
SINO-PLATONIC PAPERS

Number 129

December, 2003

Linguistic Evidence for Cultural Exchange in Prehistoric Western Central Asia

by
Michael Witzel

Victor H. Mair, Editor
Sino-Platonic Papers
Department of East Asian Languages and Civilizations
University of Pennsylvania
Philadelphia, PA 19104-6305 USA
vmair@sas.upenn.edu
www.sino-platonic.org

SINO-PLATONIC PAPERS

FOUNDED 1986

Editor-in-Chief
VICTOR H. MAIR

Associate Editors
PAULA ROBERTS MARK SWOFFORD

ISSN
2157-9679 (print) 2157-9687 (online)

SINO-PLATONIC PAPERS is an occasional series dedicated to making available to specialists and the interested public the results of research that, because of its unconventional or controversial nature, might otherwise go unpublished. The editor-in-chief actively encourages younger, not yet well established, scholars and independent authors to submit manuscripts for consideration. Contributions in any of the major scholarly languages of the world, including romanized modern standard Mandarin (MSM) and Japanese, are acceptable. In special circumstances, papers written in one of the Sinitic topolects (*fangyan*) may be considered for publication.

Although the chief focus of *Sino-Platonic Papers* is on the intercultural relations of China with other peoples, challenging and creative studies on a wide variety of philological subjects will be entertained. This series is **not** the place for safe, sober, and stodgy presentations. *Sino-Platonic Papers* prefers lively work that, while taking reasonable risks to advance the field, capitalizes on brilliant new insights into the development of civilization.

Submissions are regularly sent out to be refereed, and extensive editorial suggestions for revision may be offered.

Sino-Platonic Papers emphasizes substance over form. We do, however, strongly recommend that prospective authors consult our style guidelines at www.sino-platonic.org/stylesheet.doc. Manuscripts should be submitted as electronic files, preferably in Microsoft Word format. You may wish to use our sample document template, available here: www.sino-platonic.org/spp.dot.

Beginning with issue no. 171, *Sino-Platonic Papers* has been published electronically on the Web at www.sino-platonic.org. Issues 1–170, however, will continue to be sold as paper copies until our stock runs out, after which they too will be made available on the Web.

Please note: When the editor goes on an expedition or research trip, all operations (including filling orders) may temporarily cease for up to three months at a time. In such circumstances, those who wish to purchase various issues of *SPP* are requested to wait patiently until he returns. If issues are urgently needed while the editor is away, they may be requested through Interlibrary Loan. You should also check our Web site at www.sino-platonic.org, as back issues are regularly rereleased for free as PDF editions.

Sino-Platonic Papers is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 2.5 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/2.5/> or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

Linguistic Evidence for Cultural Exchange in Prehistoric Western Central Asia

Michael Witzel

Abstract

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 Archaeological 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 by a western off-shoot, the Mitanni Indo-Aryan of Syria/Iraq and by the language of related tribes indicated by some Indo-Iranian words in Kassite. All these represent series of intrusions by Indo-Iranian speakers into the world of the great Mesopotamian, Bactro-Margiana, and Indus civilizations and their acculturation.

§ 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 (BMAC), 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,

* This study is a detailed follow-up on earlier notes (Witzel 1995: 103, 1997b: xx-xxiv), lectures (Erlangen, Indogermanische Gesellschaft, Oct. 1997 = Witzel 2000c; Philadelphia, *Contact and Exchange in the Ancient World*, May 2001 ~ 3rd Harvard Round Table on the Ethnogenesis of Central and South Asia, May 2001, preprint: http://www.fas.harvard.edu/~sanskrit/images/C._ASIA_.pdf; Leiden, Third Intl. Vedic Workshop 2002 = forthc. b), and investigations (Witzel 1999a: 58-60, 1999b, 1999c: 388-393; Witzel 2000a, 2000c). - Special thanks are due to John Colarusso: he has suggested, just before printing, a substantial number of additions and corrections, especially from Caucasian; they are quoted below as "J. Colarusso, pers. comm."

¹ Sarianidi 1992, 1998a, 1998b Dani 1992, Francfort 1989, 1990, 1994, 1998, 1999, 2001, Hiebert 1988, 1992, 2001.

² Hiebert 2002, Colarusso 2002, however, see Mair 2001. A few Elamite seals have been found in S. Turkmenistan.

Akkadian, Sumerian, Elamite)³, and those west of the Indus area. The language(s) of the Indus civilization 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 which have unfortunately been overlooked 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.

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 to have been discussed or described in details in their texts.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 3rd to 1st mill. BCE, but they deal just with the border areas of Mesopotamia⁸ and furnish only some vague references such as that to Aratta, probably Arachosia.⁹ Similarly, we have only a few vague reminiscences in the earliest Indian texts (*R̥gveda*) composed in the Greater Panjab (c. 1200 BCE-1000 BCE)¹⁰ which seem to refer back to the area along the Volga (*Rasa*) and secondly, to the people along the River *Sindes* (Tacitus' name for the Merw or Tedzhen river): the *Dasa* or O.P. *Daha* (whom the Greeks called *Da[h]ai*), the *Arii*,

³ Languages known from barely more than the names given to their speakers in Mesopotamian sources, such as those of the Gutu and Lullubi at c. 2250 BCE, are neglected here. For the contemporary situation on the Iranian plateau, see Vallat 1980, 1985, 1993, Steinkeller 1982, 1989, Blažek 1999. For (possible) connections between Elamite and Vedic names see Blažek 2002.

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

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

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

⁷ See, however, V. Mair 2001; yet note Proto-Elamite seals close by, at Tepe Hissar, as well as at Shahdad, Shahr-i Sokhta, etc., and recent finds to the west of Tehran at Tepe Uzbeki.

⁸ Such as the Gutu and the 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 (just south of Arachosia/Aratta) and in Badakhshan. Note Steinkeller 1982: 250 with details about a green variety, "carnelian with green spots," possibly turquoise, from Marḥaši. However, the blue Badakhshan variety is more famous, until today. For Aratta see Steinkeller 1982, Vasil'kov and Gurov 1995, Gamkrelidze and Ivanov 1989: 36, Witzel 1995: 320-2, Blažek 2002b: 215-218.

¹⁰ The lowest date depends on the date of iron, c. 1000 BCE; see Possehl and Gullapalli 1999. For present purposes, "Greater Panjab" indicates the area from Gandhāra (Peshawar) and Swat in the west to Delhi and the Upper Doāb in the east, from the lower Pamir/Himalayan ranges in the north to the borders of Sindh and the Bolan in the south; however, the clear center of the *R̥gvedic* area is western and eastern Panjab/Haryana.

and the **Parna* (Ved. *Paṇi*, cf. Ptolemy, Geogr. 6.10.2 *Parnoi*, *Daai/Parni*, *Dacae*; otherwise *Dahae*). Pinault (2003) connects *Paṇi*/**Parna*, as loan word from the west, with Common Toch. **paniya* 'that which belongs to wealthy people' > Toch. B *peñiyo*, A *pañi* 'splendor' and takes the Gr. form *Parn-oi* as reflecting a local variant of Ved. *Paṇi* with "intrusive" -r- (cf. Kuiper 1991:70-81), however see below §5. (Blažek (2002: 219-226) compares Vedic *dasyu* with Elamite *taššu-p* "people", **taššu* "man"; note Romani *das* "non-Gipsy" < RV *dasa*).

In addition, after the sparse attestation found in the old Iranian and the much later Middle Iranian sources¹¹ most of these areas became Turkish speaking after about 1400 years ago; this has obliterated much if not most of the older Iranian and IIr. record, frequently even that of topographical names.¹²

Even in this unfortunate situation, we can retrieve, based on the 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 (c. 2400-1600 BCE)¹³ and of Greater Afghanistan. However, it is precisely these Indo-Iranian sources that have largely been neglected so far.¹⁴

For some years (1995-2002) 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.

Such substrate words are quite common in languages that have occupied the territory of an earlier people speaking a different language.¹⁵ In English, for example, such common words as *sheep* (Dutch *schaap*, German *Schaf*) belong to the Neolithic substratum¹⁶ of the North Sea coast of Northern Germany and Denmark, the homeland of Anglo-Saxon.

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

¹² We may note the proliferation of Central Asian place names ending in Turk. *-su*, *-kul*, *-kum*, etc. See the paper by P. Golden in Mair (forthc.) for information concerning the rise and the spread of the Turks. Needless to say we do not have adequate etymological dictionaries of Turkmen, Uzbek, or even for the Ir. languages (with the exception of one for older Turkic by G. Clauson, the dated one of P. Horn for Persian, and G. Morgenstierne for Pashto, H. Bailey for Khotanese Saka).

¹³ Based on new carbon dates, see Francfort and Kuz'mina 1998: 468; 2400-1500 BCE (post-urban: 1800-1500 BCE) in Francfort 2001: 152. "Greater Afghanistan" signifies the territory covered by this country and some adjacent surrounding areas.

¹⁴ For initial suggestions see Witzel 1995, 1999a,b; see below n. 158, 195, 204.

¹⁵ Exceptions are the territories of Australia, Polynesia, and the Americas when first settled.

¹⁶ Cf. Huld 1990, Polomé 1986, 1990, Vennemann 1994, 1998.

§ 1.2. Sources

In order to evaluate the scarce materials at our disposal properly, a brief look at our sources is in order.¹⁷ 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 *R̥gveda* (RV); it is composed in archaic Indo-Aryan (Vedic Sanskrit). It is followed by a number of other Vedic texts, usually listed as *Samhitās*, *Brāhmaṇas*, *Āraṇyakas*, and *Upaniṣads*. Linguistically, however, we have to distinguish five distinct levels: *R̥gveda*, other *Samhitās* (Mantra language), *Yajurveda Samhita* prose, *Brāhmaṇas* (incl. *Āraṇyakas* and *Upaniṣads*), and the late Vedic *Sūtras* (Witzel 1987, 1997a)¹⁸

The language of the RV is an archaic form of Indo-European. Its 1,028 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 rote learning, until today. It must be underlined that the Vedic texts are "tape recordings"¹⁹ of this archaic period. Not one word, not a syllable, not even a tonal accent were allowed to be changed.²⁰ The oral 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, and rivers (Witzel 1999c), and for loan words²¹ from contemporary local languages.²²

The *R̥gveda* was composed in the Greater Panjab²³ and is to be dated before the introduction of iron in the northwestern subcontinent around 1000 BCE (Possehl and Gullapalli 1999). Later texts cover all of northern India up to Bengal and southwards towards the Vindhya hills.

Some 4% of the words in the *R̥gvedic* 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 pre-IA substrate(s).²⁴

¹⁷ For the sparse Mesopotamian sources, see below (Steinkeller, Vallat).

¹⁸ For abbreviations of the names of texts see attached list.

¹⁹ The middle/late Vedic redaction of the texts has influenced only a very small, well-known number of cases, such as the development *Cuv > Cv*.

²⁰ They even preserve very *special* cases of sentence intonation, see Klein 1997, Witzel 2001a.

²¹ Summary and discussion for RV words by Kuiper 1991; for post-RV texts, see Witzel 1999a,b.

²² 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 scope of the present paper. 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). For the toponymy of present day Iran, see the useful web site at Tokyo Gaikokugo Daigaku: <http://www.aa.tufs.ac.jp/~kkami/AbadIranE.html>.

²³ See Witzel 1997a, 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 Bolan area.

²⁴ See Kuiper 1991, Witzel 1999a,b. This situation is remarkable: if one were to apply it to a Near Eastern context, it would mean that an ancient Jerusalem temple ritual might contain Philistine, Lebanese, Akkadian, Egyptian, or other "heathen" words. The Indian situation also differs remarkably from that of the Hittite empire, where the preceding non-IE language, Hattic, was actually used as the ritual language.

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 *Gāṇās* (Yasna 28-53), and his(?) contemporaneous ritual text embedded among the *Gāṇās*, the Yasna *Haptaṇhāiti*, 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.²⁵ 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)²⁶ 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 even less amenable to absolute dating than the Vedic ones. Mesopotamia (or early China) simply do not figure in all these texts.

The older Avestan texts (*Gāṇās*/Yasna *Haptaṇhāiti*) 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).²⁷ The few Old Persian inscriptions that have survived date from 519 BCE onwards. However, other than is the case with old Indian texts,²⁸ 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 (see n. 14, 158, 195, 204).

§ 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 and O.Ir. that do not conform to Indo-European/Indo-Iranian word structure (including sounds, root structure and word formation) and have no clear IE/Ir. etymology.²⁹

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*),

²⁵ Modern recitation depends on these written texts and cannot be used in the same way as Vedic recitation.

²⁶ Humbach et al., 1991.

²⁷ Discussion by Skjærvø 1995. However, the YAvest. local name of Bactria (*Bāxδi*) 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 in Old Persian (519 BCE) already one word, A^[h]*uramazdā*. For the transfer of Zoroastrianism into Persis (the modern province of Fars, i.e. southwestern Iran) see K. Hoffmann 1992.

²⁸ Note the ongoing debate, since the mid-19th century, especially S. Lévy, Przyluski, Kuiper, and the relevant summaries in Mayrhofer, *KEWA* and *EWA*; last update in Witzel 1999a,b.

²⁹ Lubotsky (2001) adds also some less indicative features: limited geographical distribution, specific semantics, i.e. a category which is particularly liable to borrowing.

while the Russian, Indian and Japanese *chai/cha*, 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 they belong to a long disappeared language. The reason is that all languages follow certain patterns, allowing only certain sounds or groups of sounds while others that are 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*).³⁰

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 thousands of years ago. For example, we have the Rhine (Lat. loan word *Rhenus* < Celtic **Rēnos* < IE **reinos*), Danube (Lat. *Danubius* ~ N. Iran. *Dān-*), Don, Gr. *Tanais* (from pre-W. Circassian *t'āna/t'ane* "Don", J. Colarusso, pers. comm.), 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-fratar* "good brother"), Sum. *Buranuna* / Akkad. *Purattum* / Elam. *ú-ip-ra-tu-iš*, all from a pre-Sum. substrate more than 5,000 years ago.

³⁰ 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 < *sitomu*, or woodchuck (*Marmota monax*) from Algonkian *otchek*, *ochig*, *odjik* "fisher, weasel"; nevertheless, the English folk etymology gives the word away (Witzel 1997b).

The early river names of most of Europe belong to one and the same old system.³¹ A different prehistoric system is found in Greece and the Aegean area, with the typical pre-Greek *-s(s)-*, *-nt-*, *-mn-* suffixes. This phenomenon has been extensively discussed for much of the 20th century.³² 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 *Lugodunum* "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*, *they*, *she*, *he*, place names in *-vik*, *-ay*), then of Norman French with a large number of loans (*beauty*, *ancestor*, *-ville*, etc.), and finally an equally 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*, *lentil*, *land*, *delve*, *prick*) derived from a long-lost prehistoric Northern European language.³³

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.

³¹ Explained, since H. Krahe, as an "old European" layer of IE (summary by W. P. Schmid 1995); 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.

³² Summary by Th. Lindner 1995.

³³ Cf. Hamp 1998: 328, Huld 1990, Vennemann 2001.

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
(dev. of various Prakṛts)

Arabic superstrate loans
(dev. to Middle Iranian)

Old Greek loan words
Old Persian/Iranian loan words

some Old Greek loans
Old Persian /Later Avestan

(development from Vedic to MIA)

Later (R̥g)Vedic / OIA dialects
Dravidian adstrate
<----immigrant Old Indo-Aryan

Old Avestan
Old Iranian superstrate
<-----immigrant OIA in Iran

*Sarasvatī, Sarayu, etc.)

Central Asian substrate

Harappan language (see below)

Indo-Iranian in C. Asia, south of Uralic,
Ket (Yen.)

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, which has obliterated much of the older Iranian, 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, even in Northern India.

However the combined data of the reconstructed PIIr. language allow us 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* (see §6). 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.³⁴ The large amount of PIIr. data permits us to find loan words from IIr. languages and accompanying cultural contacts with neighboring languages and language families.

These are those of the Uralic (Finno-Ugric) and Yeneseian (Ket) languages to the north, the two 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 the various Indo-Iranian languages themselves in their post-immigration homelands, in Greater Iran and in the Greater Panjab, to the South.³⁵

³⁴ Some of these deities may be reflected in the BMAC, see Witzel 2000a, forthc. b, and below cf. n. 150-154 for the same origin of some of the IIr. deities.

