906).[126] The post-RV change from *-an-* > *-o-* is untypical for the Panjab but common in Sindh (Witzel 1999a,b). It must have influenced, in some way (such as Drav. influence on the late RV, Witzel 1999a), the actual form of Ved. *go-dhūma* as well. Note that the Drav. word, too, seems to be based on a popular etymology.[127]

Interestingly, irrigation channels, the only means allowing of sustainable cultivation in these dry lands, are indicated by a loanword as well, *\*ya(u)vyā* 'streamlet, channel'.[128] Irrigation agriculture is said to exist from 2200 BCE in W. Central Asia.[129] The difference in sound between Vedic and O.Persian indicates a slightly different source of the loan, e.g., more eastern vs. a more western or Gorgan origin: Ved. *yavyā* 'stream, channel', but O.P. *yauviyā* > M.P., N.P. *jō, jōy* 'stream, channel', Parachi *žī* 'rivulet' (EWA II 405). Both forms cannot go back to exactly the same source; this kind of difference often is indicative of loan words.[130] In the sequel, a few more important words relating to material culture are listed briefly.

§ 3.4. Other agricultural terms.

Items from agriculture and settled life include words for 'sheaf', 'seed, semen', 'pillar', 'to heal, healer', 'lute', 'mark', 'lump', and even colors: 'blue', 'brown'. These items are indicative of the expanded agricultural base and the settlements of the BMAC.
• *\*parš* 'sheaf': Ved. *parṣa* 'sheaf (of corn)', Y.Avest. *parša* 'sheaf', EWA II 101;
• *\*bīj* 'seed, semen': Ved. *bīja* 'seed, semen', OIran. *\*bīza* (in names), Buddh. Sogd. *byz'k*, Parachi *bīz* 'grains';
• *\*sthūnā* 'pillar': Ved. *sthūna*, YAvest. *stūnā, stunā*, O.P. *stūnā* 'pillar'; unless ~ Ved. *sthūra* 'tall, thick', Avest. *-stura*, Khot. *stura* (thus EWA II 768);

---

[126] There is archaeological evidence now: While Renfrew had maintained the introduction of wheat by the speakers of the (hypothetical) Elamo-Dravidian, reaching India as early as 6000 BCE, Dorian Fuller (University College, London) now shows that wheat and barley arrived in southern India only at c. 2200 BCE, after having stalled for 3000 years in northwest India before farmers developed a monsoon-tolerant variety, see Science Magazine, Volume 294, 2 November 2001, page 989, http://www.sciencemag.org/cgi/content/full/294/5544/988.

[127] The unfamiliar (Iranian) *\*godum*, adopted into India, has been reconstructed for Drav. by Southworth (1988: 658, 660) for c. 1000 BCE as *\*kō-tumpai* 'low red plant' (perhaps because of DEDR 3334 Tam. *tumpai* etc. 'nettle, weed'?) The development from *\*tumpai > di*, however, is not clear. At the *supposed* date of 1000 BCE, *\*kōtumpai* could even be based on RV *godhūma!* This late date (along with its speculative, glottochronological basis), anyhow, is now doubtful as wheat is found in S. India already at 2200 BCE (see previous note). -- The various Elamian, Sumerian, etc. loans into Drav. will have to be compared, and must be reconciled with Blažek (forthc.)

[128] See Dani 1992: 116, 222 etc.; note, later on, also the extensive network of underground channels in Afghanistan, *karēz*, etc.

[129] Adams 1998: 376, quoting F. Hiebert (1996, cf. 1998: 231); cf. however, Francfort 1999: 451.

[130] Interestingly, the O.P. word is so close to Vedic; this suggests, like several other (grammatical) items, a previous closer relationship in N.(E.) Iran/Central Asia -- *yav(i)yā* may have been influenced by PIIr *\*naHwíya* "boatable" (from *\*naHu* 'boat'), > Avest. *(āfš) nāuuaiiā* 'water channel', Skt. *nāvyā* 'navigable river' (cf. Adams 1998: 373). Adams thinks that Tocharian borrowed key words for irrigation from E. Iranian sources (which he dates much too early, Francfort 1999: 451) though the Tocharians knew terms of primitive agriculture (plow, draft ox, harness, grain/wheat, harvest, to plow) before; cf. Shishlina and Hiebert 1998: 231.

• *pind 'lump': Ved. *piṇḍa* 'lump, ball (of food)' Khotan. *piṇḍaa*, Armen. *pind* 'compact, firm' < Iran;[131]
• *ling 'mark' : Ved. *liṅga* 'mark, penis' : Avest. *haptō-iriṇga* 'the seven marks' = the seven stars of the Great Bear/Wain (*ursa maior*), s. . EWA II 478sq;
• *vīnā 'lute' : *vīṇā* 'lute' (YV+ RVkh), Khot. *bīna* 'harp, lute', Sogd. *wyn'* 'lute', M.P. *win* 'lute', Armen. *vin* 'lute', -- unless these are loans from India, cf. EWA II 568;
• *kapaut 'blue': Ved. *kapota* 'pigeon', O.P. *kapauta* 'blue'; Khot. *kavūta* 'blue', M.P. *kabōd* 'grey-blue', *kabōtar* 'pigeon', s. EWA I 303, Kuiper 1991
• *kadru 'brown': Ved. *kadru* 'red-brown', *Kadrū* 'a snake deity', Avest. *kadruua.aspa* 'with brown horses, N.P. *kahar* 'light brown'.
• *bhiš, bhiš-aj' 'to heal, healer' : IIr *bhiš-aj > Ved. *bhiṣ-aj*, *bheṣaja* ~ Avest. *baēsaz-iia;* note the unusual noun formation, and the important combinations with *\*sauma* (Soma/Haoma), see EWA II 264.
• *sīs 'lead', EWA 734 from southwest Iran. (O.P.) *\*siça* 'white' which would be too early[132] for a loan into Old Persian as *sīsa* is found already in AV. Rather the word must be due to a local substrate, perhaps one found[133] in the Kandahar area.[134]

§ 3.5. Local Plants.

Importantly, some Central Asian plants are included in this list as well.
• *bhang 'hemp': Ved. *bhaṅga,* Iran. *banga* 'hemp, hashish', if not ~ *bhañj* 'to break', cf. Ir. *Hauma-varga* "Soma twisters", a Saka tribe; cf., however, Macrocauc. words such as Bur. *baṅ, baṅgí* 'hemp', Caucasian: Chechen *baga* 'pine tree', Lezgi *muk'-rag* 'fir tree, etc. < PEC *\*bħ·nḵ'wV* (Witzel 1999b, Bengtson 2001).
The following two cases, however, indicate a much deeper level and a wider spread, extending, via N. Iranian languages (such as Scythian), to the borders of Europe. These *wanderwörter*, words of culture such as 'coffee' or 'tea', have traveled the length and breadth of Eurasia.
• **kan-/k'an- 'hemp' : Ved. *śaṇa* 'hemp, cannabis', M.P. *šan* 'hemp', Khot. *kaṃha*, Osset. *gœn, gœnœ*, note : Sumer. *kunibu;* Gr. *kánnabis,* Russ. Church Sl. *konoplja,* OHG *hennep* < Scythian *\*kanap;* Kirgiz, etc. *kändir,* Turkish *kendir,* > Hung. *kender; Mordwin. kańc't, kańf;* Cheremis *kəńe;* -- for the change between *k'/ś* in E. Afghanistan/Panjab as in *Karkōṭa/Śarkoṭa* see Witzel 1999a, b;

---

[131] Note also *kubja, kubhra* 'crooked', and N.P. *kūž, kūz*, Sogd. *kwzz* 'humpbacked', Khot. *kūysa* 'bent', with unclear etymology; cf. Skt. *kuṇṭha* 'defective' CDIAL 3260, 3290 ~ Iran: N.P. *kund,* Bal. *kunt* and the many words in Turner, CDIAL meaning 'defective'. -- For Kubhā as a river name, see however, below n. 180.

[132] The Persians moved into the Persis and Anšan from NW Iran only after c. 700 BCE., see summary in Skjaervø 1995.

[133] Lead is found in the Kandahar area, see maps in Kenoyer 1995.

[134] Further a large number of verbs can be added, such as *\*kan* 'to find pleasure, please', *\*kram* 'to stride', *\*kroc'* 'to shout', *\*kc'ā* 'to regard, look', *\*kšad* 'to serve food', *\*kšam* ' to be patient', etc., or the culturally important *\*kaiš* 'to apportion magically, to teach', *\*mark/marc* 'to damage', *\*bhiš* 'to heal' with its strange derivative *\*bhiš-aj* 'healer', or nouns such as *\*kaufa* 'mountain, hump (of a camel)', see further, below n. 157.

• **sinšap 'mustard': Ved. *saṣarpa* 'mustard' (Brāhmaṇas) > MIA, NIA *sāsapa* 'mustard seed', Khot. *śśaśvāna*, Parth. *šyfš-d'n*, Sogd. *šywšp-δn*, M.P. *span-dān*, N.P. *sipan-dān* 'mustard seed' (cf. Gr. *sínapi*), < pre-Iran. *sinšapa* < **sinsap* (Henning's *s₁ens₂ap,* see KEWA s.v.)[135]

§ 3.6. Local animals.

Even more importantly, some names for local animals pinpoint the area of the 'BMAC' substrate more clearly. On would imagine that the Indo-Iranians had a word for the common tortoise, however, this is a loan which has been given to a river as well.

• *kac'yap 'tortoise': Ved. *kaśyapa* KS+, YAvest. *kasiiapa*, Sogd. *kyšph*, N.P. *kašaf, kaš(a)p* 'tortoise'; cf. *Kashaf* Rūd, river on the border of Turkmenistan and Khorasan (cf. Humbach 1984, 1991)

The leopard, a common Central Asian animal (often found in early C. Asian art, Dani 1992), later used for sport (see elsewhere in this volume), again has designations that differ considerably in the various language concerned. It is already seen as a prominent motif on the oldest pottery from the area, in W. Turkmenistan.

• **pard/pandh 'spotted animal, panther': Ved. *pṛdāku* 'snake' RV, *pṛdakū* AV 'panther snake' AV, *pṛdākhu* BŚS, W.Panj. *parṛā*, O.Iran. *pard-*, cf. Khowar *purdūm* < *pṛdhūma*?, Bur. (Yasin) *phúrdum* 'adder, snake', see EWA II 163, KEWA II 335, CDIAL 8362.

The overlap between 'panther' and '(spotted) snake' is due to the fact that snakes are frequently named after a number of characteristics of other animals and plants, in case, the spots of the panther (differently EWA II 163). Lubotsky (2nd Intl. Vedic Workshop, Kyoto 1999, in press) has pointed out that the AV snake *pṛdāku* is called after the leopard.[136]

The Iranian forms differ considerably: N.P. *palang 'leopard, felis pardus'* < O.Iran. *pard-*,[137] Kurd. *pilink*, Pashto *pṛang,* Parachi *parö:n,* Lahnda *parṛā*; cf. Gr. *párdalis, párdos, léo-pardos* 'leopard', all < *pard* 'spotted, wild animal?' (cf. EWA II 133), or following Henning, from an older **parϑ*. However, the alternate Greek word, *pánthēr*, must be taken into consideration as well. The original C. Asia word seems to have had the dialect variants **pard/pand*.[138]

Even more intriguing is the case of the word for the rhinoceros. It seems to go back to a local word *kart/kard with a 'suffix' *-ka/ga* (cf.Lubotzsky, forthc., on the common BMAC suffix *-ka* added directly to roots).

• **kart-ka 'rhinoceros': Ved. *khaḍga* 'rhinoceros' MS+; cf. N.P. *karka-dān*, Arab. *karkaddan*, Aelianus *kartázōnos* (*kargazōnos*) 'Indian rhinoceros'; however, (cf. Kuiper 1948: 136 sqq).