³⁵ Occasionally the Mesopotamian sources can provide some data such as on Aratta, see Steinkeller (1982), and in great detail Vasil'kov and Gurov (1995) who discuss Dravidian possibilities. Chinese sources are geographically too distant for most of the time before Zhang Qian (2nd cent BCE), though some loans from IE into Sinitic can be discerned, see below, n. 61.

2.1. The Northern Border

§ 2.1.1. Uralic, Finno-Ugrian, and Yeneseian

Starting in the extreme north, it is very important to note that early IIr. loan words are preserved in Uralic and more specifically, its branch, the Proto-Finno-Ugric (PFU) languages. They establish the existence and spread of P-IIr. in the areas bordering the PFU homeland³⁶ in the East European and Siberian woodlands and, thus, also the relative age of the speakers of IIr. This evaluation is based on the exact form of IIr. that the various loan words preserved in Proto-Uralic and its somewhat later western branch, Proto-Finno-Ugrian, have retained: in other words, these loans form virtual "archaeological" layers of a contact situation that persisted for several millennia.

Koivulehto (2001: 236-238) adduces 11 loans from PIE into Proto-Uralic (PU) and PFU, such as PIE **wed-er/en-* "water" (cf. Rédei 1986: 43) > Finn. *vesi/ved-*, PSamoyed **wit*; PIE **wosā* "bought object, merchandise, ware" (Hitt. *waš* "to buy", Ved. Skt. *vas-nā* "price") > Cheremis *uža* "price", Finn. **wos-ta* > *osta* "to buy"; PIE **g'halgho-* "long thin pole" > PFU **šalka*, Finn. *salko*, Mordvin *salgo*. The rest of the words are, surprisingly, verbs: to fear, to plait/spin, shall/must, to walk/wander, to exchange/sell, to wash, to bore, to lead/draw (PIE **wedh* "to lead, marry" > PFU **wetä*, Finn. *vetä*, Hung. *vezet*). To be noted is the preponderance of words relating to exchange, commerce and cultural borrowing. Koivulehto adds 15 words that have come from PIE into western FU languages, among which PIE *pork'o-s* (see below), and 5 that are *not* attested in northwestern IE languages.

The remaining 26 words are from the stages of (P)IIr. and (P)Ir. In 1986, Rédei had mentioned only a few loans that might be attributed to the Proto-IE period, such as PFU **mete* "honey" < PIE **medhu* (cf. Koivulehto 2001: 247) or PFU *wete* "water" < PIE **wed-or-*.³⁷ Harmatta (1992) has mistakenly subdivided the IIr. loans into 17 stages³⁸ which must be collapsed into just a few linguistically attested stages.

About half of the loans discussed by Koivulehto (2001) and even more of them in Rédei (1986) come from the Proto-IIr. (or pre-Proto-IIr.) period during which PIE **k^w*, *k^{wh}*, *g^w*, *g^{wh}* became **k*, *kh*, *g*, *gh*; thus, FU **warkas(e)* "wolf", P-Samoyed **warkə* "bear" < PIIr. **vrka-s* < PIE **wlk^wo-s*, etc. Another early, pre-Proto-IIr. loan is **ket^stro* "spindle" > Finn. *kehrä*, *keträ*, Mordvin *št'efe*, *kšt'ir* (Koivulehto 2001: 249), with the retention of PIE *-e-* and *-tst-*.

³⁶ See the brief summary by Kuz'mina (2001: 291) and other relevant papers in the same volume (Carpelan et al. 2001).

³⁷ Rédei has the following oldest loans: **miye* "to give, sell", **muške* "to wash", **nime* "name" (which may rather be Nostratic, cf. Jpn. *na-*, etc.), **sene* "sinew", **toye* "to bring", **waške* "metal", **wete* "water".

³⁸ 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, 1999) 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. - For further details on Uralic and IIr. see Joki 1973, Rédei 1986, 1988, Katz 1985, 2001, Koivulehto 2001.

Later on, *P-IIr. **k'*, *k'h*, *g'*, *g'h* developed to IIr. *é*, *éh*, *j*, *j'h* as seen in FU, F-Volg. **poréas*, *poršas* "piglet" (Koivulehto 2001: 242 derives this from PIE, but runs into problems with W. FU derivatives); still later, IIr. *é* > *ś*: FU **śata* "100" (Koivulehto 2001: 248), FU **śaka*, *śawa* "goat" (Rédei 1986: 59), FU **rešmä* < **račmi*.³⁹

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. *bogŭ* "god".⁴⁰

It is important to underline that the bulk of the data comes from the PIIr. period and this allows us to posit speakers of PIIr. at the northern fringes of the steppes, just south of the *taiga* belt, -- in other words, the linguistic and cultural ancestors of the later Vedic and Iranian tribes were the southern neighbors of the Uralic tribes (cf. below, §6).

§ 2.1.2. Another northern neighbor is the isolated Siberian Ket language that is still spoken on the middle Yenesei river.⁴¹ 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 Yenesei 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 Yenesei-Angara rivers (Vaijda 1998: 10, Blažek 1998: 27). Their hydronyms are typified in Ket *se's* "river", as well as in names ending in *-ces*, *-tet*, *-set*, *-šet*; *-ul*; note also Yen. **xur*₁ "water", Kott *kem*, Pump. *tōm* "river".

Like Uralic, the Ket (Yeneseian) languages have a number of old loan words: Kott *art'a* "true, veritable" < IIr. **ṛta*, Ved. *ṛta*, OAvest. *arəta*, O.P. [*ərta*], Median *arta*, YAvest. *arəta*, Mitanni (and Greek historians) with the spelling *arta-*; Kott *čak* "force", Kott *čaga* "strong" < IIr. *čak* > Ved. *śak* "to be able; force," perhaps also Ket *ku's*, Yug *ku's*, Kott *husa*, Arin *kus*, Pumpokol *kut* "cow" < Pre-IIr. **g^wāus*, PIIr. *gāus*, Iran.

³⁹ See Rédei 1986; Koivulehto 2001: 250 reconstructs an unattested (pre)IIr. source **rečmo/a* (Ved. *rašmi-*). However, while the early vocalism may be in order, early PIIr. would still have had **rek'mi-*. There are indeed problematic representations of certain vowels in Uralic when compared with their successive sources in (P)IIr. or older (pre-IIr.): old IE, Pre-IIr. *-o-* + PIIr. *-é-* in F-Volg. = IIr. **poréas*, or Common/Late PIIr. (*k'* > *é*) in F-Volg. **oraše* < IIr. **varājha* < PIIr. **wərojho*, and in later IIr. forms (*é* > *ś*, *j* > *j*), as in F-Volg. **wašara* < IIr. *vājra* < PIE **h₂weg'-ro*, or Koivulehto's "preserved IE *e*" in a form taken from early Plr. (FU **serä* "old" < Plr. **dzero* ~ Ved. *jarant*). The FU representation is often due to the necessities of the FU sound system, FU vowel harmony, and certain substitutions (*e* > *i*, *o* > *u*), as seen in F-Volg. **rešmä* < **račmi*. Cf. also Rédei 1986: 33 sqq. for similar substitutions of Iran. vowels in Permian (Votyak) loans, such as *a* > *o*, *u*; *ā* > *u*, etc. -- Harmatta's FU and IIr./Ir. forms (1992) are not reliable.

⁴⁰ This word, via normal Slavic sound shifts, derives from N. Iran. (Scythian), i.e. *baga-* "god"; likewise many of the Ukrainian/S. Russian river names (Don, Donets, Dnyeper, Dnyestr, 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 directly from IE. PFU **pakas(e)* "luck" > Mordvin. *pavas* "luck" shows the older meaning, from IE **bhagos* > PIIr. *bhagas* "share", see above.

⁴¹ For a detailed discussion see several articles in the journal *Mother Tongue* IV, Boston 1998: 4-32.

gauš and Tocharian A *ko*, B *keu*, but note, rather, an origin of the Yen. words in P.Yen. **ku's* "horse".⁴²

§ 2.1.3. Some words may shed more light on the old (west) Central Asian language(s) of the area that are now lost. IE **medhu* "honey, mead"⁴³ 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'at* < IE **medhu*, Gr. *méthui* "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. *mür*, Chin. *mi* < **mjit/mit*,⁴⁴ Sino-Kor. *mil*, Jpn. *mitsu* < **mit(u)*. Its Iranian form, Iran. **maδu* > Turk., and Mong. *bal*, Korean *bel* (*beol*) "bee" (cf. Jpn. *hachi*; note Arab. *mādī*?); finally also into 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 *miliþ* point to a more western source, **melit*, perhaps in the Balkans/Anatolia. Note that the early reconstructions of Nostratic⁴⁵ list both forms under **majla* > Ural. *majδ'a*, Drav. *maṭṭ*, *miṭṭ* (DEDR 4662 *maṭṭu*), Altaic */m/ala*, *bala*.⁴⁶

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.⁴⁷

⁴² Yugh *ku's*, Kott *huš*, Arin *kus*, *qus*, Assan *huš* < Proto-Yen. **ku's* "horse" (Blažek 1998: 27); see n. 66.

⁴³ Cf. Lubotsky 1998: 379, Koivulehto 2001; on the eastwards spread of the honey bee, Carpelan and Parpola 2001.

⁴⁴ Since Polianov (1916), see Lubotsky 1998: 379; see Carpelan and Parpola (2001: 115 sq.) for the eastward spread of beekeeping.

⁴⁵ Nostratic is the reconstructed source, at c. 10,000 BCE or earlier, of IE, Uralic, Altaic, Dravidian, Kartvelian (SW Caucasian: Georgian, etc.), and Afro-Asiatic (formerly, Hamito-Semitic), see Illich-Svitych 1976: 38 sq.

⁴⁶ Note, even beyond this area, in Polynesia, though the area had no bees or honey before European colonialization: Samoan *meli*, Hawaiian *mele*, *meli*; *mele*, *melemele* "yellow", Maori *miere*; Rarotongan *meli* "honey", Mangareva *mere* "honey". (However, Tongan *melie* "sweetness, sweet, delicious" apparently stem from Proto-Pol. **malie* "pleasant" -- pers. communication by R. Clark, 4/4/2002). Much of this could be old, assuming an older S.E. Asian/Melanesian origin of the underlying concept: A variety of the bee, *Apis cerana* (or *Apis indica*) is found in India and S. Asia up to Java and Borneo, and stingless bees, *Melliponinae*, occur all over the southern hemisphere, including Australia, New Guinea, and Melanesia (Carpelan and Parpola 2001: 116), an area where the Proto-Polynesians have passed through. Thus, the Pol. words must not necessarily be an introduction by missionaries (from French *miel* "honey", though R. Clark tells me that there were early French speaking missionaries on the South Island of New Zealand). For, the word for is also seen in Haw. mythology: *mele-mele* "a star name": *Melemele* ("yellow star"?; MW) 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". -- Finally, for the spread of the word for "honey", note the role of *cire perdue* (also typical for the S.E. Asia) with the inherent use of bee's wax, see Andrew Sherratt, in V. Mair (forthc.)

⁴⁷ Note, again, A. Sherratt on the spread of bronze smelting and *cire perdue* casting along the *taiga* belt.

The old Central Asian word for "lion", **sengha/singha* has a similar spread and variation (cf. Behr n.d., Blažek, n.d.): Ved. *simha* "lion" < **sinjha* < **singha*. However this differs, even within Indo-Iranian, from the Proto-Iran. form **sarg* that has resulted in Khot. *sarau*, Khoresm. *sary*, Sogd. *šryw/šrw*, Parth. *šarg/šgr*, Pahlavi *šgr*, *šyr*, N.Persian *šēr* (Horn 1893, no. 803); cf. also Elamite-O.P. **Šargu-dāta* (*Šá-ir-ku-da-da*, etc.). Blažek also derives Ved. *śrgala* "jackal" from an Ilr. form **šrgu* (Pali *sigāla*, Kashmiri *śāl*, Nirlami *šakāl*, Nuristani *šyōl* "wolf", (Prasun) *šil*, *šēli*; cf. Bur. *hal* "fox").

Perhaps some Dravidian words belong here as well (Blažek, n.d.): *ci(v)vaṅki* DEDR 2579 < *ciṛu(C)t-ai* as in Telugu *ciṛu puli* (*puli* "tiger"), Tulu *ciṛteṇṇi* "leopard", Tamil *ciṛuttai* "panther".

However, one will have to compare Tib. *seṅge*, W. Tib. *seṅge*, *siṅge*, Zhañ Zhuñ *saṅgo*, Lepcha *sung-gi* (Behr, n.d. 15); Chin. *suan-ni* < Middle Chin. **swan-ŋej* < O.Chin. **aso[n/r]-aŋe* (Behr, n.d.: 10, **suân-ŋei* Karlgren, Henning). However, Starostin (1989: 402) reconstructs **Cwân-ŋ(h)ē* and S.E. Jakhontov (in Blažek n.d.): Old Chin. **sōr-ŋē*;⁴⁸ for the Central Asian interchange of *-n/-r-* see below §5.⁴⁹ Another word contains variations of O.Chin. **tsu[r/n]-ŋe* (Behr, n.d.: 10 sq.). The common word, however, is mod. Chin. *shi-zi*, from "Arch. Chin." **ši-* (Karlgren, or < **šəj-*, **srij-*, see Behr: 5, derived by Pulleyblank, via **šəjčá*, from Toch. *šecake*); cf. also Jpn. **si-* > *shi(-shi)*.

Further west, Toch. A *śisäk*, B *šecake* "lion" < *šecake* < **šec-äke* with the common, borrowed Ilr. suffix *-a-ka* (Pinault 2002: 331; for other etymologies, see Behr n.d.: 17-20), and perhaps also Armen. *inc*, *inj* seem to go back to a S. and E. Central Asian variety of the word such as **si(n)é-*, which is close to pre-Ved. **sinjha*. Behr (n.d.: 20), too, thinks of an unknown Central Asian language as the ultimate source of the Toch. B and Chin. words.

Such forms are indeed found in N. Caucasian: PEC **čä:nq V* "lynx, panther" (PNEC **-oniq:qo-* J. Colarusso, pers. comm.), Nakh **čōq* "snow leopard" > Chechen *čoq*, Avar-Andian **čirq:q V* > Avar *čirq:q*, Akhvakh *čiq:q-o* "lynx", Godoberi *čirq:q-u* "snow leopard"; Dargwa: Akusha *čirq* "panther", Lak *činiq*. Note again the interchange between forms with *-r-* and *-n-* that have not been explained previously (Blažek, n.d., also draws attention to Assyrian *simkurru* "a hunted mountain feline, gepard", and

⁴⁸ See now Pulleyblank 1995; Lubotsky 1998: 379. For details of the Chinese attestations see Behr, n.d.: 5 sqq.; note that there is no early attestation of the lion in China, except for the recent finds of some small lion figures at Erlitou III-IV (late 3rd mill. BCE, Behr, n.d.: 3). -- There existed a sub-variety, the "Iranian maneless lion" (Masson 1992: 39; apparently different from the one seen in Persepolis sculptures; for this *panthera leo persica* see Behr n.d.: 3); another remnant population of the Asian lion is still found in India, in the hills of Girnar (W. Gujarat, cf. Behr : 2).

⁴⁹ For an inner-Chinese explanation of the *r/n* change in this word see Behr n.d.: 15; the rare and early *suan-ni* is glossed "shizi" in Guo Pu's commentary on the Mu Tianzi zhuan (3rd c. BCE), < O.Chin. **so[n,r]-ŋe*, close to the Tibetan form. -- The word was reconstructed by Henning (see KEWA, s.v. *simha*) as **siengha*; however, there must have been another form from another Central Asian dialect or language, based on **šer(gh)*, or now with Blažek (n.d.) Iran. **šargu/sargu* < *šorg(h)u* < *šrg(h)u*; cf. below §5, on the interchange of *r/n*. -- Not related is Turk. (etc.) *arslan* "lion," as is, incidentally, the often quoted Suahili *simba* < PBantu **-cimbà* "wild cat" (Behr, n.d. 14).

Elamite place names such as *Šenkuru* > mod. *Senkereh*, as well as Akkadian *zirqatu* "wild cat, lynx").

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*, and *leōn/le(w)on(t)*- > Lat. *leōn*- (cf. Behr, n.d.: 16 on Toch. *lu*, genitive *lw-es* "beast").

It is imperative that more such data be identified and collected to arrive at a closer picture of the northern parts of W. Central Asia.⁵⁰

§ 2.2. The Eastern Border : Eastern Central Asia

It is much more difficult to specify the linguistic eastern border of Western 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, roughly present Xinjiang) is as lacuneous as that in Western Central Asia. To some extent, place names that can fill in the gap in the evidence available so far. 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 our first notices: a few personal names and relatively more place names. However, most of them (such as *Mara-kanda*,⁵¹ *Zari-aspa*) are already of clear O. Iranian etymology and hardly go beyond the eastern boundary 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),⁵² or by those from the Pamirs.⁵³ Gryunberg (1980: 168) gives a long list of relevant place names from Afghan Badaxšān⁵⁴ and specifies that these "substrate" names point to a widespread language,

⁵⁰ Leaving aside the speculations of Harmatta (1992) and worse, Sergent (1997) on Dravidians in C. Asia, including Harmatta's unlikely *Haftyār* from Ir. *haft* < Ilr. **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). See now Vasil'kov and Gurov 1995: 36, who derive the word *Aratta* from Drav. *DEDR* 372 *er* "mountain ebony, *Bauhinia racemosa* L." + *DEDR* 101 **aṭṭV* "ravine", thus "ebony ravine".

⁵¹ E.Iran. *kanda* "town", cf. Pāṇini's 4.1.103 *kanthika* "inhabitant" of Varṇu, mod. Bannu; see below n. 160, 161 for a list of Sogdian names; cf. Khromov 1960, 1980, Èdel'man 1980.

⁵² Khromov 1960 specifies suffixes such as *-ef*, *-uf*, (< Sodg. *-uv*), *-ic* (< Sodg. *-yc*), *-kam* (*-kand*) < *kt* "town", *-yār* < *yr* "mountain", *-zoi* < *z'yh* "earth, land", *-rāt* < Sodg. *rwt* "river" - *rovut* (cf. Tajik *rāvad*) "meadow".

⁵³ See Savina 1980: 140-155; Èdel'man 1975.

⁵⁴ 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*, *Ghezv*, *Rāj*, *Purzārg*, *Našer*, *Imj*, *Kher*; *Yāsic*, *Iwinak*, *Sucu*, *Yakhcew*, *Ezwān*, *Pijangīw*, *Bušt*, *Winj*, *Afrij*, *Korkhu*, *Bahārak*, etc. Gryunberg 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.

or at least to a 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.⁵⁵ 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.

Data⁵⁶ from the surviving E. and N. Iranian (Saka) languages could be added, including the only eastern descendent of Saka, Sariqoli, on present Chinese territory.⁵⁷ Further, the contribution of Burushaski and its earlier forms⁵⁸ as well as an unknown substrate present in the Pamir area (Berger 1960, Jettmar 1975: 190, Tikkanen 1988, Blažek 1998: 449 sq.), and in the IA language Khovar⁵⁹ as well as in the Hindukush (Èdel'man 1968: 58) should be compared.

A closer study of the local names in the (Gāndhārī) Niya Prakrit in S. Xinjiang⁶⁰ with personal names such as *Šekrase*, *Bhimase*, *Bugosa*, and of the substrates preserved in Tocharian,⁶¹ with place names such as *Yursa*, *Kuci*, *Hippuka*,⁶² would yield further information on the eastern neighbors of prehistoric Western Central Asia.

In fact, G. Pinault (2003) has recently pointed out that some words that have been identified as stemming from the general area of the BMAC (below, §3.2) are also found in Tocharian. Such words appearing in both Tocharian languages (A, B) must go back to early Common Tocharian. Like the BMAC loans in O.Iran. and Vedic, they do not have IE word structure and etymologies and also cannot have entered Tocharian at the later stage of contact with Iranian and MIA as they follow the general sound shifts from PIE to Tocharian (*a > æ > Toch. B e, A a, *ā > ā > Toch. B o, A a). They include words such as *išt(i)* "clay, mud brick", *anēu* "(rusty) brown", *éarwa*, "hunting, living in/from the forest", *pani* "wealthy", *āṇi* "hip", *athr* "superior, over-coming force" (see §3.2, 3.4, 3.5, n. 146, 150, 150). Pinault sums up his new evidence as being very similar to the BMAC one: voiceless aspirate stops, syllabic liquids, several palatal stops and palatal clusters, even retroflexes (see, however, §5). This new evidence now provides the Eastern rim of influence of the BMAC language(s).

⁵⁵ 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.

⁵⁶ See also the discussion of many linguistic details of the Greater Pamir area in Èdel'man's (1968) discussion of Indo-Iranian linguistic geography.

⁵⁷ Still widespread in the hills and mountains south of Khotan. The western variety of N. Iranian is preserved in the north and south of the Caucasus range as Ossete, a descendant of Alan.

⁵⁸ Old Burushaski names in v. Hinüber 1995; comparison with Basque and Caucasian by Bengtson 1992, 1999, 2001, cf. Harmatta 1992, Tuite 1998 for a comparison of Bur. and Caucasian grammar and religion, Parkes 1987 for a comparison of Pyreneean, Caucasus, Pamir ethnology, with a binary system (male/fem. : pure/impure) of activities and correlated wild and domestic animals.

⁵⁹ See Kuiper 1962: 14, Witzel 1999a,b.

⁶⁰ See discussion by von Hinüber 1995: 663.

⁶¹ For the contacts between Tocharian and Sinitic see Pulleyblank 1966, 1995, between Tocharian and Altaic, see Clauson 1972, 1975, Róna-Tas 1974, cf. also Róna-Tas 1980.

⁶² O. von Hinüber 1995: 662; cf. the place names in Pinault 1987, 1998.

As can be expected, Tocharian, which seems to have split off from PIE at an early stage (Hamp 1998), and moved into Eastern Central Asia at perhaps c. 2000 BCE, also shows contacts with Uralic (Ivanov 1985) and Altaic languages (Róna-Tas 1974, 1988, Reinhart 1990), especially early Turkic (Clauson 1975, Pinault 1998), and since Shang times with Sinitic (Pulleyblank 1966, Lin 1998)⁶³ as well as with Tibeto-Burmese (Sapir 1951, Blažek 1984).

All of this leads, however, much beyond the frame of the present study.⁶⁴ Nevertheless, it should be pointed out that there are additional, 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 whole area, from Armenian and Iranian to Tibetan, Tocharian and Chinese, and so does ***med(h)-/melit-* from Finno-Ugrian and PIE to Chinese and Japanese on the one hand and to Hittite and Latin on the other. In this connection it should be noted that one incentive for the eastwards spread of the word for bees and honey may have been that of the use of bee's wax in the early copper casting technology of *cire perdue* (see A. Sherratt, in V. Mair, forthc.) that spread along the *taiga* belt of S. Russia and S. Siberia (see n. 46-47).

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*⁶⁵ :: FU **lox* :: Yeneseian (Ket) **ku's*⁶⁶ :: 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*,⁶⁷ Hittite

⁶³ Candidates for late IE/Toch. loans into Chinese include such words as **lak* "milk" (note the Himalayan NIA language Bangani with non-IA *bktə* "milk", and cf. below §6 for a possible arrival of the Proto-Bangani speakers from C. Asia); **mit* "honey"; **khwin'* "dog" (cf. TB **kwi* "dog/puppy").

⁶⁴ Early contacts of Sinitic with Tocharian and/or other IE, and those of Tocharian with Altaic are excluded here.

⁶⁵ Including Toch. A *yuk*, B *yakwe*, Hieroglyphic Luwian *á-zú-(wa/i)*, Cuneiform Luwian *azzu(wa)*, Lycian *esbe*, Hurrian *ešši*, *iššiya* < Mitanni IA, and other loans such as into Uralic; other reconstructed forms include: **Heh₃k'ú-* / **H₂ek'u-*, **oH₂ku-*, *H₃eku-*, *H₃eH₁k'u-/H₃H₁k'u* (Blažek 1998). For the IE designations of the horse see Hänsel et al. 1995, Raulwing 2000; Blažek 1998: PIE *Hekwo-s* from Proto-Yen. *ik-ku's* "stallion", **kus* "horse > cow", that is close to the Neolithic (horse) hunters at Botai in Kazakhstan, 3300-2700 BCE; J. Colarusso, (pers. comm.) regards Yen. *ik-kūš*, *kus* as a borrowing from IE.

⁶⁶ Arin *kus*, *qus*, Assan, Kott *huš*, Yugh *ku's*, cf. Werner 1998: 20.

⁶⁷ Tuite 1998: 464, cf. Andij *k'otu* "horse" (Andic, belonging to the Avar-Andic branch of NE Cauc.). J. Colarusso (pers. comm.) adds, substantially: "The NEC forms for "horse" show, as is usual, a number of distinct roots (hachek = pharyngealization, C' = tensed and prolonged consonant, unaspirated if voiceless or strongly ejective if glottalized): Avar *ču*, Andi *k'otu*, Akhvakh *určé*, Chamadal *určé*, Dargi *urci*, Lak *č^ou-*, *ču-*, *duč-*, *šugu*, *ča'tu*, *k'otu*, *k'^oatu*; but Tabasaran *haywán*, Aghul *hăywán*, Tsakhur *balkan*, Lezgin *p'alk'an*, Khinalug *ps-i/ps-o-i*, Udi *čk-*, *čkw-* (oblique stem). From which I would reconstruct the following: **ču-*, with derived stems **d/rV-ču-*, metathesized to **ur-ču-*, with the odd **k'otu-*, **p'alk'an*, and **hăywán*. These may have meant "mare, stallion", etc. The Udi clearly reflects some very old IE form. The

kur-ka "foal" (EWA 373), Sumerian *anše kur* "mountain ass" (rather the *kur*-donkey?), Drav. *kutir(ai)*,⁶⁸ Munda (Koraput) *kurtag*,⁶⁹ Korku *gurgi* (= *kurki*), Sabara (Sora) *kurtā*, Gadaba *krutā* < **ghurta*, Tib. *rta*, but Tsangla (in Bhutan) *kurtā*,⁷⁰ Meithei (Manipuri) *sa-gol*,⁷¹ Bur. *ha-yur* < **sayúr*,⁷² Drav. **gur-*,⁷³ Brahui (*h*)*ullī*, Khasi [*u*] *kulai*, Amwi *kurwa*, Indones. *kudja*; note also Late Vedic loan word *ghoṭaka*, (modern NIA *ghor-*, 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 derive from an isolated, hypothetical IE, or rather pre-Germanic **ghūl-*. Taking into account also Yeneseian **ku's* and Turkic *(*x*)*at*, may one posit an older Eurasian form ***kuC* : **kur/kul* :: **ku'* / **h₁k^w* - / **k'ot-u*?

Finally there are further, isolated designations such as Munda *sadom*,⁷⁴ Eastern Himalayan (Dhimal) *ōnyha*, Lepcha *on*; Drav. *(*h*)*ivuli*,⁷⁵ Tib. *rta*,⁷⁶ Elam. *lakpilan* (Blažek 1999: 64).

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 Modern Mongolian *morin*, *mörin* < **morī*, Tunguse *murin* (borrowed into Ghilyak as *mur*, *murng*), Korean *mar*, *mal* < *māl*), Japanese *uma* (*mume*) / Ryukyū *nman* < **uman*, **mVrV*, Chinese *ma* < **mraq* (*mra'*) < **mVra*, (or **mwa*, Benedict 1972: 189), Dun-Huang Tib. *rmang*, Burm. **mrang-h* > *myin-h*,

Khinalug might reflect an Iranian loan (older than that for "colt"): **aspa-* > pre-Khinalug **apsa-* > **psa-* > *ps-i-/ps-o-*, or it may be just one more odd root. -- The Northwest Caucasian forms are all straightforward (W.Circass. /ɬ'ə/ (retroflexed, aspirated), Ubykh /ɬ'a/ (laminal) "horse", /ɬ'əda/ "donkey", Abkhaz-Abaza /(a-) ɬs/ "horse"). These are clearly cognate with the first NEC root, and point to a common NC **č'u, The PNC form with an initial syllable, perhaps **hača* (*h* = voiceless pharyngeal, č = laminal), is the word for 'brother.' This form might even be **ay-č'ha*, with a frozen reciprocal prefix and the root for 'horse' used as a totemistic reference for 'brother'. Blažek has cited an Abkhaz form, /ača/, with the ubiquitous indefinite noun marker /a-/."

⁶⁸ Drav. (DEDR 1711 b) Tam. *kutirai*, Tel. *kudira*, *kudaramu*, Kota *kulyr*, Toda *kīṛīr*, Kodagu *kudīre* 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?).

⁶⁹ Drav. (DEDR 1711) *kutirai* "horse" > Koraput Munda **kuXrtag*, see Zide & Zide 1976: 1331.

⁷⁰ Shafer 1954: 25 **ghurta*, contra J. Bloch's **ghutr-* (Bagchi 1929: 24).

⁷¹ Berger 1959: 27, n. 36, p. 33 sqq. (Ling. Survey I, pt. II, p. 105).

⁷² Berger 1959: 27 n. 36.

⁷³ DEDR 1711(b): Tel. *gurramu*, Kolami *gurram*, Naikri *ghurram*, *kuramam*, *kurrmam*, Parji *gurrol*, Gondi *gurram*, Konda *guram*, Kuwi *gūrumi*, *gurromi* "horse"; (for Burrow on Tel. *gurramu* < Skt. *ghoṭaka*? see IJDL 1, 23-24).

⁷⁴ Pinnow 1959: 78 §70: "hardly to be connected with Khmer *seh*, Bahnar *yseh*".

⁷⁵ Tam. *ivulī* and Brahui (*h*)*ullī* "horse" (< "half-ass, hemione/onager", Burrow 1972, McAlpin 1981: 147; DEDR 500), see now Blažek 1999: 64 for a possible connection with Semitic **ib(i)l-* "camel"; -- other Drav. words include: DEDR 500 Tam. *ivulī*, Brah. (*h*)*ullī*, 1711 Tam. *kutirai* etc., DEDR 3963 Tam. *pari* "runner", 4780 Tam. *mā* "animal" (horse, elephant), Tel. *māvu* "horse".

⁷⁶ Tamang, Gurung *ta*, but cf. *kurtā(g)* (above).

Kachin *gù-mrà*; Thai *ma*, note Drav. *mā(v)*⁷⁷, Nahali *māv*. Janhunen (1998: 415) considers them to have been "introduced to East Asia from a single source,"⁷⁸ possibly by "a single wave of cultural impact" -- perhaps all from a Proto-Altaic **mVrV*, which proposition is dismissed by Janhunen (1998: 424).

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. *meorh*, Engl. *mare*, O. High Germ. *marah*, fem. *meriha*, *Mähre*, cf. *mar-shall*. Due to their isolation, Janhunen (1998: 425 sq.) does not want to connect western IE **mar(ko)-* with the Central Asian **morī*, 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" would appear to 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,⁷⁹ 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 (Raulwing 2000). Genetic analysis⁸⁰ seems to point to various centers of origin of the domesticated horse within the steppe belt, just as the various reconstructed words that designate it. From these areas, the horse has been secondarily introduced into the Near East, South and South-East Asia (see Meadow 1998).

Finally, 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 the 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-* "axe" > Greek *pélekus*, Ved. *paraśu-* (< Ilr. **paraću*, cf. CDIAL 7947 *parśu*), O.P. **paraθu* < W.OIr. **parasu* < Ilr. **paraću*. However, there also is a variant PIran. **paratu* > Osset. *færæt*.⁸¹

⁷⁷ 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".

⁷⁸ Similarly, Benedict (1972: 189) thinks of a Central Asian form (***mərən*) as seen in Mong. *mörin*, which may also have resulted in Tib.-Burm. **mrañ, srañ*, from a bisyllabic word **m[rañ]*. -- Note that Mong. *-in* in *mörin* is unexplained from the point of view of Altaic.

⁷⁹ 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).