The rhinoceros is by definition a South Asian (Indus Civilization etc.), not a Central Asian animal. However, its prehistoric spread still needs to be investigated. The local climate and water

---

[135]Also found in SE Asia: Malay *sawi, səsawi*, Austro-As. *sapi, sV(r)-sapi*; cf. further the similar word, EWA 712, 727: *śiṃśápā* RV+ 'Dalbergia sissoo' N.P. *šīšam*, Pashto *šəwa* < *śīṣampā*, CDIAL 12424, Elam. *še-iš-šá-ba-ut = [šeššap].*

[136] Note that the use of * *pard* (*pṛdāku*) has been narrowed down in India to '[spotted] snake' while old word for 'panther' has been substituted early on by *dvīpin* 'the one having islands (spots)', AV+, see KEWA II 87, EWA II 769; probably *pṛdāku* has been influenced in form (and meaning: 'reptile') by the local Indian *sṛdāku/gu* MS, *sṛdāku* 'lizard', *sṛdara* 'snake'.

[137] For the sound changes involved cf. Avest. *paridaēza* > N.P. *pālēz* 'garden.'

[138] Cf. § 5, on the interchange of r/n; *pa^r/_n d* has been substituted in India early on by *dvīpin,* see above. -- Note that Steinkeller 1982: 253 considers the "speckled 'dog' of Meluḫḫa" (*ur gùn-a,* not: 'red dog' as often quoted!), a diplomatic gift from Marhaši, to be a leopard (Panthera pardus). One wonders, however, why one would send such a common animal as the leopard to Ibbī-Sin, the last king of the Third Dynasty of Ur.

table around the Sistan lakes and the reed thickets of the Helmand river,[139] would have allowed for its existence in the area, while the comparable situation in the Central Asian oases and their rivers, such as the Balkh river and the Amu Darya, must be studied more closely and the question rises whether this is a loan from S.Asia or a local word as well.[140] At any rate, based on this word alone, the southern and southeastern limits of the BMAC language cannot be decided. However, it offers a welcome possibility of checking the origin, track and spread of such loans.

§ 3.7. The religious sphere

Importantly, in addition to and beyond the items of material culture listed above, the religious sphere,[141] too, is strongly involved in the C. Asian loans into O. Iranian and OIA.

The most prominent words are those of certain rituals, deities and priests: *anc'u 'Soma plant', *yātu 'black magic', *atharwan 'priest' (however see EWA I 60), *ṛši 'seer', *uc'ig 'sacrificing priest', *c'arwa 'name of Rudra', *indra, *g(h)andharw/b(h)a 'demi-god or demon'.[142]

All these words are at the center of much of Vedic and also (pre-) Zoroastrian religion but have not been considered as being non-IIr (non-IE) so far. The advances made in the study of the C.Asian substrate now allow to place these items in perspective.

First of all, the older IE ritual stressed the ritual drink made from honey, Greek *ambrosia* (the drink of 'immortality' < IE *ṇ-mṛto-); it is called *madhu* 'mead, fermented honey' in India (see above, on 'honey'). Reminiscences of this drink remain in the poetry used for Vedic ritual, where mead has been substituted altogether by the new 'pressed out' (*su-*) drink, *Soma*, whose preparation and use developed into the most important Indo-Iranian ritual.[143] The Soma plant, whose botanical nature is still uncertain,[144] originally seems to have been called *anc'u* (Ved. *aṃśu*, Avest. *ạsu*, see G. Thompson[145]). As has been mentioned, the best variety grows, according

---

[139] See Falk 1997. with relevant literature.

[140] The tiger (along with the rhinoceros??), will have occurred in the swamps of Sistan. The last Central Asian tiger was reportedly shot on the Aral Lake in the Seventies (cf. Masson 1992: 39), and the last one in Afghanistan, on the Amu Darya islands near Kunduz, in 1970 (Bucherer-Dietschli 1986: 95). However, the tiger (*bebr*, Horn 1983: 42, cf. Ved. *vyāghra, EWA II 593*, with folk etymology < C. Asian **bagr?*) is still found in Iran, in the Elburz and Kopeh Dagh mountains. Its absence in the RV in contrast to its prominence in the seals of the Indus civilization (along with that of the leopard that has been found depicted from the early Turkmenistan pottery onwards), may be due to the fact that it not belong to the traditional imagery of the Indo-Aryans who preferred the lion (*siṃha*). A new study of early IIr animal designations is a desideratum. -- In view of these uncertainties, it remains of course entirely doubtful whether the languages of the BMAC and of Harappa were related at all as Lubotsky (forthc., see above, n. 112) will have it.

[141] First collected and discussed, as far as represented in EWA, by Lubotsky (forthc.).

[142] The last three words may, however, be late loans into Vīdēvdād, from OIA, in the late 1st. mill. BCE.

[143] Indo-Iranian *sauma,* Ved. *soma*, O.P. *hauma*, Avest. *haoma*, from the root *su* 'to press', thus *sauma* 'the pressed drink'; note also the name of the Saka *Haumavarga* 'the Soma twisting/pressing Sakas' in Central Asia, attested in the O.P. inscriptions.

[144] See Parpola 1995, Nyberg 1995 (= Ephedra), and the recent Leiden workshop on the problem; its results are forthcoming in EJVS 2002.

[145] Presentation at the 3rd Harvard Round Table on the Ethnogenesis of South and Central Asia, May 2001. -- Note that the word for Asafoetida (a plant producing resin) is similar in shape: IIr *anc'u* could go back to an older pre-IIr form *ank'u* which is found in Toch. as *aṅkwaṣ,* Chin. *yāngkuì,* (a resin of *Asafoetida sinica,* a low grade stimulant), see Pulleyblank 1962; however, Lubotsky 1998: 379 simply subsumes it under *Wanderwörter* of unknown etymology. Interestingly, Arrianos, Anabasis iii, ch. 28 is aware of it (*silphion*) as growing in the 'Caucasus' (Hindukush); this is identified as *Asafoetida,* liked by sheep; see Wirth 1985: 887, n. 111.

to Avestan and Vedic sources, on the high mountains (Tien Shan/ Pamir/Himalayas),[146] and that is exactly where the more potent variety of Ephedra is found. From there it is traded and brought in by the local (mountain) people.[147] There are indications, so far largely neglected, of a C. Asian ritual involving smaller and larger (Soma?) vessels, usually found turned upside down in archaeological sites.[148]

Not unexpectedly then, the names of some of the most important IIr. priestss and composers of ritual poetry also belong to this substrate: *atharwan '(fire?) priest,' *ṛši 'seer, poet', *uc'ig 'sacrificing priest', *yātu 'black magic'.[149] IE etymologies, however, not very convincing ones, have been proposed for some of these designations, especially in the light of the analysis given below (-arwa suffix).

If these points are evaluated against a discussion of the names of deities, it becomes obvious that a major change in ritual and religion took place among the immigrating speakers of Indo-Iranian in C. Asia. For, another feature of IIr religion is the emergence of a group of "gods of law and order" (asura, Āditya) that regulate the "truthful' behavior of people, from a single person to a tribe.[150] Other important gods that seem to be based on a substrate designation include *c'arwa 'name of Rudra', *indra, *g(h)andharw/b(h)a 'a demi-god or demon'.[151] These are, however, of some uncertain attribution, as they are, outside Vedic India, found only in a late Avestan text in (Vīdēvdād) and may represent loans from OIA, when Zoroastrian religion was confronted with Indian 'unbelievers'.

---

[146] See now Staal (forthc.) with a localization; the RV speak of the best variety, *Maujavata*, from the mountain *Mūjavant* (AV), c   f. Avestan *Muza,* and note even now the mountain (range) *Muzh* Tagh Ata, and the R. Muzkol, on the Xinjiang/ Tajikistan border, see Witzel 2000.

[147] See Witzel 2000 for a discussion of non-IIr Hindukush and Pamir mountain peoples.

[148] Sarianidi 1992: 34; this custom contined even down to the Greek period in Bactria (Ai Khanum).

[149] In detail: *atharwan, Ved. *Atharvan* '[ancient]priest, sorcerer, Avest. *āϑrauuan* 'priest', EWA I 60; -- *ṛši 'seer, poet', Ved. *ṛṣi*, cf. O.Avest. *ərəšiš* '(person) full of enthusiasm of god,' Y 31.5, *ərəšišiia*, EWA I 261; --*uc'ig 'sacrificing priest', Ved. *Uśij, Auśija,* Avest. *usij* 'sacrificer following non-Zorastrian ritual/belief', EWA I 234; -- *yātu 'black magic', Ved. *yātu*, Avest. *yātu* 'sorcery, sorcerer', N.P. *jādū* ' sorcerer'; -- note also *kaiš 'to apportion magically, to teach' (cf. EWA I 359).

[150] The Vedic/Avestan lists have: 1. Ved. *Varuṇa,* Mitanni *Uruna* (substituted by Avest. *Ahura Mazdā*?; with a local(?) -una suffix otherwise rarely found in IE, IIr (in: *Yamunā, taruṇa,* cf. Avest. *tauruna, auruna,* 2. Ved. *Mitra,* Mit. *Mitra,* Avest. *Miϑra,* 'god  Agreement', 3. *Arya-man/Aⁱriia-man* ('Arya-hood', god of marriage exchange, an artificial formation from an adjective with the deverbal derivative suffix -*man*, similarly Avest. *x̌aē-tu* 'self-hood' > 'belonging to a family'), 4. *Bhaga/baγa* (O.P. *baga*) 'god Share', 5. *Aṃśa* 'god Lot', 6. sqq. *Dhātṛ/dātar-* 'god Apportioner/Creator', *Dakṣa* 'god Cleverness', *Mārtāṇḍa Vivasvant/Gaiiō marətan, Vīuuaŋhuuaṇt.* These new deities are active in the social relationships, respectively in the world, tribe, clan (marriage), family, individual; *Vivasvant* is the ancestor of *Manu* and *Yama/Yima* and thus, of human beings. They are not found in IE (Slav. *boguᵘ* 'god' is a loan from N. Iranian *baga* 'god'). Their (already IIr.) origin may be due to socio-religious developements in steppe conditions (water rights, oral agreements, etc.) or based on interaction with the BMAC khanate style settlements (?, and subsequent backchannel influence on the North Iranians??). However, the BMAC population had an altogether different religion, see Francfort.

[151] Details: *c'arwa 'name of a god', Ved. *Śarva* 'name of Rudra', Iran. *Sauruua* 'a demon', *indra 'name of the 'king' of the present generation of the gods', Ved. *'Indra'*, Mitanni *Indara,* Avest. *Indara,* 'a demon', *g(h)andharw/b(h)a, Ved. *Gandharva* 'name of a semi-divine spirit of lust and procreation', Avest. *gaṇdarəß a* 'a demon, monster'.

Importantly, Lubotsky (forthc.) draws attention to the common 'suffix' *-arwa-[152] in *atharwan, g(h)andharw/b(h)a, and perhaps also in *c'arwa.[153] Interestingly, this is seen only in religious terms, which at a minimum points to intentional, if not artificial formations involving these words.[154]

However that may be, the surprising cluster of words related to religion indicates that the Indo-Iranians were strongly influenced by a Central Asian population having its own peculiar religion; this whole problem is in need of a close evaluation in comparison with Vedic and Avestan religion.[155]

The remaining group of words only attested in Indian sources[156] and words only attested in Iranian sources[157] will be evaluated in another paper.


§ 4. Place Names and Hydronomy

An extensive survey of the older place names of Turkmenistan, S. Uzbekistan, Afghanistan, etc. is still to be done. The oldest sources are found in the Avesta and in the Old Persian inscriptions. Both include only a few Central Asian names, usually of Iranian origin.[158] The Greeks, starting with Herodotos[159] and the historians[160] of Alexander's trek towards India, provide a number of interesting toponyms and personal names, again mostly of Iranian nature.