⁸⁰ The domesticated horse seems to have several (steppe) maternal DNA lines, see *Science* 291, 2001, 474-477; *Science* 291, 2001, 412; cf. *Conservation Genetics* 1, 2000, 341-355.

⁸¹ Loan word in PPerm., Votyak etc. *purt*, Tunguse *purta* "knife", see Rédei 1986: 76. Note that the often compared Akkadian *pilaqu*, *pilakku* does not mean "axe" but "spindle, dagger." (EWA, s.v. *paraśu*). J. Colarusso (pers. comm.) points out that **paluqa* "may form part of a complex with a root **bel/*pel*, as in

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 changes in meaning between "stone" and "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* (or the Slavic *Perun'*).

All these sporadically available terms and names indicate close contacts between the regions east and west of the Pamir-Tian Shan-Altai ranges, but do not allow us, 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 still later, Turkic,⁸² can supply some of the desired data.

First of all, IE or rather Ilr. influence is seen in some of the Caucasus languages,⁸³ some of which may have retreated from the plains into the mountains, just as the N. Iranian (Saka) language Ossete did more than a thousand years ago. Such influence can be seen in Udi *ek'* "horse" < PIE or, pre-PIIr. (*h₁*)*ek'wo-*. From PIIr. *ac̥wa* (or inherited from Proto-N. Caucasian **h₁[n]čī / *ič̥we*) derive Avar-Andian **ič̥wa* > Andi *ica*, Avar *ču*, Lak *č̣wu*, Akhvakh *ič̣wa*, Chamalal *iša* "mare", Lezgin *šiw* "mare", and NW Caucasian: **č̣'wə* > Akhas *a-č̣ə*, Kabardin *šə*, Ubych *č̣'ə* (Blažek 1998), and Dagi *urci*, Khinalug *p̣ši*; as for PNC **ič̥we* cf. also Kartvelian: Georgian *acu'ca* (*acu*, to urge on horses), *ačua* "horse" (nursery word).⁸⁴ 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 Kabardian */g̊əwə/* "bull ~ PIE **g̊^hou* "cow".⁸⁵

However, the case is different when it comes to the designation of the old domesticated animal, the goat. Domestication has begun in the mountains of the fertile crescent (Zagros Mts.), from where it spread at the end of the Younger Dryas (c. 9,600

Russian *bol'soi*, Greek *Pelastike*, *Pelasgioi*, perhaps also *pēleús*/Peleus, Latin *(de)bilis*, Keltic *Belgae*, Welsh *balch*, and which may also occur in the wanderwort *pelewan* "strong man, hero," all related in some way to "might, power, big."

⁸² Cf. Nichols 1997, 1998.

⁸³ Note the materials supplied by J. Nichols 1997: 125-129, 143.

⁸⁴ Cf. the materials in Harmatta 1992: 369. See now additional discussion above, n. 67 (J. Colarusso).

⁸⁵ J. Colarusso, pers. comm., adds: "The Kabardian */g̊əwə/* (W. Circass. */čə'/* "bull, ox") clearly is an IE borrowing, but not without its mystery to its vowels." -- However, Chechen *gowr*, Ingush *gawr*, Nakh *govr* "horse" ~ Ilr. **gāura*, Ved. *gaura* "whitish-yellowish-reddish, wild ass, wild buffalo"; Iran. **gaura* > N.Pers. *gōr* "wild ass"; cf. also Pashto *yyara* "wild ass" (Morgenstierne 1927, no. 337), which must be kept separately. A similar development might have occurred with Ilr., 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 *xəɹ* "muddy, turbid, dirty brown", etc. Cf. also Rau 1980/81.

BCE). It is often assumed that Ilr. **āja* was loaned into Circassian *ača*, or that P-Iran. **aza* > Kabardian *aza* "goat for breeding". However, the etymology of Ilr. *āja* is not clear at all. Usually it is compared to Greek *haik'-s*, *haig'-*, itself without clear etymology.⁸⁶ 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 Macro-Caucasian. This by now firmly established language family (Bengtson 2001) includes the former language isolates Basque, North Caucasian (NE Caucasian = Nakh-Daghestanian;⁸⁷ NW Caucasian = Abkhas-Adygian), and Burushaski in northernmost Pakistan.

It is Macro-Caucasian that seems to have delivered to IE languages the word for the "goat". Apparently the word has been taken over several times, and in varying Macro-Caucasian dialect forms. Proto-North Caucasian has **ʔjZ'wē* (Adyghe *āča*, Dargwa/Akushi *ʔəža*, Chirag *ʔəč:a*, etc., Bengtson 2001), and Burushaski has *acás*, both of which are closest to PIE **Hag*' (Indo-Iranian *āja*- Skt. *aja*-, Avestan *aza*-, Lithuanian *ožys*, Latvian *āzis*).

However, there are other, divergent IE forms that must be reconstructed for PIE or for the stage of IE immediately preceding the forms actually attested in its daughter languages. These reconstructed forms indicate typical dialect divergences in the source language(s) or the intermediary (now lost) languages that have transmitted these words into IE. There is PIE (or late Common IE) **Haig-* in Gr. *haig-*, *haiks*; PIE **ghaid-* in Germanic (Engl. *goat*) and in Italic (Latin *haedus*); PIE **kag*' in Slav. *koza*; and finally PIE: **sk'ag* in Indo-Iranian **ścāga/ścaga-* (Ved. *chāga-*, Ossete *sæg(æ)* "goat", with a further loan into Uralic: Mordwinian *šava*, *šeja*, see EWA I 558); this particular form is apparently related to Proto-North Caucasian **ʔikV* / **kʔZV* (Karata *c'ik'er* "kid", Lak *c'uku* "goat"), and to be connected further with Burushaski *çigir*, *çigir*, *çhiir*, Basque *zik(h)iro* "castrated goat".

To the same substrate may belong Engl. *buck*, attested in Celtic and Germanic, in Avestan (*būza* "male goat") and in Proto-N. Caucasian PEC **b[a]c'V* (Bengtson 2001): Lak *buxca* < **buc-xa*? "young he-goat", Rutul *bac'i* "small sheep", Khinalug *bac'iz* "kid", etc.; cf. also Nakh **b'ok*' "male goat" (Nichols 1997: 128) and Burushaski

⁸⁶ See EWA s.v. *aja* for correspondences. 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*, etc. (below). -- Again, J. Colarusso (pers. comm.) adds: "The word for "billygoat" in W.Circass. is /*ače*/, Kabard. /*aže*/, Abkhaz-Abaza /*abáʔ*/, Ubykh /*p'q'ýáya*/. The Circ. word for "goat" generically is /*pč'ənə*/, Kab. /*bženə*/, Abaza /*ž, yma*/ (but Ubykh /*st'á*/ related to Circ. /*t'a*/ "two year old goat"), which suggests a link to Circ. /*č'ame*/ (actually, the /*a*/ is predictable) "cattle", perhaps PNWC **čima* > *p-čəma* > **p-čəna* > *pč'ane*, (where **p-/* is a grammatical class marker). I don't see the usual links between Ilr and any of the Caucasian material, except for my claimns in my "horse" paper. - The NEC words for "goat" are again varied, but point to PNEC **c'e-r/n-*, *(*a*)*č(a)-*, **č'(a)-* > *č'a-* since /*i*/ is facultatively glottalized in NEC. The word for "kid" in NWC is varied, W.Circ. /*č'ec'ə*/ Kab. /*k'əc'*/, (PCirc. **k'ec'e*/, Ub. /*st'á-š*/, Abkhaz /*a-žəs*/. In NEC the forms /*c'ik'*/, /*čik'*/, /*čepuš*/ occur (no metathesis, as is the case with some bisyllabic roots), suggesting links to the roots **c'e-* or *(*a*)*č(a)-*. No PNC form is recoverable. In Basque "z" is /*s*/, so Basque *zik(h)ir* is /*sik'ir*/. Even so, this resembles the NEC form.

⁸⁷ Subdivided into Avar-Andic, Tsezic, Lak-Dargwa, Lezgian.

buc, which appear as a loans in the Sanskrit substrate word **bokka* (Turner *CDIAL* 9312), Skt. *bukka* (Nepali *boko*, etc.).

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*).⁸⁸ Its southward 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 **gho-* and on the Iranian plateau⁸⁹ with a later, suffixed form **gant-um-*, and it is first seen in S. Asia in E. Baluchistan (Mehrgarh) at c. 6500 BCE.⁹⁰ Its progress beyond the Indus valley was held up for some 3000 years (see n. 127; a detailed discussion of the word for "wheat" is given below, §3.3, 5).

It is also 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/Hupanis; *Sindikē* area, see n. 180), *Sindes* in Turkmenistan (R. Tedzhen), E. Iran (Avest. *həndu*) and in the Indus valley: Bur. *Sinda* / Skt. *Sindhu* (see below, §4).⁹¹ Tuite (1998: 449) has indeed given some ethnological and linguistic

⁸⁸ According to Bengtson (2001): Bur. (H,N,Y) *gur*, *gurgán* (H,N) "autumn wheat", Cauc.: Tindi *q'ëru*, Archi *qoqol*, etc. < PEC **Gōl'e*, Basque *gari* "wheat" (combinatory form *gal-*). Here belong also J. Nichols' (1997-8) other words of culture such as **wino* "wine", etc. (Georgian *ywini* suggests that the culture word for "wine" started with a uvular, J. Colarusso, pers. comm.).

⁸⁹ For the beginnings of agriculture in Southern Turkmenistan (Kopet Dag, 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.

⁹⁰ 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; cf. n. 127.

⁹¹ The following list of NE Caucasian designations should be counterchecked against C. Asian Iranian and NW-Indian place names. They follow Colarusso (in Mair 1998), who based them on the work of Kibrik, A. E. and S.V. Kodzasov 1990; cf. the E. Cauc. reconstructions of Bokarev (1981). I have compared them with Bengtson 1999 and with his PEC forms which, according to his private communication of March 2002 (henceforth B. in this list), in part follow Starostin's PNC reconstructions of 1994: **tTem-* / **tTu-*, "water" (with tense voiceless lateral fricative [ɬ], secondary from **x-*, PEC **xānñi*, B. 2002); **sTor-* "water", (~ PEC **sVrV* "river, lake", Bur. *du-šór* "to melt", B.; cf. Syr Darya); **Tenc'o-* "river" (cf. *Sindes*, *Sindhu*, Bur. *sende*); **xTul-* "river" (PEC **hwǎ-IV*, **wirǎ* "river", Bur. *hur* "water conduit", Yen. *hur* "water" Bengtson 1999: 49; cf. Nuristani, Dardic *-gol*, *-gul*, *-gal* in river names, further (?): W. Nepali *-gād*); **ōr*, 'ar "lake" (PEC "lake, pond", B.); **rēto-* "sea" (cf. PNC **jārλw/λä(j)rǎ* "sea," B.); **myērō-* / **muγrō-* "mountain" (PEC **muhalV* "mountain," B.; cf. Late Vedic, Class. Skt. *Meru* "central, world mountain", NIA Dardic *Tirīč Mīr* "central mountain of Chitral", **Devameru* > Shina *diāmer* "Nanga Parbat" *CDIAL* 6533, Bur. *mərt* "earth cliff"?); **qTunt'u-* / **qTunt'o-* "hill" (PEC **Gwint* "mound, hill," B.); **arcTi*, "land" (PEC **jōmēV* "earth," B.); **naq'o-* "land" (PEC **neqwi* "earth, dirt," B.). Colarusso (1998) mentions as criteria for substrate words: ethnonyms, non-native vocabulary, oronymy, and hydronymy are the most conservative toponymy in surviving languages. He gives toponyms and hydronyms from three Caucasian families (with no apparent C. Asian connections except for NEC/PNC); cf. also Murzaev 1980.

reasons for a possible gradual movement of speakers of Proto-Burushaski eastwards towards the Pamirs.⁹² The NEC *š^Tor-, PEC *š^{Vr}V "water" is reminiscent of (Gr.) *Silis* "Syr Darya", Bur. *du-šór* "to melt" ~ -šór as a river name in the Murghab area of N. Afghanistan, and *ʎ^Tenc'o- "river" of (Gr.) *Sind-es*, Ilr. **Sindhu*, and Bur. *sende*. I leave the decision to specialists of Macro-Caucasian.

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

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)⁹³ and a still earlier one of pre-PIE south of the Caspian.⁹⁴ These locations are not born out, and are in fact contradicted by the host of words discussed in this paper.⁹⁵ 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 Ilr. In other words, they belong to the language of the BMAC area (e.g., place names such as *Xnənta*, *Baxδi(?)*, *Suyda(?)*, or etyma such as *uštr* "camel," *anču* "Soma," etc.) Further, agriculture in this area is not as early as she seems to assume: its arrival in the BMAC of Bactria only at c. 2400 BCE is much too late to have influenced PIE. That Bactria/Sogdia could be the locus of PIE therefore is at the least very doubtful, if not simply impossible. If the localization were indeed correct, *all* IE languages should have received the same "BMAC" substrate words that are typical for Old Iranian and Old Indo-Aryan. (The same argument destroys the revisionist and "autochthonous" Indian fantasy, the so-called "Out of India theory" of PIE, see Witzel 2001a).

He adds (pers. comm.): I would collapse the cognates down to one original stem *ʃ^{em}- and take the forms as mid-Daghestani (Lak-Dargwa, and Lezgian) shifts, with some of the southern languages showing velar or even uvular reflexes of this original tensed lateral spirant. A typologically driven shift of *^m- to -ⁿ- would open the door to *n* ~ *r* alternation (as in the BMAC language), and in some the *-^r- shifted to /-^d-/. The use of superscript "T" [in the 1998 paper] to denote tenseness seems to have arisen from some font mismatch. The usual representation for "river" in PNEC would be *ʃ^{en}čo or *kʃ^{en}čo*, with a ligature under the initial cluster. -- Starostin's assumption that *^x- is primary in NEC "water" is not supported by the development of velars or uvulars in these languages. I also find no evidence to support his form for "river." Apart from *ʃ^{en}čo, there are two other roots *^{ork}'u (from Avar and Dargwa evidence), and *^{onq}'o- from Khinalug and Udi forms). The match for "mountain" is very good. I see no reason to alter my original PNEC form for "hill" *q̃^{untu}-/*q̃^{onto}*-, except to see pharyngealized uvulars as secondary developments of the tensed *^k' under conditions of tone (an imponderable, really). But "Guti" is clearly a southern Daghestani (Macro-Caucasian) word for "Hill (People)."

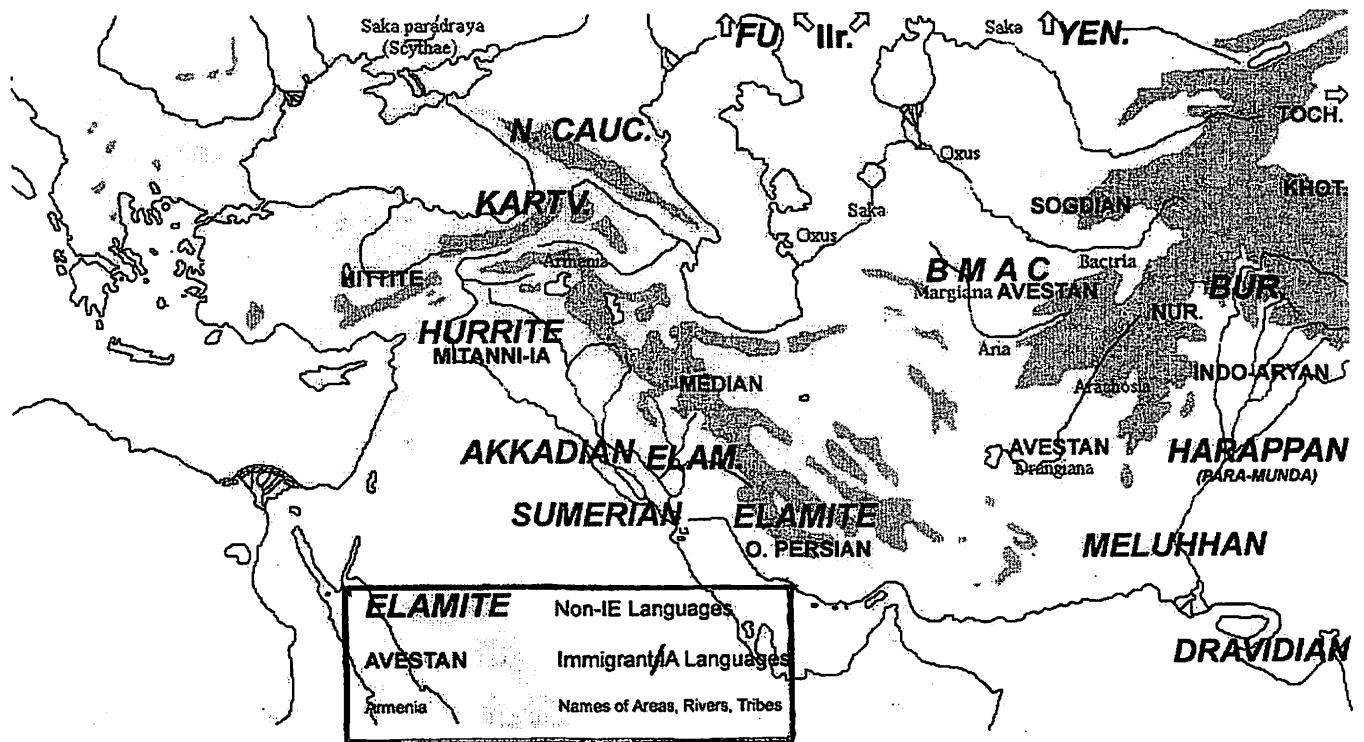
⁹² In fact, it has been suggested that the population of Gorgan and S. Turkmenistan (Kara-Kamar) "reflects the spread of Mesolithic people from the Zagros mountains to the northern foothills of the Hindu Kush via the Caspian coast." (Sarianidi 1992: 124).

⁹³ "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).

⁹⁴ Viz., north of the Black Sea. Certainly not in the clearly non-Ilr. Mazenderan, as her map seems to indicate.

⁹⁵ Note also the early criticism, in part based on prepublication materials, by Mallory 1998.

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 PU and PIE are located (PIE **wed+r/n-*, PFU **wete* "water", etc. (cf. Nichols, 1997: 146).⁹⁶



Distribution of languages, c. 2000-500 BCE

⁹⁶ 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). In contrast, 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, some 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; Witzel 1995, 1999a,b,c). 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 Yeneseian speaking peoples. We know that, in this civilization, the domesticated camel was used,⁹⁷ that it continued the large scale use of unburned 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. will hardly have moved into this complex earlier than the introduction of donkeys.⁹⁸ Pinault (2003) shows that the word for "brick", *išt(i)*, has also been taken over into early (Common) Tocharian; (cf. further below, *passim*, on *anču*, *čarwa*, *paŋi*, *aŋi*, *athr*).

However, as will be seen, some of the IIr. loan words have been taken over independently of each other, twice or thrice into various IIr. languages in different areas of Greater Iran (cf. Lubotsky 2001: 302 sq.), 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, 2000a, Lubotsky 2001) -- 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 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 isolated⁹⁹ language Elamite was spoken, not just in the Susiana (Khuzistan) and *Anšan* (Fars and surroundings), but also in

⁹⁷ By 2500 BCE there are camels (figurines) drawing vehicles in S. Turkmenistan.

⁹⁸ They could, however, have learnt about camels from the Central Asian hunters, as represented by the Kelteminar culture.

⁹⁹ A genetic link between Elamite and Dravidian has not been established, see the extensive discussion in *Current Anthropology* (McAlpin et al., 1975).

Southern and Central Iran,¹⁰⁰ in the areas of Tepe Yahya (*Simaški*),¹⁰¹ and Shahdad (*Tukriš*),¹⁰² but apparently not farther east than Bampur (*Marhaši*).¹⁰³

According to F. Vallat (1985: 52, cf. 1993) the language of *Marhaši*¹⁰⁴ differs from that of *Simaški*, and is only very partially Elamite related. Here and in Meluhha (Baluchistan, Sindh)¹⁰⁵ the language changes, though Elamite cultural influence extends even to the borders of Sindh. However, eastern names are neither Elamite nor Dravidian.¹⁰⁶ 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 (sufficiently) different from Elamite or Sumerian is obvious: the Mesopotamians needed a "translator from Meluhha" (Possehl 1996: no. 2), whose name was reported as *Šu-ilīšu* (Parpola 1994: 132).¹⁰⁷

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.¹⁰⁸ 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

¹⁰⁰ The insurrection of Nāram-Sīn (text from the end of 3rd mill. BCE) mentions the rulers allied against Akkade: "the man of Meluhha, the man of Aratta, the king of Marhaši, ... (another country : gap in tablet), the king of all of Elam" (Vallat 1995: 53); note also the list by the Sargonic king Rimus who conquered Parahšum, Zahar, Elam, [Ba]sin(?), and Meluhha (see Gelb 1997: 594).

¹⁰¹ Perhaps one of the Elamite capitals.

¹⁰² 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 *Tuṣṣrān*, *Tuṣṣrastān*, *Tokharoi*, etc. (Henning 1978), and with a possible Mitanni-IA name *Tugra*, see EWA I 651 s.v. *tugra*. Some RV passages involving *tugr-* may provide a link with this Central Iranian area as well. Cf. also Gamkrelidze and Ivanov 1989 who equate the Near Eastern Guti and *Tukriš* with the later "Tokharians" (Kuchaeans).

¹⁰³ For various Elamite sources see Vallat 1985, 1993. However, Steinkeller (1982: 255, map p. 265) tentatively locates it a little farther west, "in the perimeter of Kerman and Eastern Fars." The original name of *Maḥarši* seems to have been **m̥parahši*, represented in Akkadian as *Parahšum* (Steinkeller 1982: 237-8). Cf. the modern *Maškai* and other rivers in Baluchistan beginning with *Maš-*?

¹⁰⁴ Steinkeller 1982, 1989 (cf. also Hiebert 1998: 147). Steinkeller connects the name of *Marhaši* /*Parahšum*/ *Paraši* with the name of the Persians; for this, however, note the *Paršu* reference in RV, BSS (Witzel 1999c) and a possible connection with mod. Pashto, see below, n. 201. For further E. Iranian/Vedic correspondences see Witzel 1989, ch.10, and for some Saka-like characteristics of O.P. see n. 202.

¹⁰⁵ See Possehl 1996.

¹⁰⁶ Which speaks against Drav. as the language of the Indus Civilization, see further Witzel 1999a,b, 2001b for other possibilities. -- Note, however, Gurov and Vasil'kov (1995) on a Drav. etymology of Aratta.

¹⁰⁷ The loan word links between Sumerian and Drav. are in further need of investigation (see Blažek and Boisson 1992 and Blažek, 1992, 2002a). 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 area of Elamite (from Susa up to Tepe Yahya) and by the separate linguistic area of *Simaški* (Bampur). Following the observations made in Witzel 1999a,b, I suggest that early Dravidian may have received its agricultural terminology via Aratta (Sistan) or even directly by maritime contacts.

¹⁰⁸ Possehl 1997, with a date of 2500-1900 BCE for the Kulli/Quetta phases.

areas east of it, the Elamite areas west of it, 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 (see §6).

While the southern belt of Greater Iran thus is excluded, the 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 mountains starting around the end of the 3rd millennium. First came the Lullubi and Gutí, in the time of Naram Sín of Akkade (c. 2250 BCE). On the Iranian highlands, the Simaški state of the 3rd mill. was replaced by the Tukriš state in the second millennium, which is connected, in some texts, with the Gutí.¹⁰⁹ 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) contain some OIA linguistic elements.¹¹⁰ However, we know much less about Media (the area around Tehran, Hamadan), 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).¹¹¹

Indeed, **anću* "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 2001, Witzel forthc. b), where according to both the Avesta and the R̥gveda the best Soma grows (Ved. *Mūja*-vant mountain, Avest. *Muza*, mod. *Muzh Tagh Ata* in the Pamirs). There are, indeed, some indications of non-Ir. speakers in the high mountains of Afghanistan even at the time of the earliest Ir. texts. Such "foreigners" indeed still survive¹¹² in the Pamirs as the Burushaski speakers and are visible in the substrate in Khowar (Witzel 1999a,b, 2000a).

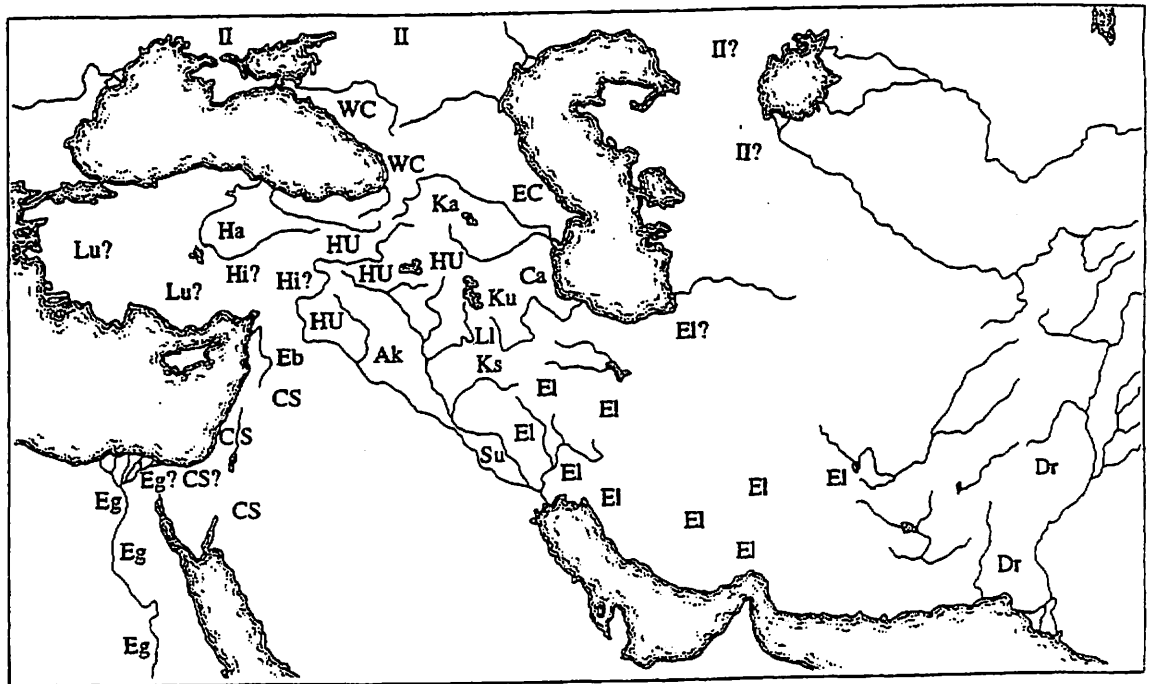
¹⁰⁹ Hammurabi inscription, see Henning 1978, 220. Perhaps Gutí means "hill (people)", from N. Caucasian **k:unt'i* (J. Colarusso, pers. comm.).

¹¹⁰ Needless to say, this fragmentary information will have to be compared with the Central Asian materials presented here. At first sight, I have not detected correspondences beyond the words for "jewel", Mit. *mani-nnu*, Ved. *mañi* (see below), Soma (Mit. *Sauma*-, EWA II 749), and the Central Iranian name *Tukriš*, possibly retained by Mit. IA **Tugra*, above n. 102.

¹¹¹ For Margiana etc., cf. also Erdosy 1998.

¹¹² Note that the central Afghan highlands have been occupied by Mongolians (Hazara) after the expansion of the Mongol empire -- wiping out possible traces of older populations; for some details on the Central Highlands see Witzel 2000a; cf. the similar situation in the Tian Shan/Pamirs, where Mt. Muzh [Tagh Ata] of our maps is derived from the old, non-Ir. name Ved. *Mūja*-, Avest. *Muža*, but was re-interpreted as a Turkic name. The impressive, Kailash-like 24,767 ft./8.255 m high *Muzh Tagh Ata* means "ice mountain father" in Kyrgyz. Another *Muzh Tagh* is found northeast of Skardu in northernmost Kashmir.

Abbreviations	
Ak	Akkadian
Ca	Caspian
CS	Central Semitic
Dr	Dravidian
Eb	Eblaite
EC	East Caucasian
Eg	Egyptian
El	Elamite
Ha	Hattic
Hi	Hittite
HU	Hurro Urartean
II	Indo-Iranian
Ka	Kartvelian
Ks	Kassite
Ku	Kutian
Ll	Lulubean
Lu	Luivian
Su	Sumerian
WC	West Caucasian



Distribution of languages
 (from: Blažek 1999: 53)

§ 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 2001).¹¹³

¹¹³ I cannot enter here a detailed discussion of Lubotsky's recent paper (2001). 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/stāpa* "tuft", **pāpa* "evil", **kaśyapa* "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 (first yellowish milk of a cow after it has given birth)", **warāḥja* "wild boar" (but note the early loan into FU: F-Volg. **oraše*, Finn. *oras*, Rédei 1986: 54!). -- He had originally assumed (as per van Driem 1999) 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 (or, as he now formulates, "intimately related," Lubotsky 2001: 306; however, see Witzel 1999a,b and forthc. a). This close connection is, however, unlikely and does not account for loan words with the suffix *-śā*, *-śa*, *-śa* 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 *śardāla* "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 us to determine what is securely

Shibboleths for time and place are, as mentioned, the local words **uštr* "camel", **khara* "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. / c. 2000 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 this case Mesopotamia, where an Akkadian (Mari) source has *ḫārum*, *ajarum* "male donkey" (EWA I, 447, 473; cf. perhaps also Kushitic **dəq^War*, Elamite /*dranku?*/ Blažek 1999: 64). Various types of donkeys may be distinguished in Vedic India,¹¹⁴ 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 *kercao*¹¹⁵ :: PDrav. **gar-(?)* > Tamil *kaḷutai* "donkey".¹¹⁶ 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 2002a).

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, Lubotsky 2001: 311). Ved. *iṣṭi*, *iṣṭakā*, *iṣṭikā*, Avest. *ištiia* "brick", *zəmō-ištuua* "clay

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 Avestan, is the *only* secure basis for analysis, since 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 (such as Avest. *araēka*, *Ainīāuua*, *Aiēhīā*, *ayažāna*, etc. see n. 158). Lubotsky so far neglects the typical substrate words found *only* in Iranian such as *kaufa* "hump, mountain" or *tātuk* "clay" (see n. 158) which can serve as a corrective. -- Lubotsky's new data, however, support my initial analysis of Central Asian loans (Witzel 1995, 1999a,b,c), though he does not localize this substrate beyond pointing to the towns of the BMAC and to the Panjab (echoed by van Driem in MT, Special Issue, Oct. 1999; for a discussion and refutation see Witzel forthc. a).

¹¹⁴ For instance *khara*, *rāsabha*, *gardabha*, etc. see Rau 1980-81.

¹¹⁵ Both with the common Indian animal suffix *-bha* (?) as in *garda-bha* (~ *grda* "penis"?), *sara-bha* which is preserved in the northwestern NIA Dardic Kalasha language as *šāra* "markhor mountain goat", *ṛṣa-bha*, which would point to a slightly different source.

¹¹⁶ The overlap of Dravidian *kaḷu-* 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. *kaḷu-tai* DEDR 1364) could, or rather should, be independent loans from a language of Greater Iran or Central Asia. - Some want to make the BMAC a Dravidian speaking area (cf. Lamberg-Karlovsky 2002: 73), however, the close loan relationships between Drav. and Sumerian and Elamite point to a more western trail; this is confirmed by the lack of any Drav. words among the BMAC loan words found in OIA and OIr.

brick", O.Pers. *išti*, M.P., N.Pers. *x-išt* > Pashto *xašta*? (Morgenstierne 1927: 98), Baluchi *išt*. One also may compare Toch. B *išcem* (or *išcem*) "a kind of clay", B *išcake* "clay". Interestingly, this goes back, as in Indo-Iranian, to an *i*-stem **išti-* > Toch. A **išce*, A **išāc*, borrowed into Uighur as *išč* "earthen cooking vessel" (Pinault 2003, cf. Pinault 2002: 326 sqq.). Finally, we may add, unnoticed so far, some words from south of the Hindukush and Pamir mountains: Bur. *d-iščik*, Shina *d-ištik*, Kalash *kh-išti-poktá* "brick", and, surprisingly, even Marathi *v-iṭ* "brick" (instead of expected **iṭ*). While all of these words lead back to a C. Asian source for "(clay) brick", **išt(i)*, the South Asian words have unclear pre-Indic "prefixes" *d-*, *kh-*, *v-* (M.P. *x-* is secondary).