---

[152] Lubotsky thinks that this group of words belongs to a *still older* layer of loans in IIr.

[153] Doubtful at least for *c'arwa,* unless one assumes a root *c-, as can happen in Caucasian languages; cf. for NEC river names, n. 90.

[154] For details see above n. 150, such as *Arya-man.* Is the IIr. suffix -*man/van* replacing a BMAC suffix -*arwa* whose use would have resulted in the difficult form Ved. *arya-rva/aryārva? --Cf., finally, the rare -*vala* suffixes in Vedic (RVC *a-kṛṣi-vala*)?

[155] For an initial investigation, see Witzel forthc. (Third Intl. Vedic Workshop at Leiden, May 2002).

[156] One may also investigate, for example, the names of clearly non-IA gods and demons in the RV, only a few of which fit the trisyllabic pattern with long middle syllable (see above, n. 112): *Araru, Arbuda, Ilībiśa, Uraṇa, Emuṣa, Karañja, Kaulitara, Khela(?), Cumuri, Jarūtha, Dṛbhīka, Namuci, Paṇi, Parṇaya, Pipru, Bṛsaya, Raji, Vaṅgṛda, Śambara, Śāṇḍa, Śimida, Sṛbinda.* Unfortunately these words do not have counterparts in Old Iranian and (many or most) may come from South Asian substrates; for details see EWA, *s.v.*

[157] Such as O.P. *kaufa* 'hump of a camel, mountain'; Avest. *aδu* 'channel, rivulet', *aϑā* 'land, property', *aoniia* 'fire place', *akana* 'receptacle', *aku* 'scissors', *ayažāna* 'an agricultural instrument', *tutūk* 'clay', -*bata* 'threshed', *muštəməša* 'myrrh', etc., or the many names of illnesses and other words describing 'evils' of all sorts (cf. the many local Indian ones listed in CDIAL !) such as *aka/aɣa, axti, afša, afsman, afšman, aŋra* (= Ved. *asra*), *ara, -yaska, -skanda, -stairiia*, or those of insects and others pests such as *anāiriti, araēka, aširiia*, and finally personal and place names such as *Axtiia, Apaxšīrā, Ankasa, Ainiiāuua, Amru*, as well as those of demons such as *Arəzura, Aiiēhiiā, Ara, Asabana*, etc.

[158] See above § 2.3: Choresmia, Sogdia, Bactria, Margiana, -- all already Iranian names, or mythical rivers such as the *Vaŋhī*/Wehrūd; see Vamberi 1891 etc., Witzel 2000.

[159] Histories, i 205: the Saka tribe of the *Massagetai*, their queen *Tomuris*, her son *Spargapisēs*; iii 92 the tribal names *Kaspioi, Pausikai, Pantimathoi (*on the Oxos), *Dareitai,*
 *Baktrianoi, Aigiloi* (on the Iaxartes); iii 117 the *Khorasmioi, Hurkanioi* (in Gorgan), *Parthoi, Sarangai* (= O.P. *Zranka*), *Thamanaioi*; vii 64 sqq: *Baktrioi, Sakai, Skuthai, Khorasmioi, Sogdoi, Kaspioi.*

[160] For example, from Arrian's Anabasis, iii 29 the town of *Baktra*, west of Mt. *Aornos*, the tribe of the *Da[h]ai* from other side of the R. *Tanais* (Iaxartes); the R. *Oxos*, the town of *Drapsaka* (Kunduz), Mt. *Aornos*; iv: the towns *Zariaspa, Kurupolis, Gaza* (in Sogdia), R. *Tanais*, the town *Marakanda*, R. *Polutimetos* (Zarafshan), R. *Epardos* in *Marda* land (near Merw?); *Marakanda, Zariaspa, Ga[za]bai* (near Bukhara), *Nautaka, Paraitakēnē, Khoriēnos, Baktra*; cf. also Curtius Rufus, Historia Alexandri Magni (De gestis Alexandri Magni).

Ptolemy' Geography[161] adds a host of names, especially for the areas on and beyond the Yaxartes, in the Saka territories and in Xinjiang.[162]

There is an increasing number of mythical and real names found in Middle Iranian Zoriastrian and Manichaen texts, the old Sogdian letters or documents from Mt. Mug with more than 60 toponyms. A brief overview of such early, medieval, and more recent data up to the 19th century is provided in a suvey article by Khromov (1980, 132-139). All of these later data cannot be dealt with in this short paper.

Modern data for these area as well as for other former Soviet Central Asian republics exist in a number of publications in Russian that can be compared with the older data. In a survey article Khasanov (1960: 156-160) discusses the multi-lingual names of rivers and towns. He underlines that the hydronomy is often taken from the names of local provinces or towns, a feature already seen (often vice versa) in the Avesta and O.Persian, such as Balx river/town < *Bāxδī*, Merw river/town < *Margu*, etc. Taken from his map (1969: 159), the Syr Darya also goes by these names (note Russian |x| often for [h], *as in Šah*): (Gr.) *Tanais, Turkestan Daryasi, Otrad/Otrar Daryasi, Karalun Okuz, Seixun, Naxraš Šax, Benaret Daryasi, Xojent Sui, Kasart Aksu Suvi, Fargana Daryasi, Uzgan Suvi* (Gr. *Silis* is missing); similarly, the Amu Darya: (Lat.) *Oxus, Urgenc Daryasi, Xorezm Daryasi, Oks, Okus, Araks, Jeixun, Kelif Daryasi, Balx Daryasi*, and the Zerafšan: *Buxara Daryasi, Xaramkam, Rudizar, Obe Sogd, Obe Kuxak, Samarkand Daryasi, Zaravšan*. Out of these, obviously only the (parts of the) old designations not stemming from IIr (*darya, ob*) and Turkic (*si, su*) are of interest here.

A useful overview of the literature and of some major toponymical features of the area has been given by Murzaev (1964: 3-13), Postelov (1980, 118-123), Khromov (1980, 132-139), E`del'man (1980, 21-32). Other areas included in these studies written in Russian are the Upper Altai, Kirgizstan, Kazakhstan, Tajikistan (Khromov 1959, Rozenfel'd 1964), Tajik Badakhshan (Rozenfel'd 1953, 1980), Afghan Badakhshan (Gryunberg 1980), the Pamirs,[163] Iran,[164] (and Turkic toponymy in general as also Mongolian, etc. hydronymy).[165]

However, these publications (Murzaev et al., 1960, 1980, Bandaruk et al., 1964) usually deal with the formations visible on the surface, that is with "suffixes" and "prefixes" of geographical designations such as "X-river, Y-meadow, upper-Z", etc. Useful as they are, they hardly deal with the pre-Iranian toponymy.

As far as the vast stretches between the Caspian and Bactria are concerned there are, however, some intriguing names that are worth discussing. J. Bengtson[166] has recently shown beyond doubt that Northern Caucasian and Burushaski belong to the Macro-Caucasian language

---

[161] See Ronca 1968/1971 and Humbach 1972, 1998.

[162] Such as the names on the eastern bank of the Yaxartes: the Saka tribes of the *Karatai, Komaraoi*, and east of them, the *Grinaoi, Toornai, Bultai* and *Massagetai*; further east, beyond the *Askatankas* and *Imaon* Mts.: *Kasia* (cf. Avest. *bərəzaṇt- [aṇtarə-]Kaŋha- < *kasa,* with the pass *Xšaϑrō.suka*), the *Khatai, Akhasa*, the *Khauranaoi*; and finally, in Xinjiang (from north to south): the *Piadai, Asmiraia*, the *Oikhardai, Issēdones, Throanoi, Thagouroi, Aspakarai, Batai* and *Ottarokorai* (= *Uttarakuru*, "beyond the Himalayas" of AB); the rivers *Oichardēs, Bautisos;* the place names *Damna, Issedōn, Piada, Asmiraia, Khaurana, Orosana, Ottarokora, Solana, Thogara, Daksata, Sēra* (*mētropolis*). Some of them clearly are Iranian (*Aspakara*), others are *prima facie* unclear.

[163] Dodykhudoev 1975, E`del'man 1975.

[164] Savina 1964 , 1980; Rozenfel'd 1964.

[165] See in: Murzaev et al., 1980, 72-88

[166] Bengtson 1992, 1999 and at 3rd Harvard Round Table on the Ethnogenesis of Central and South Asia, May 2001.

family.[167] This vastly distributed group of languages, stretching from the Basque country to the Pamirs is now left with just a few remnant populations.[168] One may expect therefore, that the area between the Caucasus and the Pamirs contain some place names that reflect the languages of peoples later replaced by Indo-Iranian and Turkic.

Indeed, there is the wide-spread river names *Sind-. It is attested north of the Caucasus, as the *Sindes*, people on the Kuban river.[169] The *Sindes* also is a river -- with persisting, pre-/non-Iranian s-, that is mentioned by Tacitus (Annales X.10); it divides the *Dahae* from the *Arii*, and thus refers to the Merw (Murghab) or Tedzhen river. Note further one of the major Central Asian rivers, the *Sir Daryā,* which still was called *Silis* by the Scythians in Plinius the Elder's time (Natural History VI. 49). Burushaski seems to preserve the etymon with *sinda, sindmiṅ* 'river'.[170] Finally, the major river of the Burusho area and beyond is the Indus, first met with in the RV *Sindhu* which clearly is the river that begins in the Himalayas and receives the Kabul, Kurram, Gomal as well as all the Panjab rivers (RV 10.75). P. Thieme (1967-91) understood the Indus as the 'border river' dividing IA and Iran. tribes and has derived it from IE with an etymology from the root *si(n)dh* 'to divide.'[171] The word is also found in O.Iranian as Avest. *həṇdu* 'border river, ocean (at the western and eastern ends of the world'[172], O.Pers. *ha̱ndu* [*həndu*] 'Indus'. This points to an IIr coinage with the meaning 'border river, ocean'.

The question remains why this reconstructable PIIr word is s similar to Macro-Caucasian words that reflect something like *(t)sin(d)/ *(t)sir;[173] (see above on NEC *ɬ$^T$enc'o- 'river' and NEC *š$^T$or- 'water'.)

---

[167] While Bengtson (1992, 1999, 2001) reconstructs a Basque-N.Caucasian-Burushaski (Macro-Caucasian) language family, this has also been extended to Sinitic and Na-Dene (in Alaska, Yukon, etc., and incl. Apache and Navajo).

[168] One may regard it as the remnant of the early (Mesolithic) settlement of Europe and C. Asia by Homo Sapiens sapiens. K. Tuite (1998), indeed, deliberates on an emigration of the speakers of Proto-Burushaski from their NE Caucasian home (or nearby), discusses a number of links between the folklore of the Caucasus and the Pamirs, and lists some grammatical similarities and words that would suppose such a link. He is, however, cautious whether this evidence proves a genetic link or just loan-word relationship (1998: 461, 467). -- Cf., the different opinion of J. Nichols 1997: 128 (and cf. 1998: 226 for Nakh-Daghestanian in the eastern Caucasus foothills); she assumes a proto-Kartvelian homeland SE of the Caspian. If that were correct, Kartv. would, improbably, just be the Cauc. remnant of the "BMAC language."

[169] Mayrhofer 1979.-- Note L.G. Gulieva 1960. The older name was *Sag-dan* (Ptolemy), from Skythian/Ossetic *don*, Ir. *dānu*, see Dul'zon 1964: 16.

[170] Bur. *sinda*, Yasin/Werchikwar dial. *sénde* (< NIA Shina *sin*?), see Pinnow 1953: 12-13; cf. also Bur. *tsil/ts.hīl, ts.hilmiṅ* 'water', Werk./Yasin *tshel*; cf. NE Cauc. *ɬ$^T$enc'o- 'river' (see n. 90); -- Harmatta (1990: 376-377) compares the wide-spread river name *Sind-* with Burushaski *sinda*.