Since there are no brick buildings at the northern rim of C. Asia¹¹⁷ 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 Uralic and Yeneseian, would first encounter bricks and would need a term for them. As discussed above, the domesticated camel was used here, and the donkey had been introduced from the Near East at the time. These three *leitfossils* provide the time frame.¹¹⁸ This 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. *raṭa*, Mitanni *Tuš-ratta*, Kassite *Abi-rattaš*) and related words (**rathin*, *rathī/raṭī* "chariot driver", and the archaic compound word *rathesṭhā/raṭaēšta* "chariot warrior").¹¹⁹ All of this leads to a successful triangulation of material culture, time, and space in W. Central Asia at c. 2000 BCE or somewhat 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 Caucasus¹²⁰ 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

¹¹⁷ For example, in Sintashta there are only reinforced palisades with clay inside; however, the area just north of the Caucasus reportedly has clay bricks, see Witzel 2000a.

¹¹⁸ 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). For the passing through the BMAC area by IA speakers note also a few words transmitted in Mitanni IA: *mani-nnu*, *Sauma*, *Tukriš* as **Tugra* (above, n. 110, 102).

¹¹⁹ Such as "chariot warrior (bow shooter), spoke, fellow", etc. (Raulwing 2000); (pre-)chariots (**ratha*, not **anas* "wagon") are first found in the northern steppes at Sintashta and also in Near East about 2000 BCE, see Witzel 2000a.

¹²⁰ See the recent summary on Ofer Bar Yosef's work, at <http://www.harvard-magazine.com/online/09016.html>.

Ved. *godhūma*, Drav. (Kan.) *gōdi*, Bur. *guriñ*, *gureñ* (pl.), *yárum*; all words are derived from W. Asia.

Ved. *godhūma*, Avest. *gañtuma* must go back, with folk etymology on the Indian side (see below)¹²¹ to a common Ilr. < **gant-um*,¹²² which in turn, echoes a Near Eastern source, Semit. **hnt* (Arab. *hintatum*), Hitt. *kant*, Egypt. *xnd* (EWA II 499). On the other hand, a Caucasian/Bur./Basque (Macro-Caucasian) source is seen in: Bur. *gur* "wheat"¹²³ ~ Basque *gari* "wheat", Proto-East Caucasian (Daghestani, etc.) **Gōl'e*, PKartvelian (Georgian *yomu* "millet", etc.) **ghomu* (see Witzel 1999a,b,c). Obviously the ultimate Near Eastern source for all these words must have been something like ***gər* // *q/gən-d* (ə = uncertain vowel; for the variation of *r/n* see §5).¹²⁴ These relationships, along with those pointed out above for "goat" and "buck", and those discussed by J. Nichols (1997-8) establish an early, widespread network of cultural interactions between the populations of the Fertile Crescent, the Caucasus, the steppes, the Urals, Iran/India and the speakers of Macro-Caucasian (including Basque and Burushaski). Such relationships will be pursued elsewhere.

The Ilr. source, **gant+um-a*, differs from its Near Eastern source ***gənd-* by a suffix *-um* which is attested not only in Iranian (Avest.) *gañtuma*¹²⁵ and Vedic *godhūma*¹²⁶ but also in Burushaski **yond-um* (and perhaps in P.Kartvelian (Georgian, etc.) **ghomu*, cf. Proto-East Caucasian (Daghestani, etc.) **Gōl'e* (see Witzel 1999a,b,c).

¹²¹ 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, n. 128).

¹²² 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. **yund-um*) may therefore have originally meant "wheat grain". (Cf. also the link between PDrav. **var* "seed, grain", Elam. *bar* "seed", Southworth, 1988: 659-660.)

¹²³ Bur. *guriñ*, *gureñ* (pl.), *-yárum* < **yor-um* < ***yund-* (Berger), ***yund-um*; cf. also Bur. *gur* "barley, wheat colored", *bur* "buck wheat," Berger 1959: 43. -- J. Colarusso (pers. comm.) adds the following N. Caucasian forms. "For "wheat" note the W. Circass. /*k'ec'ə*/, Kabard. /*g'ec'*/. The words for "wheat" in NEC are (capital L = velar lateral fricative): Avar *roL-*, Andi *muGa*, Akhvakh *q'iru-*, Chamalal *q'ew*, *cibāL'*, *q'erū-*, *cibū-rēL'u*, Tindi *q'eru*, *cibāL'ab*, Inkhokhvari *at'*, Tsez, Hunikh *at'*, Bezhta *q'ibo*, Lak *lač'a*, Archi *qoq(ó)l-*, *sot-*, Tabasaran *daxín-i*, *ž' aǝn-i*, Aghul *"ak'*, *ek'y*, *ek'én* (" = ejective pharyngeal), Lezgi *q'ul*, *næx̣̌* (/u/ fronted), Rutul *ym̩l*, *q'ir*, *nax̣̌*, Tsakhur *suk*, Kryz *Gul*, Budukh *Gul*, Khinalug *li*, *l-i*, *lok'-i*, Udi *arum*. The voiced uvular stop proposed by Bengtson is the Kryz-Budukh development of **q'*. The words are clearly borrowings (with perhaps a few native forms for earlier grains, such as Tsezic **at'*), but from an unknown source. There are two originals: **q'uli-*, **qol-*, and **lok'-i*, with assimilated variant **loL'i*. Note how final /-kV/ syllable can lateralize, Russ. *babka*, Archi /*babL'*/."

¹²⁴ See below, §5, on the interchange of *r/n*.

¹²⁵ Avest. *gañtuma*, M.P., N.P. *gandum*, Shughni *žindam*, Pashto *yanəm* < **gandūma?*, Khot. Saka *ganama* < **gandama* (Berger 1959: 40f, EWA II 498); however, Brahui *xōlum* < IA **yolum* (CDIAL 4287).

¹²⁶ As seen in Nur. *gām*, 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 *gandār* "a kind of grain". The actual Ved. form *go-dhūma* must be due to local influence by the Southern (Meluhhan) **gōdi* (see above n. 121, on Konda *goyi*) on a northern **gantum/gandum*; it may be due to Dravidian influence on the Panjab in the Middle/Late Rgvedic period as *godhūma* appears only in early post-RV texts.

One may assume a BMAC word **gant-um* that has entered Iran and India via the northern Iranian trade route (Media-Turkmenistan-Margiana/Bactria-Aratta/Sistan), while the forms with *go-* (PECauc. **Gōl'e*, PKartv. **ghomu*), which are reflected in Dravidian, must have come via the Southern route (Elam/Anšan - Simaški/Tepe Yahya - Marhaši/Bampur), resulting in Drav. **gōdi* (Kan. *gōdi*, Tam. *kōti*, cf. DEDR 1906).¹²⁷ The post-RV change from *-an-* > *-o-* is untypical for the Panjab but found in Sindh (Witzel 1999a,b). It must have influenced, in some way (such as Drav. influence on the late RV, Witzel 1999a,b), the actual form of Ved. *go-dhūma* as well. Note that the Drav. word, too, seems to be based on a popular etymology.¹²⁸

Interestingly, irrigation channels, the only means allowing of sustainable cultivation in these dry lands, are indicated by a loan word as well, **ya(u)vyā* "streamlet, channel".¹²⁹ Irrigation agriculture is said to have existed from 2200 BCE in W. Central Asia.¹³⁰ 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 *ži* "rivulet" (EWA II 405). Both forms cannot go back to exactly the same source; this kind of telling difference often is indicative of loan words.¹³¹ 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

¹²⁷ There is archaeological evidence now: while Renfrew had maintained an 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: 989, <http://www.sciencemag.org/cgi/content/full/294/5544/988>.

¹²⁸ 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 Elamite, Sumerian, etc. loans into Drav. will have to be compared, and must be reconciled with Blažek 2002a.

¹²⁹ See Dani 1992: 116, 222 etc.; note, later on, also the extensive network of underground channels in Afghanistan, *karēz*, etc.

¹³⁰ Adams 1998: 376, quoting F. Hiebert (1996, cf. 1998: 231); cf. however, Francfort 1999: 451.

¹³¹ It is interesting that the O.P. word is as close as it is 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. **naHwiya* "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.

items are indicative of the expanded agricultural base and the settlements of the BMAC. The words below are given here without the IIr. nominal stem suffix *-a-* as we do not know the original (stem?) form that they may have had in the BMAC area.

- **parš* "sheaf": Ved. *parša* "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 *biz* "grains";
- **sthūnā* "pillar": Ved. *sthūna*, Y.Avest. *stūnā*, *stunā*, O.P. *stūnā* "pillar"; unless ~ Ved. *sthūra* "tall, thick", Avest. *-stura*, Khot. *stura* (thus EWA II 768);
- **pind* "lump": Ved. *piṇḍa* "lump, ball (of food)" Khotan. *piṇḍaa*, Armen. *pind* "compact, firm" < Iran.;¹³²
- **ling* "mark" : Ved. *liṅga* "mark, penis" : Avest. *haptō-iriṅga* "the seven marks" = the seven stars of the Great Bear/Wain (*ursa maior*), s. EWA II 478 sq.;
- **vīna* "lute" : *vīṇa* "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;
- **mani* "jewel", Mitanni IA *manni-nnu*, Ved. *mañi*, Avest. *-ma'ni*, -- if not from IE, cf. Lat. *monile*, O.Eng. *mene*, etc., s. EWA II 293 sq.
- **āni* (*āṇi*?) "lynch pin", Ved. *āṇi* "lynch pin", part of leg above the knee" (Suśruta). Pinault (2003) connects Toch. B *oñiye*, Loc. Sg. *oñi-ne* < **āni-ēn* "hip" (with a productive Toch. suffix). The shift in meaning is explained by the common identification of vehicle parts and body parts (*ratha-mukha*, *ratha-śrīṣa* "head of a chariot", *nabhi* "navel, nave"): both sides ("hips") of a vehicle ~ sides of the body; Pinault also analyses *kalyāṇī*, *kalyāṇa* "beautiful" < *kali-āṇi* "having beautiful hips = *pṛthu-śroṇi*"). For the retroflex *-ṇ-* cf. *Paṇi/Parna* and see §5.
- **kapaut* "blue": Ved. *kapota* "pigeon", O.P. *kapauta* "blue"; Khot. *kavata* "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". Words for non-primary colors tend to be taken over as loan words (black, white, red, blue : orange, crimson, indigo, violet, etc.).
- Lubotsky 2001: 307 now adds: **aka* "bad", **karuš* "damaged (teeth)", **pāpa* "bad" (actually, "evil").
- **bhiṣ*, *bhiṣ-āj* "to heal, healer" : IIr. **bhiṣ-aj* > Ved. *bhiṣ-aj*, *bheṣaja* ~ Avest. *baēsaṣ-ia*; note the unusual noun formation, and the important combinations with **sauma* (Soma/Haoma), see EWA II 264.
- *sīs* the metal "lead", EWA 734 from southwest Iran. (O.P.) **siṣa* "white" which would be too early¹³³ 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 found¹³⁴ in the Kandahar area.¹³⁵

¹³² Note also *kubja*, *kubhra* "crooked", and N.P. *kāž*, *kaz*, Sogd. *kwzz* "humpbacked", Khot. *kāysa* "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.

¹³³ The Persians moved into the Persis and Anšan from NW Iran only after c. 700 BCE., see summary in Skjærvø 1995.

¹³⁴ Lead is found in the Kandahar area, see maps in Kenoyer 1995.

- Lubotsky (201: 307) now adds a list of items of water/irrigation **khā* 'well, source', *čāt* 'pit, well', **yavīya* 'canal'; of agriculture: **nagna* 'yeast, bread', **(s)phāra* 'plough-share'; of building technology: **išt(y)a* 'brick', *ḡharmiya* 'firm structure, permanent house', **mayūkha* 'wooden peg', **sikātā/čikatā* 'sand, gravel'; of artifacts: **kapāra* 'dish, bowl', **naij(s)* 'spit', **wači* 'axe, pointed knife' ('seems cognate with Circassian /wəʃʷə/ 'small axe for splitting rails,' J. Colarusso, pers. comm.), **gadā* 'club', and of clothing: **atka* 'cloak', **dača* 'hem, thread', **dṛca/dṛca* 'coarse garment', **pawasta* 'cloth', **sūči/čūči* 'needle'.

§ 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, Macro-Cauc. words such as Bur. *bañ*, *bañgi* 'hemp', Caucasian: Chechen *baga* 'pine tree', Lezgi *muk'-rag* 'fir tree,' etc. < PEC **bhīnk'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. *gæn*, *gæncæ*; note: Sumer. *kunibu*; Gr. *kánnabis*, Russ. Church Slav. *konoplja*, OHG *hennep* < Scythian **kanap*; Kirgiz, etc. *kāndir*, Turkish *kendir* > Hung. *kender*; Mordwin. *kañčt*, *kañf*; Cheremis *kəñe*; -- for the change between *k'/ś* in E. Afghanistan/Panjab, as in *Karkōṭa/Śarkōṭa*, see Witzel 1999a, b;
- ***sinšap* 'mustard': Ved. *saśarpa* 'mustard' (Brāhmaṇas) > MIA, NIA *sāsapa* 'mustard seed', Khot. *śśaśvāna*, Parth. *šyṣ-d'n*, Sodg. *šywšp-δn*, M.P. *span-dān*, N.P. *sipan-dān* 'mustard seed' (cf. Gr. *śnapi*), < pre-Iran. **sinšapa* < ***sinsap* (Henning's *s1ens2ap*, see KEWA s.v.)¹³⁶

§ 3.6. Local animals

Even more importantly, some names for local animals pinpoint the area of the "BMAC" substrate more clearly. One 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.

¹³⁵ Further, a large number of verbs can be added, such as **kan* 'to find pleasure, please', **kram* 'to stride', **kroc* 'to shout', **kōā* 'to regard, look', **kšad* 'to serve food', **kšam* 'to be patient', etc., or the culturally important **kaiš* 'to apportion magically, to teach', **bhiš* 'to heal' with its strange derivative **bhiš-aj* 'healer' (see Lubotsky 2001), perhaps **mark/marc* 'to damage', or nouns such as **kaufa* 'mountain, hump (of a camel)', etc.; see further, below n. 158.

¹³⁶ Also found in S.E. Asia: Malay *sawi*, *sasawi*, Austro-As. **sapi*, *sV(r)-sapi*; cf. further the similar word, EWA 712, 727: *śimśāpā* RV+ 'Dalbergia sissoo' N.P. *šīšam*, Pashto *šəwa* < **šīšampā*, CDIAL 12424, Elam. *še-iš-šā-ba-ut* = [*šeššap*].

- **kaśyap* "tortoise": Ved. *kaśyapa* KS+, YAvest. *kasiapa*, Sogd. *kyšph*, N.P. *kaśaf*, *kaś(a)p* "tortoise"; cf. *Kashaf* Rūd, river on the border of Turkmenistan and Khorasan (cf. Humbach 1984, 1991)

- Lubotsky (2001; 307) has: *(H)*uštra* 'camel, **khara* "donkey", **kaśyapa* 'tortoise', **kapauta* "pigeon", **jaja/uka* "hedgehog", **matsya* "fish", **mṛga* "game", **warāḥa* "wild boar" ("reflected in Ossetian Uryz(mæg), Circassian (Iranian loan) Warza(meg). The Nart name must reflect what Johannes Knobloch thought "wild boar". He would be the sole theriomorphic warrior relic in that tradition." J. Colarusso, pers. comm.).

The leopard was a common Central Asian animal that is often found in early C. Asian art (Dani 1992), in fact a prominent motif on the oldest pottery from W. Turkmenistan. Later on, it was used for sport (see Th. Allsen in V. Mair, forthc., van Binsbergen, forthc.). Its designations differ considerably in the various languages concerned.