[171] Some compare Irish *Shannon*, etc., see KEWA, EWA II 729; these names are discussed in some detail by Mayrhofer (1979); he denies any connection with Skt. *Sindhu.*

[172] Cf. *us.həndauua* 'beyond the natural frontier' (Thieme), that is: the (mountain) 'rising from the ocean (Milky Way)' / or 'beyond the Milky Way', *zraiiah vourukaša*, Witzel 1984).

[173] For details on the distribution of *r/n* see § 5. The word is attested in Mesopotamian sources as *si-in-da-a*, var. *si-in-du*), see n. 177.

In the case of IA one may also think of an adaptation of Bur. *sinda,*[174] from Proto-Burušaski,[175] to the well known IIr and IE *sidh*/*sindh*- 'to divide'. Pinnow, indeed, connected IIr *Sindhu* with Burušaski as he could not find a cogent IE etymology and as he rejected Near Eastern ones (Pinnow, 1953: 12-13.)[176] He also points to a number of river names in the northwestern subcontinent which have the same suffix -*u*, but are clearly IA (1954: 14 sqq). It must also be noted that the word *sinda* is attested fairly early in the context of Mesopotamian trade.[177]

However the wide spread of these words east of the Caucasus rather seems to indicate a Macro-Caucasian source *sind-/sir*- that has been connected, *already in P-IIr times*, with IIr *si(n)dh* 'to divide' and etymologized as *Sindhu* 'border river.'[178]

If the connections made above are correct we get a band of river names, from the Caucasus to the Pamirs, that reflect the language of a previous C. Asian population.

This impression is reinforced by a few other names. The people north of the Caucasus, the *Sindes*, live on a river now called Kuban that emerges from the Caucasus mountains of the Cherkes republic. This name, attested at least since 7th cent. CE,[179] seems to echo[180] that of the

---

[174] Cf. Witzel 1999a,b,c.

[175] Cf. Tikkanen 1987, 1988, 1998. Early loans include Ved. *kilāla* / Class. Skt. *kīlāṭa* 'a milk product' (RV), Bur. *kīlāy* 'curds'; Ved. *meṣa* 'sheep', Bur. *mēṣ* 'skin bag' (cf. Slav. *mex$^u$*, Lith. *maišas* 'skin bag'); cf. also Bur. *baluqá* 'big hammer' (see above, § 2.2; further *śon* 'blind', Ved. *kāṇa* with the same northwestern variation of *k/ś* as seen in *Kar-koṭa/Śar-koṭa* (Witzel 1999a,b); cf. also a cultural loan from the Indus: *ɣupas* 'cotton', Ved. *kārpāsa;* all of this is in need of further investion. -- For pre-Bur. substrates see the summary in Kuiper 1962, 1991.

[176] Cf. also Mayrhofer 1979, on the *Sindes*, a people on the Kuban R., north of the Caucasus mountains and the *Sindes* river (Tedzhen, in Tacitus, *Annales* 10*;* Witzel 1999c, § 5 and n.139-140; cf. however, n. 170.

[177] Cf. Neo-Babylonian sources for Indian products, as detailed by B. Landsberger, *Welt des Orients* 3, 261 (cf. OIP 2, 123: 35), e.g., pillars of ivory, cypress, cedar, *duprānu*, juniper and 'Indian wood' (*si-in-da-a*, var. *si-in-du*), apparently derived from a predecessor of Bur. *sinda*/OIA *sindhu*. -- Note however, also the alternative explanation by Southworth (1988: 659): P.Drav. *kīntu > *cīntu [cīndu, sīndu] 'date palm, phoenix sylvestris' > Munda *sindi, kindad, kita* etc. (Sant. *kinde't,* Mundari *kitā*). Southworth sees here the origin of the name of Sindh, however without explaining -*dh*- in Ved. *Sindhu* (s. EWA II 729, Witzel 1998, 1999).

[178] Note some early loans from Pre-Vedic IA into Iranian when O. Iranian still had *s* (later > *h*, as in Assyrian *As-sa-ra ma-za-aš = Assara Mazaš, Ahuramazdā,* see Hintze 1998). Ir. *Hindu*- thus can be a loan from an older IA substrate. However, the 'eastern and western' *həṇdu,* 'oceans' Y. 57.29, and the name of the mythical central mountain, *us.həṇdauua* 'emerging from the river/ocean [*Vourukaṣa*]' indicate *həṇdu* '(mythical) ocean' in Avestan (Witzel 1984), derived from an IIr (loan) coinage 'border river, ocean' which would fit P. Thieme's etymology (1967-91) from the IE root *si(n)dh* 'to divide.'

[179] L.G. Gulieva, 1960, 140 quotes *Kum* in the Armenian geography of Moses Xorenac'i; *Kofina* in the *Brevarium,* the history written by the Constaninopolitan patriarch Nikephoros (c. 750-829 CE, when describing the time of the emperor Maurikios, -602 CE); *Kuba* in the 17th cent. Russian records, *Bol'shoi Chertezh*. -- Gulieva 1960: 135 includes other names: (Gr.*) Hupanis* (she writes Gr. Gipauis, Gipanis)*, Antikites, Bardan, Bardanus, Bardanis, Psalmis, Kuba, Kopa, Sopa, Kufus, Ukrug, Psyzh', (Psishe), Burlik, Bal, Kum.*

[180] It seems to come from a Caucasian language. Gulieva (1960: 136) points to the Karachevo-Balkar word *koban* "(demoniacally) furious, racing (like a horse)," and she compares (incorrectly) the Greek 'Gipanis' (= *Hupanis*), which she connects with horses as well. The word rather should be derived from an Iranian *Hu-pāni*- 'having good drinking water(?)', but note that *pā* 'to drink' is little attested in Iranian (O.Avest. *vīspō.paiti*, N.P. *nabīd*), and OIA *pānīya* 'water' is attested only later, in the Epic; however, *pāni* has become the common word for 'water' in NIA. -- Other forms are *Kuman, Kuban'* (in Turkic languages). Gulieva (1960: 138) also compares the river name *Kopa* 'lake.' The Indian *Kubhā* (RV, Gr. *Kophēn, Kophēs,* 'Kabul River') can only vaguely be connected with such words as Ved.

Kabul river, RV *Kubhā*, Gr. *Kophēs, Kophēn*." Another interesting river name is that of the Indra River in S. Tajikistan, *Indar-āb*,[181] and the *Inder* lake (Russ. *ozero Inder*) on the lower Ural river in W. Kazakhstan. In light of the proposed non-IIr etymology[182] of the name of the god *Indra* (see above) these widespread names may reflect the C. Asian substrate language as well.[183] Much more research is needed, however, to turn these proposals into something closer to certainty.

As far a Greater Iran itself is concerned, the O.Iran. sources (O. Persian inscriptions and the Avestan texts)[184] add a number of interesting names. Beginning in the Northwest, the non-IIr name of the Gorgān plains south of the Caspian Sea is found in the Avesta in V.1.9 as *Xnənta* < *\*khnanta* which Humbach (1991) derives from < *\*khr-* (cf. Ptolemy, Geogr. 6.9.5 *Khrindoi*, Lat. *Chrindi*). This, incidentally, is another case where we see a shift between *r/n* in Central Asian etyma and names (see below, § 5). The IIr name of the area is *Vəhrkāna* > mod. *Gorgān*, "the (country) of the wolf (men)." In the northwest of Iran the O.Pers. inscriptions mention the districts *Asagartiya* (Witzel 1980: 112 n. 76 with lit. on place names), *Kampanda* and *Nisāya*, and the fortress *Sikaya(h)uvati* (cf. O.P. *ϑika* 'gravel', Ved. *sikatā*), some of which have been explained by Eilers (1982,1987).

Further east, the name of Bactria may have a non-IE etymology as well. The E. Ir. name *Bāxϑī-* (wiztel 1980), may be derive from O.Ir. *\*Bāxϑ(r)ī-* < IIr. *\*bhāk-trī-* which may mean the 'distributing (river);'[185] however, the loss of *-r-* remains enigmatic (but cf. O.P. *Skud(r)a*, below). Y.Avest. *Bāxδī* corresponds to the AV loan *balhika* and would indicate a Y.Avest. dialect form *Bāxδī-* already at the time of the AV, c. 1000 BCE (Witzel 1980, 1997). If the local form of the name *Bāxδī* is original, and the Iranian forms in *-tri/ϑri-* were to be regarded as popular etymologies, it may reflect a local name.

Indeed, the name of a neighboring tribe and area, that of the Sogdians, is without clear etymology as well. Y.Avest. *Suγδa*, O.P. *Suguda*, which Szemerényi 1980, however, derives from O.P. *Skud(r)a* 'archer', the older name of the Saka.

To the South of this area, in Arachosia, the O.Persian sources have the names of two fortresses, *Aršādā*[186] and *Kāpišakāni* and a *Gandutava* district. *Kāpišakāni* is close to *Kāpiśī*, a town in the Kabul area (Gandhāra, see Pāṇini 4.2.99; Patañjali further specifies *Kāpiśāyana/-ī* 'inhabitant of *Kāpiśī*'.[187] In the same area may be found the Vedic *Kamboja* tribe, reflected by the O.Pers. royal name *Kambūjiya* 'Cambyses'.[188] Finally, there are the SE Iranian provinces of

---

*kubja* 'bent, crooked', *kubhra* 'humped bull' (KEWA I 232; EWA I 368, CDIAL 3300, 3261; Pinnow 1959: 340 §483, Kharia: *kubja* 'crooked') and may rather be be explained, as the *Kuban* above, from a related Cauc. word.

[181] The N. Afghani river *Andar-āb* has to be separated from this as it is attested in mid-first mill. Chin. sources with *an-*, see Witzel 2000, n. 7.

[182] The IE etymology from PIE *\*ind* 'to swell may be reflected by the French river name *Indré* in central France, cf. however the next note.

[183] Taking onto account the Macro-Caucasian family, the French river name *Indré* (south of Tours) may reflect a Basque substrate.

[184] There is urgent need of a detailed study of the many name of mountains in Yt 19; for some initial attempts see Witzel 2000.

[185] Cf. also the *Vāxəδrikā* mountains in Yt 19.4, see Witzel 1980, 2000.

[186] One may try, however, the IIr. etymology, such as *arša(n)-d(h)ā* 'placing/containing men.'

[187] Cf. *Kapisa* in Ptolemy, Geogr. 6.18.4; see Witzel 1980: 108: n. 49, Eilers 1982: 21; on Kandahar < Gandhāra, see Eilers 1982: 22.

[188] Cf. Witzel 1980, n. 81, also n. 16, 32, 47, 52, 82, 96, 105, 106, 108. Note also the import of horses from Kamboja, Bactria, Sauvīra (Sindh) in Arthaśāstra 2.30.9.

*Karmāna* and *Maka* (modern Makrān, Gedrosia) whose inhabitants are called *Maciya* 'person from *Maka.*'[189] These may reflect the old names of the area (see above, on *Marḫaši*/Bampur).[190] (The Southwest of Iran is proper Elamite territory and therefore left out here).[191]



The major languages and language families of W. Central Asia

§ 5. Some Characteristics of the BMAC language

In spite of the (so far) limited number of etyma, anthroponyms and toponyms we can summarize a few characteristics of this language. For example, just as in the case of the rather limited dialect materials available for the northern and southern Indus languages (Witzel 1999a,b), we can discern a dialect difference involving an interchange of *r/n*. Naturally, we are very much in the realm of speculation here, as the available data are still very sketchy and from a variety of quite different languages. It also still is somewhat difficult to pin them down in time and place. Nevertheles, a few observations are added here in the hope of drawing attention to the facts themselves.

(1) *r/n*. Assuming the existence of the Macrocaucasion language family, the terms for wheat have the structure **gVr/l (Basque *gari*, PECauc. *\*G̲ōl'e,* Bur. *gur* 'wheat'; other languages of the Near East have **kant,* and in Iran/India the "suffix" *-um* (Ir. *gantuma*) or they have changed the initial vowel to *-o-* (PKart *\*ghomu*?, Ved. *godhūma*, Drav. *\*kotumpai*?) In sum, the more northern population (Macro-Caucasian?) which derived its word for wheat directly from the Fertile Crescent, has a predilection for *r/l* while other, more southern languages from the area have *-n-*.

---

[189] Witzel 1980: 112 n. 76 and Eilers 1982: 30 with lit. on place names.

[190] Incidentally, the Greek name of the capital of this area, *Poura,* does not seem to be connected with IA *pura* 'city', cf. modern *Bampūr*; see Mayrhofer 1979.

[191] In the SW, the towns *Kuganakā, Tāravā, Māru, Raxā* and the mountains *Arakadi* and *Parga; Pātiš(h)uvari* (Gr. *Pateiskhorēis*), a Persian tribe Eilers 1987: 49, "gegen die Sonne gerichtet" in Bartholomae's *Wörterbuch.*

The same distribution is reflected by the word for 'water, river': *Macro-Cauc. \*(t)sir*, NEC *\*šᵀor-* 'water' which may be preserved in the name of the Iaxartes, *Silis*, Bur. *ts.hil* etc., (cf. Basque *(h)ur*; not from Yenesseian *\*xur*).

The loanword for 'leopard' has a similar distribution. The southern belt has -*r*- forms while the west has -*n*-: \*\**pard* 'spotted wild animal?' > O.Iran. *\*pard-* 'leopard' (N.P. *palang*, etc.), Lahnda *parṛā*; taken into Gr. as *párdalis, párdos, léo-pardos* 'leopard'; however another loan (from Anatolia?) is seen in Greek *pánthēr*.

Similarly, the word for 'lion' has -*r*- in the southern belt, represented by Iran. *šer (< \*serg'h??)*, while a host of 'non-southern' languages from the intrusive IA (originally close to FU north of the steppes), Tibetan, Chinese etc., have a form with -*n*-, \*\**sɪeṇgha*, pre-OIA *sing'ha*, etc. Note the possible western form in Arm. *inc/inj < \*sinj'h*?

With soem reservations one may add the name of the province of Gorgan (Hyrcania, *Vəhrkāna*) in N.Iran, Avest. *Xnəṇta, next to an original \*Xrəṇta* (Humbach 1984, 1991).The name certainly is non-IIr. and as the Avestan texts were redacted in Arachosia, this would represent a split between the Macro-Caucasian areas close to the Caucasus, in Turkemenistan-Bactria and the Pamirs versus a southeastern O. Iranian dialect, close to the Vedic area with -*n*, (see above on -*śer/\*sing'ha)*.

The case of the *Sindes/Sindhu* river is more complex. NEC *\*ɬᵀenc'o-* 'river' agrees with the rivers *Sind-* in the Caucasian Kuban, TurkmenianTedzhen and the Indus areas, where it is fairly early, as Mesop. *sinda* 'wood from Meluḫḫa' (Baluchistan/Sindh) attests.[192] A Macro-Caucasian word \*\**(t)sin(d-)* was taken early over into Indo-Iranian and given a popular etymology from IE/IIr *\*sidh/sindh* 'to divide', as border rivers indeed do; consequently IIr *\*sindh-u* designates (border) rivers and the rim of the world, the ocean surrounding the world (Avest. *həṇdu* 'western and eastern ocean', *us.həndauua* '(mountain) rising from the ocean'). It was also applied, in S.Asia, to the Indus river (cf. Bur. *sende*), whence Iran. *Hindu/Həṇdu/Haṇdu*. Again, the word belongs to the Macro-Caucasian belt and was taken over by IIr on first contact.

In sum, The Macro-Caucasian belt, from the Pamirs to the Caucasus (and further west), has a predilection for forms with -*r*- where others have -*n*-.[193]

(2) <u>k : g</u> The Avesta may also supply another feature, an interchange of  *k : g*  as in *\*kaiśa > keśa* (caesar?)  :: Avest. *gaēsa*; cf. above, western (Hitt., Semitic) *kant/ḫant* 'wheat' :: local "Iranian" *gant-uma* (Kartv. *\*ghomu*?);  and further the unexplainable difference in an old loan word: Cauc. *b'ok'o* 'male goat' :: pre-Proto-Iran. *\*būg'a > \*būj'a >* Avest. *būza*  :: Indic *\*bōkka* 'he- goat' (CDIAL 9312. also *bukka* in lex.) > Prakṛt *bokkaḍa* 'male goat', Panjabi *bokkā*, Nep. *boko*, etc. (cf. Nichols 1997,1998 referring to Gamkrelidze-Ivanov 1994: 501).

(3) Perhaps, <u>kh : ka</u> may be added; note Ved . *khā*, Avest. *xan* 'source, well' : Ved. *khan* 'to dig', *khara* 'dug out mound' :: Iran. *\*kan* 'to dig' (Avest. *us.kəṇti, ni-kaiṇti, auua-kaṇta*, O.P. *ni-kaⁿtuv, kaⁿtanaiy* etc.), cf. EWA I 446.

(4) <u>t : dh</u>
is seen early on in IE *med(h)-* (Skt. *madhu*, Engl. *mead*, etc.) :: western IE \*\**melit*, Gr. *mélit-*, Hitt. *milit,* Lat. *mel, mell-,* Gothic *miliϑ*.

(5) <u>c : j</u> ?

---

[192] If indeed an old designation for the Indus, this would present an interesting aspect as far as the language of the Indus civilization is concerned. Cf. in general Witzel 1999a,b.

[193] Note also the more widely spread apophony of *r/n* (and *r/i*, named after W. Caland) in Indo-European of the type nom. *\*wedō-r*, gen. *\*wete-n-es* (cf. Witzel 1992).

may perhaps be deduced from Armen. *inc, inj* 'lion' < \*\**sinj'* ~ pre-Ved. \**sinj'h* < \**sing'h* :: Toch.
A *śiśäk*, B *śecake* 'lion'.
(6) <u>syllable structure</u> CəCɜCə,
as discovered for the Central Asian substrate by A. Lubotsky (forthc., see above, n.112).

In sum, the still expandable list of the names, etyma, phonetic and grammatical details,
discussed so far,[194] allows to posit a so far unrecognized, pervasive substrate in Old Indo-Iranian
that has certain phonetic peculiarities and that does not go back to the Sumerian, Akkadian,
Hurrite, Elamite or Indus languages. It is to be located in northernmost Iran/Afghanistan and in
Western Central Asia, in other words, in the Greater BMAC area and its surroundings.
This is the immediate contact zone for steppe populations coming from the north.[195]

§ 6. Transhumance, Trickling in, Immigration of Steppe Peoples

There is no need to underline that the establishment of a BMAC belt of substrates has
grave implication for the theory of an immigration of speakers of Indo-Iranian languages into
Greater Iran and into the Panjab.
The body of words taken over into the Indo-Iranian languages in this area closes the gap
between the northern boundaries of the IIr speaking territories (Uralic, Yenesseian) on the one
hand, and the firmly attested location of the IIr languages in Greater Iran (Avestan, O.Persian,
Median) and in the (northwestern) Indian subcontinent (Ṛgveda, later Vedic) on the other. In
form of a brief summary, this scenario would include:

• Immigration of speakers of late common Proto-Indo-Iranian or of early Old Indo-Aryan / Old
Iranian into the BMAC area and into Areia (Herat) and Sistan. These languages must have come
from areas further north, as the *old* loans into P.Uralic and P.Yenesseian clearly indicate.

• Amalgamation of Central Asian words along with their underlying concepts (\**bhiš, kapauta*),
religion (\**sauma, -rwa*), animals (\**uštra, \*khara*) and plants (\**bhanga, \*anc'u*) into the pre-Vedic
and pre-OIr. languages.

• This new, late IIr speaking entity moves ---*Kulturkugel* fashion (Mallory)-- into Iran and
Panjab, perhaps represented by the sudden expansion of the BMAC into Susa, Baluchistan and
the Indus area. This cannot (yet) be substantiated by any linguistic data.[196] Another scenario

---

[194] A comprehensive survey, especially of non-Persian place names has not yet been carried out, and a list of Iranian
substrate words in Old (not to speak of Middle) Iranian has not been drawn up either. As underlined above, this kind
of research has simply been neglected so far.

[195] It should be noted that the late BMAC shows an intrusion of steppe pottery (Hiebert, 1998, Shishlina and Hiebert
1998).

[196] Only some initial guesses are possible, for example about the ethnic nature of the *Tukriš* (see above n. 101) which
might be connected with Ved. *tugra, tugrya* (both personal names), Iran. *tuɣr-*. If true, we would have RV (and later
Vedic, BŚS) links with Bolan, Aratta, Shahdad -- recalling the more northern trail that lead the Mitanni-Indo-
Aryans westward into N. Mesopotamia.

could be that of successive spearheading forerays of mountain peoples, such as as the Guti, Lullubi, Kassites,[197] who were as yet only marginally influenced by IIr. languages and customs.

• Later, actual spread of speakers of pre-Vedic IA, that is Mitanni-IA, into N. Iraq/Syria (c.1400 BCE) and of the linguistically slightly later (still pre-Iron age) Rgvedic into Arachosia (*Sarasvatī > Harax$^v$a$^l$tī), Swat (Suvāstu) and Panjab at c.1200/1000 BCE (depending on the local date of the introduction of iron). Probably still later, the spread into Greater Iran of the earliest Iranians (1000 BCE : Hintze 1998, cf. K. Hoffmann (1976-92 [= 1941]) for pre-Ir. names in the RV).

• This move may have been independent of and actually been preceded by that of the speakers of group of IIr languages, now called Nuristani (or Kafiri) who have preserved some archaic features until today (Nur. .c. is older than RV ś or Avestan s, all from IIr c'). It may also have included that of the speakers of the non-IIr, western-IE group now represented by the substrate of speakers Bangani (in the high Himalayas east of Kashmir); however, these may just as well have come, as potential IE, non-Tocharian neighbors of the "western-IE" Tocharians, *across the mountains* from Xinjiang).

It might be added that the general path of immigration of the speakers of Indo-Aryan from the north, via the general BMAC area, is also indicated by an early loan from Nuristani, that is concentrated in the hills of NE Afghanistan (and, maybe, originally also the adjoining plains). This is Nur. *kat's'a > Ved. kāca (KS+) 'shining piece of jewelry' (K. Hoffmann 1976, *EWA* I 33),[198] taken over into O.P. as kāsa-ka 'semi-precious stone.'

• All of this is followed by the introduction of E. Iranian (Avestan) into E. Iran (1200/1000 BCE --note overlap with AV Balhika, Witzel 1980)

• and a still later one of the Median and Persian into W. Iran, by c. 900-700 BCE.[199]

---

[197] Only a few Kassite words seem to come from IIr., e.g. *Šuriiaš* 'sun god', *Maruttaš, Bugaš?;* see Balkan 1954, for horse names such as *akriyaš* '(running) in front?', *timiraš* 'black?', etc.; note the direct loan from IIr. with Nominative -s, as seen in some *old* FU loans as well (above, or cf. Finnish *kuningas* 'king' < P.Germanic *kuningaz,* as in Dutch *koning*).