- ***pard/pandh* "spotted animal, panther": Ved. *ṛḍaku* "snake" RV, *ṛḍakū* AV "panther snake" AV, *ṛḍākhu* BSS, W.Panj. *parṛā*, O.Iran. **pard-*, cf. Khovar *purdūm* < **ṛḍhuma?*, Bur. (Yasin) *phūrdum* "adder, snake", see EWA II 163, KEWA II 335, CDIAL 8362, Lubotsky 2001: 305 n. 3.

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) has pointed out that the AV snake *ṛḍaku* is called after the leopard.¹³⁷

The Iranian forms differ considerably: N.P. *palang* "leopard, *felis pardus*" < O.Iran. **pard-*,¹³⁸ Kurd. *pilink*, Pashto *ṛrang*, Parachi *parō:n*, cf. further Lahnda *parṛā*, Gr. *párdalis*, *párdos*, *léo-pardos* "leopard", all < **pard* "spotted, wild animal?" (see EWA II 133), or following Henning, from an older ***parḍ*. 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*.¹³⁹

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* (see now Lubotsky, 2001: 304, on the common BMAC suffix *-ka* added directly to roots, cf. Witzel 1999b: §15, on the word for "rhinoceros").

- ***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).

¹³⁷ Note that the use of **pard* (*ṛḍaku*) has been narrowed down in India to "[spotted] snake," while the 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 *ṛḍaku* has been influenced in form (and meaning: "reptile") by the local Indian *ṣṛḍaku/gu* MS, *ṣṛḍaku* "lizard", *ṣṛḍara* "snake" (Witzel 1999a,b).

¹³⁸ For the sound changes involved cf. Avest. *paridaēza* (source for our "paradise") > N.P. *pālēz* "garden."

¹³⁹ Greek *panth-* < substrate/loan word **pandh-*? Cf. §5, on the interchange of *r/n*; **paʾ/n* has been substituted in India early on by *dvīpin*, see above n. 137. Note that Steinkeller 1982: 253 considers the "speckled "dog" of Meluhḥa (*ur gūn-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 Ibbi-Sin, the last king of the Third Dynasty of Ur.

The rhinoceros is by definition a South Asian (Indus Civilization, etc.), not a Central Asian, animal. However, its prehistoric spread needs to be investigated. The local climate and water table around the Sistan lakes and the reed thickets of the Helmand river¹⁴⁰ would have allowed for its existence and survival 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 must be raised whether the word in question is a loan from S. Asia or a local word.¹⁴¹ 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. - Lubotsky (2001: 307) now adds terms of cattle breeding: **kšira* "milk", **pīyāša* "biestings", **ścāga/ścaga* "billy-goat" (see above, § 2.3.) and a list of body parts: **kapha* "mucus, phlegm", **kaića/gaića* "head hair" (cf. *Vāiti.gaēsa* "Bādḡīs Mt.", Witzel 1972), **kućši* "side of the body, flank", **gr̥da* "penis", **maļjha* "belly", **pušća* "tail", **stuka* "tuft of hair", **šwaipa* "tail", **wṛtka* "kidney".

§ 3.7. The religious sphere

Importantly, in addition to and beyond the items of material culture listed above, the religious sphere,¹⁴² 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: **anću* "Soma plant", **yātu* "black magic", **atharwan* "priest" (however see EWA I 60), **ṛši* "seer", **ućig* "sacrificing priest", **magha* "gift, offering, sacrifice", **ćarwa* "name of Rudra", **indra*, **g(h)andharw/b(h)a* "demi-god or demon".¹⁴³

¹⁴⁰ See Falk 1997 with relevant literature. The rhinoceros is mentioned in the Hanshu 96: 3889 A (completed 92 CE) as living the kingdom of Wuyi (Alexandria) in Afghanistan (see Behr, n.d.: 8); note that rhinoceros were still hunted by the Moghuls in the Peshawar valley, which had been heavily forested and hardly settled in prehistory, at least down to the Achaemenid period.

¹⁴¹ The tiger (along with the rhinoceros?), must 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 Kopet Dag Mountains. Its absence in the RV in contrast to its prominence in the seals of the Indus civilization, along with the R̥gvedic absence of the leopard (that has been found depicted from early Turkmenistan pottery onwards), may be due to the fact that it did not belong to the traditional imagery of the Indo-Aryans who preferred the lion (*siṃha*). A possible reason may be the preponderance of lion images in religion (and the absence of panther and tiger), a religious choice (Francfort 1994) already made by the people of the BMAC area through which the speakers of OIA must have passed, (cf. also W.J.M. van Binsbergen forthc., on the image of the leopard in Africa and Asia). A new study of early IIR. animal designations and their respective importance in religion (cf. W. Voigt 1937) is a desideratum. In view of these uncertainties with regard to the distribution of these animals, it remains of course entirely doubtful whether the languages of the BMAC and of Harappa were related at all as Lubotsky (2001: 306) believes (see above n. 113).

¹⁴² First collected and discussed, as far as represented in EWA, by Lubotsky 2001: 304; cf. now the overview of pre-Vedic religion in Witzel, forthc. b).

¹⁴³ The last three words may, however, be late loans into Vidēvdād, from OIA, in the late 1st. mill. BCE; but note Pinault on *ćarwa*, n. 152.

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 Central Asian substrate now allow us 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 **n-mrto-*); 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.¹⁴⁴ The *Soma* plant, whose botanical nature is still uncertain,¹⁴⁵ originally seems to have been called **anću* (Ved. *aṃśū*, Avest. *qsu*, see G. Thompson 2001, 2003,¹⁴⁶ Pinault 2003). As has been mentioned, the best variety grows, according to Avestan and Vedic sources, on the high mountains (Tian Shan, Pamir, Himalayas),¹⁴⁷ and that is exactly where the more potent variety of Ephedra is

¹⁴⁴ 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; discussion in Witzel (forthc. b).

¹⁴⁵ See Parpola 1995, Nyberg 1995 (= Ephedra), and the recent Leiden workshop on the problem; its results have partially been published in *EJVS* 9, (May) 2003, ed. by J. Houben. For Ephedra in Afghanistan and in the Kalash valleys of Chitral, see next note.

¹⁴⁶ Presentation at the 3rd Harvard Round Table on the Ethnogenesis of South and Central Asia, May 2001 (<http://www.fas.harvard.edu/~sanskrit/RoundTableSchedule01.html>) and now in *EJVS* 9 (2003). Note that the word for *Asafoetida* (a plant producing resin) is similar in shape: IIr. *anću* could go back to an older pre-IIr. form **ank'u* which is found in Toch. as *añkwaš*, 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. Pinault (2003) further connects the Common Toch. word for "iron" **añcuwæn* > Toch. A **añcu* (adj. *añcuwaši* "made of iron"), B *eñcuwo* (adj. *eñcuwaññe*) and cf. Khor. *hnčw* < Iran. **anśuwan*. The meaning "iron" derives from the rusty color of iron ore, just like the *Soma* plant or its juice are called "rusty brown", indeed the color of the powder of some Ephedra varieties, cf. below. The Rgvedic *Soma* is called *babhru*, *hari*, *aruša*, *aruṇa* (i.e. "brownish" through "reddish". -- Interestingly, Arrianos, *Anabasis* iii, ch. 28 is aware of it as *silphion*, growing in the "Caucasus" (Hindukush); this has been identified as *Asafoetida*, liked by sheep, see Wirth 1985: 887, n. 111. However, note that the same is said about Ephedra = Kalash *sámani* (<**sumanas*, CDIAL 13492): it grows in the higher mountain meadows, is eaten by goats..., and is used as an ingredient in making chewing tobacco (*nazwár*). It is, according to Morgenstierne who went there in 1929, a "small shrub resembling heather, with stiff, straight twigs, growing in the high mountain valleys." The latter description is the same as that given for *hum* in Afghanistan: "a stiff bush, three feet high, with closely arranged, vertical, leafless, articulated stems. It grows on stony, infertile ground. The fruits are red and fleshy and are eaten by children. The stems are used for dying yellow, and are used, as powder, for chewing [tobacco] and as snuff. The Afridi tribes crush the stems and soak them in cold water. The extract derived from it is used as medicine against fever. The plant grows in all of Turkestan, north and middle Persia, in northern and eastern Afghanistan and in the northwestern Himalayas." (K. F. Geldner, *Der Rig-Veda*, introduction to vol. iii, Cambridge: Harvard Oriental Series 35, 1951: 2; my translation). "The central portion of dried mature [Ephedra] twigs collected in autumn contains a powdery material of rusty red colour" (Pinault 2003, N.A. Qazilbash, *The Pharmaceutical Journal* 26, 1960, 499).

¹⁴⁷ See now Staal (2001, 2003) with an exact localization; the RV speak of the best variety, *Maujavata*, from the mountain *Majavant* (AV), cf. Avestan *Muza*, found even now as the mountain *Muzh* Tagh Ata and the R. *Muzh-kol* (Xinjiang/ Tajikistan border and in N. Kashmir), see Witzel 2000a, cf. above, n. 112.

found. From there it was traded and brought in by the local (mountain) people for Vedic rituals.¹⁴⁸ 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.¹⁴⁹

Not unexpectedly then, the names of some of the most important IIr. priests and composers of ritual poetry also belong to this substrate: **atharwan* "(fire?) priest," **ṛṣi* "seer, poet", **uḥig* "sacrificing priest", **yātu* "black magic".¹⁵⁰ 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 speakers of Indo-Iranian in C. Asia. A prominent feature of IIr. religion is the emergence, probably in the Ural area (Witzel forthc. b), of a group of "gods of law and order" (*asura*, *Āditya*, with IIr. names!) that regulate the "truthful" behavior of people, from a single person to a tribe.¹⁵¹ Other important deities seem to be based on a

¹⁴⁸ Pinault (2003) draws attention to the combination of *aṃśu* "Soma" and the "foreign" name *Śṛṣṭa* (*Śṛṣṭra*, *Śrṣṭra*, RV; Kuiper 1991: 7, 70); another form of this name is RV *Kṛṣṭa*, with the frequent interchange of *k/ṣ* [k'] in non-IA names and words (Witzel 1999a,b,c). -- See further Witzel 2000a for a discussion of non-IIr. Hindukush and Pamir mountain peoples.

¹⁴⁹ Sarianidi 1992: 34; this custom continued even down to the Greek period in Bactria (Ai Khanum). -- For the alleged finds of Ephedra ("Soma") at Togolok in Bactria see Lamberg-Karlovsky 2002: 71 and now Houben *EJVS* 9, 2003.

¹⁵⁰ In detail: **atharwan*, Ved. *Atharvan* "[ancient] priest, sorcerer", Avest. *āθrauuan* "priest", EWA I 60; Pinault (2003) connects this with Toch. B *etre*, A *atār* "hero" < **athr* "superior force" and assumes suffixation with the common IIr. suffix -*van* (cf. Ved. *athar-yu*, *athar-vi*, *athar-i*); **ṛṣi* "seer, poet", Ved. *ṛṣi*, cf. O.Avest. *ərəṣiš* "(person) full of enthusiasm for god," Y 31.5, *ərəṣišīia*, EWA I 261; **uḥig* "sacrificing priest", Ved. *Uṣij*, *Auṣija*, Avest. *usij* "sacrificer following non-Zoroastrian 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).

¹⁵¹ The Vedic/Avestan lists have: 1. Ved. *Varuṇa*, Mitanni *Uruna* (substituted by Avest. *Ahura Mazda*?; with a local(?) -*una* suffix otherwise rarely found in IE, IIr. (but note : **Perk^w-uno-* in Lith. *Perkūnas*/Slav. *Perun* "thunder god"); IIr. -*una* is seen in: *Yamunā*, *taruṇa*, *aruṇa*, cf. Avest. *tauruna*, *auruna*, etc.; note Blažek 2002: 233 for Elamite *Urun*, at Haft Tepe, 2nd half of 2nd mill. BCE), 2. Ved. *Mitra*, Mit. *Mitra*, Avest. *Miθra*, "god Agreement", 3. *Arya-man*/*Aṛīia-man* ("Arya-hood", god of marriage exchange, an artificial formation from an adjective with the deverbal derivative suffix -*man*, similarly Avest. *xʾaē-tu* "self-hood" > "belonging to a family"), 4. *Bhaga/baya* (O.P. *baga*) "god Share", 5. *Aṃśa* "god Lot", 6. sqq. *Dhātṛ/dātar* "god Apportioner/Creator", *Dakṣa* "god Cleverness", *Mārtāṇḍa Vivasvant/Gaiiō marətan*, *Vīuuanhūuan*. These new deities are active in the social relationships, respectively: the world, tribe, clan (marriage), family, individual; *Vivasvant* is the ancestor of *Manu* and *Yama/Yima* and thus, of human beings (or at least of all *ārya* lineages). These deities are not found in IE (Slav. *bog^u* "god" is a loan from N. Iranian *baga* "god"). Their PIr. origin may be due to socio-religious developments in steppe conditions (water rights, oral agreements, etc., see Witzel, forthc. b); this location is indicated by the several correspondences of these early words in Uralic and Yeneseian, such as *Asura* > Mordwin *azoro* "lord", Vogul *ātər* "prince" (Koivulehto 2001: 247), etc. -- However, the BMAC population had an altogether different local religion, with a preponderance of an (agricultural) fertility Goddess and her antipode, a polymorphous, male scaled-skin Dragon deity (combining several animal and human forms and allied with snakes and lions), who is fought by an eagle-headed Hero (Francfort 1994), see now Witzel, forthc. b; cf. below §6.

substrate designation. They include **carwa* "name of Rudra", **indra*, **g(h)andharw/b(h)a* "a demi-god or demon".¹⁵² These are, however, of 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".

Importantly, Lubotsky (2001: 304, 306) now draws attention to the common "suffix" *-*arwa*-¹⁵³ in **atharwan*, *g(h)andharw/b(h)a*, and perhaps also in **carwa*.¹⁵⁴ Interestingly, this "suffix" is seen only in religious terms, which at a minimum points to intentional, if not artificial, formations involving these words.¹⁵⁵

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 closer evaluation in comparison with Vedic and Avestan religion.¹⁵⁶

The remaining group of words only attested in Indian sources¹⁵⁷ and words only attested in Iranian sources¹⁵⁸ will be evaluated in another paper.

¹⁵² Details: **carwa* "name of a god", Ved. *Śarva* "name of Rudra" (commonly seen as archer, hunter), Iran. *Sauruua* "a demon"; Pinault (2003) connects the Toch. word for "hunter", Common Toch. **šærwæ*, B *šer(u)we*, A *šaru* "hunter", and N.Iran. words: Iran. **sarwa* > **saurwa* (Avest. *Sauruua*; cf. *aora* ~ Ved. *arvāñc*) > Osset. Digor *sorun*, *surd*, Iron *sūryn*, *syrd* "to hunt", Khotan. *ha-sūrā*, *hu-sure* < **fra-saura* "quarry, hunted beast" and deduces the meaning of **carwa* as "**hunting, hunter, living in the forest* > **being of the wilds* > *god of the wilds, destroyer*" (cf. also Osset. *Syr(don)*, the name of a wicked Nart, J. Colarusso, pers. comm.), **indra* "name of the 'king' of the present generation of 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. *gandarəša* "a demon, monster". -- Blažek 2002: 232-3 compares Indra with the Old Elamite names *Int(a)ri*, *Inda(p)* from *d/ta* "to put, lie". J. Colarusso (pers. comm.) thinks that "Indra/Indara, Hitt. *Inara* (Inra, Inar), all point to this name having been an epithet. Note Circassian *lyəna-ə-ra/* big-be-gerund, Abkhaz *la-yn-ar/* the-big-(be-)gerund (the name of the god of the forge), all meaning "the big, great one".

¹⁵³ Originally, Lubotsky had apparently thought that this group of words belonged to a *still older* layer of loans in IIr.

¹⁵⁴ Doubtful unless one assumes a root **c-*, as can happen in Caucasian languages; cf. NEC river names, n. 91. See, however, n. 152 for a Toch. reflex, B *šer(u)we*, A *šaru* "hunter"; cf. n. 143.

¹⁵⁵ For details see above n. 151: just like *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 (RV *a-kṛṣi-vala*)?

¹⁵⁶ For an initial investigation, see Witzel, forthc. b.