[198] However, this may also be a post-Rgvedic loan from these isolated mountain languages, the archaic third branch of the Indo-Iranians (Morgenstierne 1973) that has survived in the mountains of northeast Afghanistan and in neighboring Chitral (Pakistan). -- Note *kāsaka* 'semi-precious stone', *kāsaka kapauta* 'lapis lazuli' and *si$^n$kabru* 'carnelian' described as brought from Sogdia, and *kāsaka axšaina* 'from Choresmia' (DSf 37-40). One would expect Bactria/Badakhshan.

[199] It remains to be investigated whether the Persians (*Pārsa* < *pārc'va-) are related to the Parśu (< *parc'u) of the Vedic texts (RV, BŚS), where they are located next to the Arattas (*āraṭṭa, arāṭṭa),* thus in Afghanistan. These are likely to be the ancestors of the Pashto (*paṣtō* < *-rš/*xšt- < *parštu/parštawā or (improbably) < *paxšt-; or cf. Avest. *paršta* 'back' thus, '*the hill people'; see Morgenstierne 1927: 61; Pashto has often been compared with Herodotus' *Paktues* which however cannot reflect expected -ršt-, only -xšt-, at the time). -- Notably, whether *parc'va is connected with Pashto or not, Old Persian -s- (as in < *asa* 'horse') < *śś < śv < c'v < IE k'w shares the development of IIr c'v > śś with Saka -śś-, while the rest of Iranian has -sp- (*aspa*) and Vedic has -śv- (*aśva*). This feature and others (cf. further grammatical features in Witzel 1989, ch. 10) may point to an ultimately northeastern (Bactrian?) rather than a northwestern (Urartu/Median) origin of O.P., and thus to a track of immigration from the NE via Media to the Persis, somewhat like Nichols' (1997-98) 'southern trajectory'. A northeastern origin would be close to the location of the Ved. *Parśu.*

In sum, it can now be stated more securely that speakers of an IE language, early OIA (pre-Ṛgvedic) entered the Greater Panjab, acquired local words from the Northern Indus dialect (such as *śaṇa, lāṅgala, vrīhi, godhūma, kaṅgu, Gandhāra,* Witzel 1999a,b). About the same time, the speakers of Dravidians entered Sindh, acquired the related words from the southern Indus dialect (*gōnu, ñāñcil, variñci, godī, kaṅku/kampu*) and perhaps they even brought horses to Pirak (1700 BCE), that is, not the IA(?) Bhalānas.


The whole process of 'Aryanization' may be summed up in the words of Polomé (1990: 337), who discusses the introduction of Indo-European into Northern Europe supplanting the local language, but not without leaving many substrate words (and ideas) with the emerging Proto-Germanic speaking peoples:

> "whichever way .... [the area] was indo-europeanized, the new population initially constituted a mere *adstratum* or *superstratum* to the long-established set of peoples. When and why the language shift took place remains a widely open question, but one thing is certain : it did not take place without leaving clear traces of the prior language(s) in the lexicon."

To which we may add: 'and customs, beliefs, rituals, religion, and material culture.'

ABBREVIATIONS

| | |
|---|---|
| Akkad. | Akkadian |
| Armen. | Armenian |
| Austro-As. | Austro-Asiatic |
| AV | Atharvaveda Saṃhitā |
| Avest. | Avestan |
| Brah. | Brahui |
| Bur. | Burushaski |
| Cauc. | Caucasian |
| CDIAL | Turner 1966-69 |
| DEDR | Burrow, T. and M.B. Emeneau 1984 |
| Drav. | Dravidian |
| EJVS | Electronic Journal of Vedic Studies |
| Elam. | Elamite |
| ep. | Epic Sanskrit |
| EWA | Mayrhofer 1956-76 |
| FU | Finno-Ugrian |
| F-Volg. | Finnish-Volgaic |
| Gr. | Greek |
| Hitt. | Hittite |
| IA | Indo-Aryan |
| IE | Indo-European |
| IIJ | Indo-Iranian Journal |
| IIr | Indo-Iranian |
| Indo-Ar. | Indo-Aryan |
| Ir(an). | Iranian |
| JAOS | Journal of the American Oriental Society |
| Kartv. | Kartvelian |
| Kaśm. | Kashmiri |
| KEWA | Mayrhofer 1986-96 |
| Khot. | Khotanese Saka |
| KZ | Zeitschrift für Vergleichende Sprachwissenschaft |
| Lith. | Lithuanian |
| M | Middle- |
| Mbh. | Mahābhārata |
| MIA | Middle Indo-Aryan |
| Mong. | Mongolian |
| M.P. | Middle Persian |
| MS | Maitrāyaṇi Saṃhitā |
| MT | Mother Tongue |
| N | New- |
| NEC | Northeast Caucasian |
| N.P. | New Persian |
| NIA | New Indo-Aryan |
| Nur. | Nuristani (Kafiri) |
| NWC | Northwest Caucasian |
| O | Old- |
| O.Avest | Old Avestan |

| | |
|---|---|
| OIA | Old Indo-Aryan |
| O.P. | Old Persian |
| Osset. | Ossetic |
| P | Proto- |
| PEC | Proto-East Caucasian |
| Pkt. | Prakrit |
| PNC | Proto-North Caucasian |
| PPerm. | Proto-Permian |
| RV | Ṛgveda Saṃhitā |
| Skt. | Sanskrit |
| Sum(er). | Sumerian |
| StII | Studien zur Indologie und Iranistik |
| Toch. | Tocharian |
| V. | Vīdēvdād |
| Ved. | Vedic |
| Y. | Yasna |
| Y.Avest. | Young Avestan |
| Yt. | Yašt |
| Yen. | Yenesseian |
| YV | Yajurveda (-Saṃhitā) |
| ZDMG | Zeitschrift der deutschen morgenländischen Gesellschaft |

# BIBLIOGRAPHY

Adams, D.Q. On the History and Significance of Some Tocharian B Agricultural terms. In: Mair 1998 : 372-378

Antilla, R. *Historical and Comparative Linguistics.* Amsterdam/Philadelphia: John Benjamins 1989

Bagchi, P. C. (ed.), *Pre-Aryan and Pre-Dravidian in Sanskrit.* Calcutta : University of Calcutta 1929

Balkan, K. *Kassitenstudien I: Die Sprache der Kassiten.* New Haven: American Oriental Society 1954.

Bengtson, J. D. Macro-Caucasian phonology. In: Shevoroshkin, V. (ed.) *Nostratic, Dene-Caucasian, Austric and Amerind.* Bochum: Brockmeyer 1992: 342-352
---, A Comparison of Basque and (North) Caucasian Basic Vocabulary. *MT* V, 1999, 41-57
---, Genetic and Cultural Linguistic Links between Burushaski and the Caucasian Languages and Basque. Paper at the 3rd Harvard Round Table on the Ethnogenesis of Central and South Asia, May 2001

Berger, H. Die Burušaski-Lehnwörter in der Zigeunersprache. *IIJ* 3, 1959, 17- 43

Blažek, V. The new Dravidian - Afroasiatic lexical parallels. forthc.
---, and C. Boisson, The Diffusion of Agricultural Terms from Mesopotamia. *Archív Orientalní* 60, 1992, 16-37

Blench R. and M. Spriggs, *Archaeology and Language I. Theoretical and methodological orientations.* London/New York : Routledge 1997
---, *Archaeology and Language II. Correlating archaeological and linguistic hypotheses.* London/New York : Routledge 1998
---, *Archaeology and language IV. Language change and cultural transformation.* London/New York : Routledge 1999

Bondaruk, G.P, E`.M. Murzaev, V.A. Nikonov, V.V. Cybul'ski (eds.). *Toponimika Vostoka. Novye Issledovaniya.* Moskva: Nauka 1964

Brereton, J. P. 1981. *The R̥gvedic Ādityas.* New Haven.

Bucherer-Dietschli, P. & C. Jentsch. *Afghanistan. Ländermonographie, bearbeitet von der Arbeitsgemeinschaft Afghanistan.* Liestal : Stiftung Bibliotheca Afghanica. 1986.

Burrow, T. and M. B. Emeneau, *A Dravidian Etymological Dictionary.* Second Edition. Oxford: Clarendon Press 1984        (DEDR)

Colarusso, J. Languages of the Dead.  In V. Mair 1998: 431- 447

Clauson, G. The Foreign Elements in early Turkish. *Researches in Altaic Languages*, L. Ligety (ed.) Budapest: Akadémiai Kiado 1975, p. 44

Dani, A. H. and V. M. Masson (eds.), *History of civilisations of Central Asia*, Vol. I. *The dawn of civilisation: earliest    times to 700 BCE.* Paris: Unesco Publishing 1992

Dodykhudoev, R. Kh. *Pamirskaia mikrotoponimiia : issledovanie i materialy.* Dushanbe : Irfon 1975.

Dodykhudoeva, L. R. The problem of East-Iranian loan words in Tajik language, (conf. paper) 2000

Dul'zon, A.P. Ket toponyms of Western Siberia. *Uchenye Zapiski, Tomskii Gosudarstvennyi Pedagogicheskii Universitet* 18, Tomsk 1959: 91-111
---, Drevnie toponimi yuzhnoi Sibiri indoevropeiskogo proiskhodzheniya. In: Bondaruk et al., 1964, 14-17
---, Ketskie imena i prozvishcha. In: Murzaev et al., 1980: 5-9

E`del'man, D. I. *Osnovnye voprosy lingvisticheskoi geografii. Na materiale indoiranskikh yazykov.* Moskva: Nauka 1968
---, K substratnomu neslediju central'noaziatskogo jazykovogo sojuza. *Voprosy jazykoznanija*, 5, 1980, 21-32
---, Geograficheskie nazvaniya Pamira. *Strany i narody Vostoka.* Vyp. 16. Pamir. M. 1975

Eichler, E., G. Hilty, H. Löffler, H. Stegler, L. Zgusta (eds.). *Namenforschung. Name Studies. Les nomes propres. Ein internationales Handbuch zur Onomastik. An International Handbook of Ononastics. Manuel international d'onomastique.* Vol. 1. Berlin: de Gruyter 1995

Eilers, W. *Geographische Namensgebung in und um Iran. Ein Überblick in Beispielen.* Bayerische Akademie der Wissenschaften, Phil.-Hist. Klasse Sitzungsberichte, Jg. 1982, Heft 5. München 1982.
---, *Iranische Ortsnamenstudien.* Sitzungsberichte der Österreichischen Akademie der Wissenschaften, Phil.-Hist. Klasse 465. Wien 1987.

Erdosy, G. (ed.). *The Indo-Aryans of Ancient South Asia.* (Indian Philology and South Asian Studies, A. Wezler and      M. Witzel, eds., vol. 1). Berlin/New York : de Gruyter 1995
---, Language, ethnicity and migration in protohistoric Margiana. In: Gubaev, A., G.Kosholenko, M.Tosi. *The archaeological map of the Murghab delta. Preliminary Report 1990-1995.* Rome, IsAIO 1998: 141-147

Falk, H. The purpose of Ṛgvedic ritual. In: M. Witzel (ed.) *Inside the texts, Beyond the Texts. New Approaches to the Study of the Vedas.* Cambridge 1997, 67-88.

Francfort, H.-P. *Fouilles de Shortughai : recherches sur l'Asie centrale protohistorique.,* par Henri-Paul Francfort avec des contributions de Ch. Boisset [et al.].

[Memoires de la Mission archeologique francaise en Asie centrale] Paris : Mission archeologique francaise en Asie centrale : de Boccard 1989.

---, *Nomades et sedentaires en Asie centrale : apports de l'archeologie et de l'ethnologie : actes du colloque franco-sovietique, Alma Ata (Kazakhstan), 17-26 octobre 1987, textes reunis par H.-P. Francfort.* Paris : Editions du Centre national de la recherche scientifique, diffusion: Presses du CNRS 1990.

--- et E. M. Kuz'mina. Du nouveau dans la chronologie de l'Asie Centrale du Chalcolithique à l'Age du Fer. *Rapport du Groupe de travail 'Asie'. Actes du colloque "C14 Archéologie",* Paris 1998: 467-469

---, Compte rendu: V. H. Mair, 1991. In: *BEFEO* 86, 1999, 449-458


Gelb, I.J. Makkan and Meluhha in Early Mesopotamian Sources, *Revue d'Assyriologie* 64, 1970, 1-7 = F.R.Allchin and D. Chakrabarti, *A source-book of Indian Arcaheology*, Vol. II Settlements, Technology and Trade. Delhi: Munshiram 1997, 59- 596


Gryunberg, A.L. Zametki po toponimii afghanskogo Badakhshana. In: Murzaev 1980, 165- 169


Gulieva, L.G. O nazvaniyakh reki Kuban. In: Murzaev et al., 1960, 135-139


Hamp, E. Whose were the Tocharians? In: V. Mair 1998: 307-346


Harris, D. R. (ed.). *The Origins and Spread of Agriculture and Pastoralism in Eurasia.* London: UCL Press 1996, 370-389


---, and C. Gosden, 1996, The beginnings of agriculture in western Central Asia. In: Harris 1996: 370-389.


Harmatta, J. The emergence of the Indo-Iranians: The Indo-Iranian languages. In: Dani, A.H. and V. M. Masson, *History of civilizations of Central Asia. Volume I. The dawn of civilization: earliest times to 700 B.C.* Paris: Unesco Publishing 1992, 357-378


Henning, W. The first Indo-Europeans in history. In: Ulmen. G.L. (ed.) *Society and History. Essays in Honour of Karl August Wittvogel.* The Hague -Paris -New York: Mouton 1978: 215-230


Hiebert, F. T. South Asia from a Central Asian perspective. In: G. Erdosy 1995, 192-212
---, Central Asians on the Iranian Plateau: A Model for Indo-Iranian Expansionism, in V. Mair 1998, 148-161
---, The recently discovered Bronze Age inscription (2300 BC) from Anau, Central Asia. Paper at: *Third Harvard Round Table on the Ethnogenesis of South and Central Asia*, May 12-14, 2001 ( http://www.fas.harvard.edu/~sanskrit/RoundTableSchedule.html)
---, and C.C. Lamberg-Karlovsky Central Asia and the Indo-Iranian borderlands. *Iran* 1992, 1-15


Hinüber, O.v. Indische Namen in Zentralasien bis 1000 n.Chr. In: Eichler, E. et al. 1995: 657-665


Hintze, A. The Migrations of the Indo-Aryans and the Iranian Sound-Change *s > h.* W. Meid (ed.) *Akten der Fachtagung der Indogermanischen Gesellschaft in Innsbruck* 1996. Innsbruck 1998

Hock, H.H. Out of India? The linguistic evidence. In: J. Bronkhorst & M. Deshpande, *Aryan and Non-Aryan in South Asia. Evidence, Interpretation and Ideology.* Harvard Oriental Series. Opera Minora, vol. 3. Cambridge 1999, 1-18

Hoffmann, K. *Aufsätze zur Indoiranistik.* (ed. S. Glauch, R. Plath, S. Ziegler, vol. 3). Wiesbaden 1992

Horn, P. *Grundriss der neupersischen Etymologie.* Strassburg: Trübner 1893

Huld, M. E. The linguistic typology of the old European substrata in North Central Europe, *JIES* 18, 1990, 389-411.

Humbach, H. Historisch-geographische Notizen zum 6. Buch der Geographie des Ptolemaios. *Jahrbuch des Römisch-Germanischen Zentralmuseums*, 19, Mainz 1972, 89-98.
----, A Western Approach to Zarathustra. *Journal of the K. R. Cama Oriental Institute* 51, 1984, 15-31
---, The Gāthās of Zarathustra and the Other Old Avestan Texts. by H. Humbach in collaboration with J. Elfenbein and P. O. Skjærvø. Part I. Heidelberg 1991
---, and Susanne Ziegler. *Ptolemy, geography, book 6 : Middle East, Central and North Asia, China.* . Wiesbaden : Reichert, 1998-

Illich-Svitych, V. M. *Opyt sravneniya nostraticheskikh yazykov.* II, Moskva: Nauka 1976

Janhunen, J. The Horse in East Asia: Reviewing the Linguistic Evidence.
In: Mair 1998: 415-430

Joki, A.J. *Uralier und Indogermanen. Die älteren Berührungen zwischen den uralischen und indogermanischen Sprachen.* Helsinki 1973

Katz, H. *Studien zu den älteren indoiranischen Lehnwörtern in den uralischen Sprachen.* Habilschrift München 1985
---, *Fremd und Eigen : Untersuchungen zu Grammatik und Wortschatz des Uralischen und des Indogermanischen : in memoriam Hartmut Katz* Wien : Praesens, 2001.

Kenoyer, J. Mark. Interaction systems, specialised crafts and culture change: The Indus Valley Tradition and the Indo-Gangetic Tradition in South Asia. G. Erdosy (ed.) *The Indo-Aryans of Ancient South Asia.* Berlin/New York : de Gruyter 1995, p. 213-257

Khasanov, Kh. Kh. Istoriko-toponimicheskaya skhema Srednei Azii. In : Murzaev et al., 1960, 156-160

A. L. Khromov, Izucheniye geograficheskikh nazbanii Tadzhikistana v SSSR i za rubezhom. *Izvestiya otdeleniya obshchestvennykh nauk AN Tadzhikskoi SSR.* No. 2 (36) Duzhanbe 1959
---, Sogdiiskaya toponimiya verkhov'ev Zeravshana. In: Murzaev et al., 1960, 87-98;
---, Nekotorye osobennosti formirovaniya toponimii Srednei Azii v IX-XIII vv. In: Murzaev 1980: 132-139

Klein, J. On Verbal Accntuation in the Rigveda. In: Witzel 1997: 139-166

Krause, Wolfgang. Zur Frage nach dem nichtindogermanischen Substrat des Tocharischen. *KZ* 69, 1948, 187-203

Kuiper, F.B. J. *Proto-Munda words in Sanskrit.* Amsterdam: Noord-Hollandsche Uitgevers Maatschappij 1948
---, *Aryans in the Rigveda.* Amsterdam-Atlanta 1991.

Landsberger, B. Einige unerkannt gebliebene oder verkannte Nomina des Akkadischen. *Die Welt des Orients* 3, 1954, 246-267

Lin, Meicun. The earliest Tocharian Loan-words in Ancient Chinese. In Mair, 1998: 476-482

Linder, Th. Griechische (incl. mykenische) Ortsnamen. In: Eichler et a1. 1995: 690- 705

Lubotsky, A. Tocharian Loan Words in Old Chinese: Chariots, Chariot Gear, and Town Building. In: Mair 1998: 379-390
--, Indo-Iranian Substratum. forthc. in: *Proceedings of the Conference "Contacts between Uralic and Indo-European in the Neolithic, Eneolitic and Bronze Ages in the Light of Linguistics and Archaeological Evidence" at Tvarminne, 8-10 January 1999*, forthc.

McAlpin, D. W. *Proto-Elamian-Dravidian: the evidence and its implications.* Transactions of the American Philosophical Society, 71, Philadelphia 1981
---, et al. Comment on David W. McAlpin's `Elamite and Dravidian: Further
Evidence of Relationship. McAlpin, David W., Elamite and Dravidian: Further evidence of relationship. (With discussion by M.B. Emeneau, W.H. Jacobsen, F.B.J. Kuiper, H.H.Paper, E. Reiner, R. Stopa, F. Vallat, R.W. Wescott, and a reply by McAlpin). *Current Anthropology* 16, 1975, 105-115

Mair V. (ed.) *The Bronze Age And Early Iron Age Peoples of Eastern Central Asia,* ed. by Victor Mair. 2 vols. = *JIES* Monograph 26, Washington/Philadephia: The Institute for the Study of Man / The University of Pennsylvania Museum Publications 1998
---, Notes on the Anau Inscription. *Sino-Platonic papers,* 112, July 2001

Mallory, J.P. A European Perspective on Indo-Europeans in Asia. In: Mair 1998: 175-201

Masson, V. M. The environment. In: Dani 1992, 29-44

Mayrhofer, M. *Kurzgefasstes etymologisches Wörterbuch des Altindischen.* Heidelberg 1956-1976. (KEWA)
---, Etymologisches *Wörterbuch des Altindoarischen.* Heidelberg 1986-96 (EWA)
---, *Ausgewählte kleine Schriften* (Sigrid Deger-Jalkotzy u. Rüdiger Schmitt, eds.): Wiesbaden : Reichert 1979

Meadow, R. *The equids of Mehrgarh, Sibri and Pirak: An Osteological Evidence for the Introduction of the Horse to South Asia.* Paper presented at the seventh Conference of South Asian Archaeology, Brussels, 4-8 July 1983.

---,  The origins and spread of agriculture and pastoralism in northwestern South Asia. In: Harris 1996: 390-412.

---,  *The Review of Archaeology. The Transition to Agriculture in the Old World*, 19, 1998, (Special Issue ed. by Ofer Bar-Yosef), 12-21

Misra, S.S. *The Aryan problem, a linguistic approach*. New Delhi : Munshiram Manoharlal, 1992
---, *The date of the Rigveda and the Aryan migration : fresh linguistic evidence*. Pune : University of Pune 1999.

Morgenstierne, G. *An Etymological Vocabulary of Pashto*. Oslo: J. Dybwad 1927
---, *Irano-Dardica*. Wiesbaden 1973

Murzaev, E`.M. Central'noasiatskie toponimicheskie miniatyury. In: Bandaruk 1964, 3- 13
---, B.A. Nikonov, V. V. Cybul'cki (eds.). *Toponimika vostoka. Issledovaniya i materialy*. Moskva : Nauka 1969
---, *Onomastika vostoka*. [Sbornik statei, Ordena Trudovogo Krasnogo Znameni Institut vostokovedeniia, Institut etnografii]. Moskva : Nauka 1980.

Nichols, J.  The epicentre of the Indo-European linguistic spread.  In: Blench, R. and M. Spriggs 1997, 122-148
---, The Eurasian spread zone and the Indo-european dispersal. In: Blench, R. and M. Spriggs 1998, 220-266

Nyberg, H. The problem of the Aryans and the Soma: The botanical evidence. In:  Erdosy 1995: 382-406

Pakhalina, T.N., O proistoskhozhdenii toponimov Ishkashim, Yazgulyam i Vakhan. In: *Iranskoe Yašykoznaniye*. Moskva 1976

Parpola, A. The problem of the Aryans and the Soma: Textual-linguistic and archaeological evidence. In: Erdosy 1995: 351-381

Pinault, G.-J. Notes d'onomastique koutchéenne. *Tocharian and Indo-European Studies* I, 1987, 77-97
---, Tocharian Languages and Pre-Buddhist Culture. In : Mair 1998: 358-378

Pinnow,.H-J.., Zu den altindischen Gewässernamen. *Beiträge zur Namensforschung* 4, 1953, 217-234; 5, 1954, 1-19.
---, *Versuch einer historischen Lautlehre der Kharia-Sprache*, Wiesbaden 1959

Polomé, E. C. The non-Indo-European component of the Germanic lexicon. In*: o-o-per-ro-si. Festschrift für Ernst Risch*, ed. A. Etter, Berlin: Mouton de Gruyter 1986: 661-672
---, The Indo-Europeanization of northern Europe: the linguistic evidence. *Journal of Indo-Euroepan Studies* 18, 1990, 331-338.

Possehl, G. Meluhha. in : J. Reade (ed.) *The Indian Ocean in Antiquity*. London: Kegan Paul Intl. 1996: 133-208
---, The Transformation of the Indus Civilization. *Journal of World Prehistory* 11, 1997, 425-72

--- and P. Gullapalli. The Early Iron Age in South Asia. In: V. Pigott (ed.), *The Archaeometallurgy of the Asian Old World*. Philadelphia: The University Museum 1999: 153-175

Postelov, E.M. Gidronimicheskie arealy centraloi Azii. In: Murzaev 1980, 118-123

Pulleyblank, E.G., The Consonatal System of Old Chinese, *Asia Major* 9,1962, 58-144, 206-255
---, Chinese and Indo-Europeans. *JRAS* 1966: 9-39
---, Why Tocharians? *JIES* 23, 1995, 415-430

Rau, W. A Note on the Donkey and the Mule in early Vedic literature. *Adyar Library Bulletin* 44-45, 1980-81, 179-189

Rédei, K. *Zu den indogermanisch-uralischen Sprachkontakten.* Sitzungsberichte der Österreichischen Akademie der Wissenschaften, Philosophisch-Historische Klasse, 468 Band. Wien 1986
---, Die ältesten indogermanischen Lehnwörter der uralischen Sprachen, in: D. Sinor (ed.) *The Uralic Languages: Description, History and Foreign influences*. Leiden: Brill 1988: 638-664

Róna-Tas, A. *Tibeto-Mongolica*. The Hague: 1966
---, Tocharische Elemente in den altaischen Sprachen. *Schriften zur Geschichte und Kultur des alten Orients. Protokollband der XII Tagung der Permanent International Altaistic Conference* (1969). Berlin 1974: 499-504
---, On the earliest Samoyed-Turkish contacts. *Congressus Quintus Internationalis Fenno-Ugristarum. 3. Dissertationes symposiorum linguisticorum*. Turku, 1980: 377-385
---, Altaic and Indo-European: Marginal Remarks on the Book of Gamkrelidze and Ivanov. *Acta Orientalia Academiae Scientiarum Hungaricae* 42, 1988, 391-404.

Ronca, I. *Ostiran und Zentralasien bei Ptolemaios.* Diss., Mainz 1968.
---, *Geographie 6,9-21. Ostiran und Zentralasien. Griechischer Text neu hrsg. und ins Deutsche übertragen von Italo Ronca, mit der lateinischen Übersetzung des Jacobus Angelus, einer neuen englischen Übersetzung und texkritischen Noten.* Rom, IsMEO, 1971.

Rozenfel'd. A.Z. *Materialy po ètnografii i toponimike Vancha.* Izvestiya VGO. 85: 4, 1953
---, Zametki po gidronomii yugo-vostochnogo Tadzhikistana. In: Bondaruk et al., 1964: 175-182
---, Tadžikskaya toponimiya Badaxšana. (Gorno-Badaxšanskaya oblast' Tadžikskoi SSR). In: Murzaev 1980: 156-164

Sarianidi, V. I. *Drevneishii Merv. [In' gradymy Merv. The oldest Merv].* Moskva: Izdatel'skaya firma "Kruk" 1992
----, Food-producing and other neolithic communities in Khorasan and Transoxania: Eastern Iran, Soviet Central Asia and Afghanistan. In: Dani 1992: 109-126
---, *Margiana and Protozoroastrism*. Translated from Russian by Inna Sarianidi. Athens : Kapon Editions 1998.

Savina, V. I. O tipakh slovoobrazovaniya toponimov Irana. In: Bondaruk 1964 : 149-174

---, E`tnonimy v toponimii Irana. In: Murzaev et al. (eds.): *Onomastika Vostoka*. Moscow 1980: 140-155.

Schmid, W.P. Alteuropäische Gewässernamen in: Eichler et al. 1995: 756-762

Schmitt, R. Iranische Namen. In: Eichler et al. 1995: 678- 690

Shafer, R. *Ethnogeography of Ancient India,* Wiesbaden 1954

Sherratt, A. Patterns of Growth: Nodes and Networks in the Ancient World.
(in this volume)

Shishlina, N.I. and F.T. Hiebert. The Steppe and the Sown: Interaction between Bronze Age Eurasian Nomads and Agriculturalists. In: Mair 1998: 222-237

Skjaervø, P.O. The Avesta as a sorce for the early history of the Iranians. In: G. Erdosy (ed.), *The Indo-Aryans of Ancient South Asia*, (Indian Philology and South Asian Studies, IPSAS) 1, Berlin/New York: de Gruyter 1995, 155-173.

Staal, J. F. How a psychoactive substance becomes a ritual: the case of Soma. (forthc.)

Starostin, S.A. Praeniseiskaya rekonstrukciya i vneshnie svyazi yeniseiskix yazykov. *Ketskii sbornik: antropologiya, etnografiya, mifologiya, lingvistika*. Leningrad: Nauka, 1982: 144-237

Steinkeller, P. The question of Marḫaši: A Contribution to the Historical Geography of Iran in the Third Millennium B.C. *Zeitschrift für Assyriologie und Vorderasiatische Archäologie* 72, 1982, 237-65
---, *Marḫasi*. In: *Reallexikon der Assyriologie und Vorderasiatischen Archäologie,* Berlin: De Gruyter 1989: 381-2

Southworth, F. Lexical evidence for early contacts between Indo-Aryan and Dravidian. In: M. Deshpande and P.E. Hook (eds.). *Aryan and Non-Aryan in India.* Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan 1979, 191-233
---, Ancient economic plants of South Asia: linguistic archaeology and early agriculture. In: *Languages and Cultures. Studies in Honor of Edgar C. Polomé.* M.A. Jazayery and W. Winter (eds.), Berlin/New York : Mouton de Gruyter 1988, 659-668

Szemerényi, O. Structuralism and substratum. Indo-Europeans and Aryans in the Ancient Near East, *Lingua* 13, 1964, 1-29.
---, *Four Old Iranian Ethnic Names: Scythian - Skudra - Sogdian - Saka*. Sitzungsberichte der Österreichischen Akadademie der Wissenschaften, 386, Bd. 9, 1980

Thieme, P. *Kleine Schriften.* Ed. K. L. Janert. 2 vols. Wiesbaden. 2nd ed. Wiesbaden 1984.
---, Kleine Schriften 1970-1987. Ed. R. Söhnen. Stuttgart. 1991

Thompson, G. The relationship between Vedic and Avestan: the provenance of Soma, amshu, and its relation to the BMAC? Paper at: *Third Harvard Round Table on the Ethnogenesis of South and Central Asia*, May 12-14, 2001

(  http://www.fas.harvard.edu/~sanskrit/RoundTableSchedule.html)

Tikkanen, B. *The Sanskrit Gerund. A Synchronic, Diachronic and Typological Analysis.* Helsinki 1987
---, On Burushaski and other ancient substrata in northwest South Asia. *Studia Orientalia* (Helsinki), 64, 1988, 303-325
---, Archaeological-linguistic correlations in the formation of retroflex typologies and correlating areal features in South Asia. In: Blench and Spriggs 1999: 138-148

Tuite, K. Evidence for Prehistoric Links between the Caucasus and Central Asia: The Case of the Burushos. In: V. Mair 1998: 448-475

Turner, R. L. *A comparative Dictionary of the Indo-Aryan Languages.* London 1966 (CDIAL)

Vaijda, E. J. The Kets and their language, *Mother Tongue* IV, 4-16

Vallat, F. *Suse et Elam.* Paris : Boccard 1980
---, Eléments de géographie élamite (résumé) *PO* 11, 1985, 49-54

Vamberi, H. Die geographische Nomenklatur Zentralasiens. *Petermanns Mitteilungen*, 37. Gotha 1891.

Vennemann, Th. Linguistic reconstruction in the context of European prehistory. *Transactions of the Philological Society* 92 (1994) 215-284
---, Basken, Semiten, Indogermanen: Urheimatfragen in linguistischer und anthropologischer Sicht. In: W. Meid (ed.*), Sprache und Kultur der Indogermanen: Akten der X. Fachtagung der Indogermanischen Gesellschaft, Innsbruck, 22.-28. September 1996.* Innsbruck: Institut f  ür Sprachwissenschaft, 1998: 119--138
--- and A. Bammesberger, *Germania Semitica.* Heidelberg 2001

Werner, H. Reconstructing Proto-Yenesseian, *MT* IV, 1998: 20

Wirth, G. Arrian. *Der Alexanderzug; Indische Geschichte : griechisch und deutsch / Arrian; herausgegeben und übersetzt von Gerhard Wirth und Oskar von Hinüber.* München : Artemis Verlag 1985.

Witzel, M. Early Eastern Iran and the Atharvaveda. *Persica* 9, 1980, 86-128
 ---, Sur le chemin du ciel. *Bulletin des Etudes indiennes, 2*, 1984, 213-279
----, *Tracing* the Vedic dialects. Colette Caillat (ed.), *Dialectes dans les littératures indo-aryennes.* Paris 1989, p. 97-264
----,  Meaningful ritual. Structure, development and interpretation of the Tantric Agnihotra ritual of Nepal. *In: Ritual, State and History in South Asia. Essays in honour of J.C. Heesterman*, ed. A.W. van den Hoek, D.H.A. Kolff, M.S.Oort, Leiden 1992, 774-827
---, Early Indian history: Linguistic and textual parameters In: *The Indo-Aryans of Ancient South Asia.* G. Erdosy (ed.), (Indian Philology and South Asian Studies, A. Wezler and M. Witzel, eds.), vol. 1, Berlin/New York: de Gruyter 1995, 85-125 and : Ṛgvedic history: poets, chieftains and polities, pp. 307-354.

---, The Development of the Vedic Canon and its Schools: The Social and Political Milieu. (Materials on Vedic Śākhās 8). In: M.Witzel (ed.). *Inside the Texts, Beyond the Texts. New Approaches to the Study of the Vedas.* Harvard Oriental Series. Opera Minora, vol. 2. Cambridge 1997, 257-345

---, Substrate Languages in Old Indo-Aryan (Rgvedic, Middle and Late Vedic), *EJVS* Vol. 5:1 1999a, 1-67 (http://www1.shore.net/~india/ejvs)

---, Early Sources for South Asian Substrate Languages. *Mother Tongue* (extra number), 1999b: 1-70

---, Aryan and non-Aryan Names in Vedic India. Data for the linguistic situation, c. 1900-500 B.C. in : J. Bronkhorst & M. Deshpande (eds.), *Aryans and Non-Non-Aryans, Evidence, Interpretation and Ideology.* Cambridge (Harvard Oriental Series, Opera Minora 3). 1999c, 337-404

---, The Home of the Aryans. *Anusantatyai. Festschrift für Johanna Narten zum 70. Geburtstag,* ed. A. Hintze & E. Tichy. (*Münchener Studien zur Sprachwissenschaft, Beihefte* NF 19) Dettelbach: J.H. Röll 2000, 283-338

---, Autochthonous Aryans? The Evidence from Old Indian and Iranian Texts.
 *EJVS* 7-3 May 25, 2001a (http://www1.shore.net/~india/ejvs)

 ---, A first link between the R̥gvedic Panjab and Mesopotamia: *śimbala/śalmali,* and *GIŠ*$_{gišimmar}$? In: Karttunen, K. and P. Koskikallio (eds.) *Vidyārṇavavandanam. Essays in Honour of Asko Parpola.* 2000 (*Studia Orientalia*, published by the Finnish Or. Soc. 94): 497-508

 ---, forthc. a = Early linguistic data and the Indus civilization. In: J. Kenoyer (ed.) *Proceedings of the conference on the Indus civilization,* Madison 1998.

---, forthc. b. A structural and comparative look at Vedic Religion. Paper at the *Third Internatioal Vedic Workshop,* Leiden, May 2002.

Zide, A. and Zide, N. Proto-Munda cultural vocabulary: evidence for early agriculture. In: Ph. N. Jenner et al., *Proceedings of the First International Austroasiatic Conference.* Honolulu: University of Hawaii Press 1976, 1295-1334