¹⁵⁷ One may also investigate, for example, the names of clearly non-IA gods and demons in the RV, only a few of which fit Lubotsky's trisyllabic pattern with long middle syllable (see above, n. 113): *Araru*, *Arbuda*, *Iḥbiša*, *Uraṇa*, *Emuša*, *Karañja*, *K(a)ulitara*, *Khela(?)*, *Cumuri*, *Jarūtha*, *Ḍṛbhika*, *Namuci*, *Paṇi*, *Parnaya*, *Pipru*, *Bṛsaya*, *Raji*, *Vaṇgrda*, *Sambara*, *Śaṇḍa*, *Simida*, *Sṛbinda*, etc. 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.

¹⁵⁸ Such as O.P. *kaufa* "hump of a camel, mountain"; Avest. *aḍu* "channel, rivulet", *aḍa* "land, property", *aoniia* "fire place", *akana* "receptacle", *aku* "scissors", *ayažāna* "an agricultural instrument", *tātuk* "clay" (note the Toch. loan (?) *tuk-ri* "clay", Pinault 2002: 334), *-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/aya*, *axti*, *aḥša*, *aḥsman*, *aḥsman*, *aṇra* (= Ved. *asra*), *ara*, *-yaska*, *-skanda*, *-stairiia*, or those of insects and others pests such as *anāiriti*, *araēka*, *aširiia*, and finally personal

§ 4. Place Names and Hydronymy

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.¹⁵⁹ The Greeks, starting with Herodotos¹⁶⁰ and the historians¹⁶¹ of Alexander's trek towards India, provide a number of interesting toponyms and personal names, again mostly of Iranian nature. Ptolemy's *Geography*¹⁶² adds a host of names, especially for the areas on and beyond the Yaxartes, in the Saka territories and in Eastern Central Asia.¹⁶³

There is an increasing number of mythical and real names found in Middle Iranian Zoroastrian and Manichaean texts, and 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 survey article by Khromov (1980, 132-139). All of these later data cannot be dealt with in this paper.

Modern data for these areas 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 hydronymy 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 < *Baxδi*, 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* (< Ir. *Dān-*), *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

and place names such as *Axtiia*, *Apaxštrā*, *Ankasa*, *Ainiiauuu*, *Amru*, as well as those of demons such as *Arəzura*, *Aiēhiūā*, *Ara*, *Asabana*, etc. A detailed investigation will be presented elsewhere.

¹⁵⁹ See above §2.3: Choresmia, Sogdia, Bactria, Margiana, -- all already Iranian names, or mythical rivers such as the *Vafhi/Wehrūd*; see Vamberi 1891, Witzel 2000a.

¹⁶⁰ 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*, *Baktrianoι*, *Aigiloi* (on the Iaxartes); iii 117 the *Khorasmioi*, *Hurkanioi* (in Gorgan), *Parthoi*, *Sarangai* (= O.P. *Zranka*), *Thamaniaioi*; vii 64 sqq: *Baktrioi*, *Sakai*, *Skuthai*, *Khorasmioi*, *Sogdoi*, *Kaspioi*.

¹⁶¹ 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* (Gk. *Iaksartes*); 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 Merv?); *Marakanda*, *Zariaspa*, *Ga[za]bai* (near Bukhara), *Nautaka*, *Paraitakēnē*, *Khoriēnos*, *Baktra*; cf. also Curtius Rufus, *Historia Alexandri Magni* (*De gestis Alexandri Magni*).

¹⁶² See Ronca 1968/1971 and Humbach 1972, 1998.

¹⁶³ 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. *barəzant-* [antara-] *Kayha-* < **kasa*, with the pass *Xšaδrō.suka*), the *Khatai*, *Akhasa*, the *Khauranaoi*; and finally, in Eastern Central Asia (from north to south): the *Piadaï*, *Asmiraia*, the *Oikhardai*, *Issēdones*, *Throanoi*, *Thagouroi*, *Aspakarai*, *Batai* and *Ottarokorai* (= *Uttarakuru* "beyond the Himalayas", AB 8.14); 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.

Darya: (Lat.) *Oxus*, *Urgenc Daryasi*, *Xorezm Daryasi*, *Oks*, *Okus*, *Araks*, *Jeixun*, *Kelif Daryasi*, *Balx Daryasi*; and the Zerafšan: *Buxara Daryasi*, *Xaramkam*, *Rudizar*, *Obe* [= *āb-i*] *Sogd*, *Obe Kuxak*, *Samarkand Daryasi*, *Zaravšan*. Out of these, obviously only the (parts of the) old designations not stemming from Ilr. (-*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), Ėdel'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,¹⁶⁴ Iran,¹⁶⁵ (and Turkic toponymy in general as also Mongolian, etc. hydronymy).¹⁶⁶

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. Bengtson¹⁶⁷ has recently shown beyond doubt that Northern Caucasian and Burushaski belong to the Macro-Caucasian language family.¹⁶⁸ This vastly distributed group of languages, stretching from the Basque country to the Pamirs is now left with just a few remnant populations.¹⁶⁹ 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.

¹⁶⁴ Dodykhudoev 1975, Ėdel'man 1975.

¹⁶⁵ Savina 1964, 1980; Rozenfel'd 1964.

¹⁶⁶ See in: Murzaev et al., 1980, 72-88

¹⁶⁷ Bengtson 1992, 1999 and at the 3rd Harvard Round Table on the Ethnogenesis of Central and South Asia, May 2001.

¹⁶⁸ Bengtson (1992, 1999, 2001) reconstructs a Basque-N. Caucasian-Burushaski (Macro-Caucasian) language family.

¹⁶⁹ One may regard it as the remnant of the early (Mesolithic) settlement of Europe and Central Asia by *Homo Sapiens sapiens* (cf. the work of the geneticist Spencer Wells in Western C. Asia, e.g. paper at the 15th ICAES (Florence, July 5th - 12th 2003 : Genomic Footprints of Humans in Asia and Eurasia): "The Eurasian heartland: Central Asia in Eurasian prehistory"; Wells 2001). 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 differing opinion of J. Nichols 1997: 128 (and cf. 1998: 226 for Nakh-Daghestanian in the eastern Caucasus foothills); she assumes a proto-Kartvelian homeland southeast of the Caspian. If that were correct, Kartv. would, improbably, just be the Caucasian remnant of the "BMAC language." Note, however, Blažek 2002a on early loan relationships between Afro-Asiatic (non-Semitic) and N.Caucasian, and cf. Janhunen 2001: 213.

Indeed, there is the wide-spread river name **Sind-*. It is attested north of the Caucasus, as the *Sindes*, a people on the Kuban river.¹⁷⁰ 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*, *sindmin* "river".¹⁷¹ Finally, the major river of the Burusho area in N. Pakistan and beyond is the Indus, first met with in the RV *Sindhu*, a river that begins in the Himalayas and receives the Kabul, Kurram, and Gomāl as well as all the Panjab rivers (RV 10.75). P. Thieme (1991) 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."¹⁷² The word is also found in O.Iranian as Avest. *həndu* "border river, ocean (at the western and eastern ends of the world)"¹⁷³, O.Pers. *handu* [*həndu*] "Indus". This points to an IIr. coinage with the meaning "border river, ocean".

The question remains why this reconstructable PIIr. word is so similar to Macro-Caucasian words that reflect something like **(t)sin(d)/*(t)sir*;¹⁷⁴ (see above, n. 91, on NEC **ʔTenc'o-* "river" and NEC **ʔTor-/PEC ʃVrV* "water".)

In the case of IA one may also think of an adaptation of Bur. *sinda*,¹⁷⁵ from Proto-Burushaski,¹⁷⁶ to the well known IIr. and IE **sidh/sindh-* "to divide". Pinnow, indeed, connected IIr. **Sindhu* with Burushaski as he could not find a cogent IE etymology and as he rejected Near Eastern ones (Pinnow, 1953: 12-13.)¹⁷⁷ He also points to a number of river names in the northwestern subcontinent which have the

¹⁷⁰ Mayrhofer 1979. Note L.G. Gulieva 1960. The local Circassian name for the Kuban is *q⁰ə-bəy-na* which may mean "turn much, Meanderer" (J. Colarusso, pers. comm.). An older name was *Sag-dan* (Ptolemy), from Scythian/Osete *don*, Ir. *dānu*, see Dul'zon 1964: 16.

¹⁷¹ Bur. *sinda*, Yasin/Werchikwar dial. *sēnde* (< NIA Shina *sin?*), see Pinnow 1953: 12-13; cf. also Bur. *tsil/ts.hil*, *ts.hilmin* "water", Werk./Yasin *tshel*; cf. NE Cauc. **ʔTenc'o-* "river" (see n. 91); Harmatta (1990: 376-377) compares the wide-spread river name *Sind-* with Burushaski *sinda*.

¹⁷² 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*.

¹⁷³ Cf. *us.həndauua* "beyond the natural frontier" (Thieme), that is: the (mountain) "rising from the ocean (Milky Way)" / or "beyond the Milky Way", *zraiah vourukaša* (Witzel 1984).

¹⁷⁴ 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.

¹⁷⁵ Cf. Witzel 1999a,b,c.

¹⁷⁶ Cf. Tikkanen 1987, 1988, 1998. Early loans include Ved. *kilāla* / Class. Skt. *kīlāṭa* "a milk product" (RV), Bur. *kilāy* "curds"; Ved. *meṣa* "sheep", Bur. *mēṣ* "skin bag" (but 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: *yūpas* "cotton", Ved. *kārpāsa*; all of this is in need of further investigation. For pre-Bur. substrates see the summary in Kuiper 1962, 1991.

¹⁷⁷ Cf. also Mayrhofer 1979, on the *Sindes*, a people on the Kuban R., north of the Caucasus mountains and the *Sindes* river (Tedxhen, in Tacitus, *Annales* 10; Witzel 1999c, and below §5, n.180; cf. however, n. 171, 178

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.¹⁷⁸

However, the wide spread of these words east of the Caucasus rather seems to indicate a Macro-Caucasian source **sind-/sir-* (see §5) that has been connected, *already in Ilr. times*, with Ilr. **si(n)dh* "to divide" and etymologized as **Sindhu* "border river."¹⁷⁹ This suggests a localization of the process near the BMAC *Sindes* river rather than the *Syr Darya* (*Silis*).

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 Cherkess republic. This name, attested at least since 7th cent. CE,¹⁸⁰ seems to echo¹⁸¹ that of the Kabul river, RV *Kubhā*, Gr. *Kophēs*, *Kophēn*. Another interesting river name is that of the Indra River in S. Tajikistan, *Indar-ab*,¹⁸² and the

¹⁷⁸ 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. **kintu* > **āntu* [*āndu*, *āndu*] "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 1999a,b,c).

¹⁷⁹ 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*, *Ahuramazda*, see Hintze 1998). Ir. *Hindu-* thus can be a loan from an older IA substrate. However, the "eastern and western" *həndu*, "oceans" Y. 57.29, and the name of the mythical central mountain, *us.həndauua* "emerging from the river/ocean [*Vourukaša*]" indicate *həndu* "(mythical) ocean" in Avestan (Witzel 1984), derived from an Ilr. (loan) coinage "border river, ocean" which would fit P. Thieme's etymology (1967-91) from the IE root **si(n)dh* "to divide."

¹⁸⁰ Note the area called *Sindikā* south of its mouth into the Sea of Azov. L.G. Gulieva, 1960, 140 quotes other names of the Kuban, such as *Kum* in the Armenian geography of Moses Xorenac'i; *Kofina* in the *Breviarium*, the history written by the Constantinopolitan patriarch Nikephoros (c. 750-829 CE), when describing the time of the emperor Maurikios, reigned until 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*; J. Colarusso, pers. comm., points out that this is Circassian /*psəš* 'ə/ "water;" name of a tributary to the Kuban), *Burlik*, *Bal*, *Kum*.

¹⁸¹ 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. *vispō.paiti*, N.P. *nabid*), and OIA *pāntiya* "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. *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 explained, as the *Kuban* above, from a related Cauc. word.

¹⁸² The N. Afghani river *Andar-ab* has to be separated from this as it is attested in mid-first millennium Chinese sources with *an-*, see Witzel 2000a, n. 7.

Inder lake (Russ. *ozero Inder*) on the lower Ural river in W. Kazakhstan. In light of the proposed non-IIr. etymology¹⁸³ of the name of the god *Indra* (see above) these widespread names may reflect the C. Asian substrate language as well.¹⁸⁴ Much more research is needed, however, to turn these proposals into something closer to certainty.

As far as Greater Iran itself is concerned, the O.Iran. sources (O. Persian inscriptions and the Avestan texts)¹⁸⁵ 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 literature, on place names), *Kampanda* and *Nisaya*, and the fortress *Sikaya(h)uvati* (cf. O.P. *ṣika* "gravel", Ved. *sikata*), 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 *Bāxθr-* (Witzel 1980), may be derived from O.Ir. **Baxθ(r)-* < IIr. **bhāk-tri-* which may mean the "distributing (river);"¹⁸⁶ however, the loss of *-r-* remains enigmatic (but cf. O.P. *Skud(r)a*, below). Y.Avest. *Baxδr* corresponds to the AV loan *balhika* and would indicate a Y.Avest. dialect form *Baxδr-* already at the time of the AV, c. 1000 BCE (Witzel 1980, 1997b). If the local form of the name *Baxδr* 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. *Suyda*, 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, *Aršāda*¹⁸⁷ and *Kāpišakani* and a *Gandutava* district. *Kāpišakani* is close to *Kāpišr*, a town in the Kabul area (Gandhāra), see Pāṇini 4.2.99; Patañjali further specifies *Kāpišayana/-r* "inhabitant of *Kāpišr*".¹⁸⁸ In the same area may be found the Vedic *Kamboja* tribe, reflected by the O.Pers. royal name *Kambūjiya* "Cambyzes".¹⁸⁹ Finally, there are the SE Iranian provinces of *Karmāna* and *Maka* (modern Makrān,

¹⁸³ 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.

¹⁸⁴ Taking into account the Macro-Caucasian family, the French river name *Indré* (south of Tours) may reflect a Basque substrate.

¹⁸⁵ There is need of a detailed study of the many names of mountains in Yt 19; for some initial attempts see Witzel 2000a.

¹⁸⁶ Cf. also the *Vaxšθrikā* mountains in Yt 19.4, see Witzel 1980, 2000a.

¹⁸⁷ One may try, however, the IIr. etymology, such as *arša(n)-d(h)a* "placing/containing men."

¹⁸⁸ Cf. *Kapisa* in Ptolemy, Geogr. 6.18.4; see Witzel 1980: 108: n. 49, Eilers 1982: 21; on Kandahar < Gandhāra, see Eilers 1982: 22.

¹⁸⁹ 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, and Sauvira (Sindh) in *Arthaśāstra* 2.30.9.

Gedrosia) whose inhabitants are called *Maciya* "person from *Maka*."¹⁹⁰ These may reflect the old names of the area (see above, on *Marḥaši/Bampur*).¹⁹¹ (The Southwest of Iran is proper Elamite territory and therefore left out here).¹⁹²

§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. In addition to the materials collected by Lubotsky (2001: 303 sqq.) we can discern, just as in the case of the rather limited dialect materials available for the northern and southern Indus languages (Witzel 1999a,b), for example a dialect difference involving an interchange of *r/n*. Naturally, we are still very much in the realm of speculation here, as the available data are still very sketchy and come from a variety of quite different languages and sources. It is also still somewhat difficult to pin them down in time and place. Nevertheless, a few observations are added here in the hope of drawing attention to the facts.

(1) *r/n*. Assuming the existence of the Macro-Caucasian language family, the terms for wheat have the structure ***gVr/l* (Basque *gari*, PECauc. **ḡōl'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 cluster *-an-* to *-o-* (PKartv **ghomu*, Georgian *yomu*, Ved. *godhūma*, Drav. **kotumpai*, *gōdi*, see Witzel 1999a,b) 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-*.

The loan word 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 *parā*; 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-*, ***s₁enḡha*, pre-OIA *sinḡha*, etc. Note the possible western form in Arm. *inc/inj* < **sinḡh*.

The same distribution is reflected by the word for "water, river": Macro-Cauc. **(t)sir*, NEC **š^Tor-*, PEC **š^{Vr}V* "water" which may be preserved in the name of the Yaxartes, *Silis*, in Bur. *t.s.hil* etc., (cf. Basque *(h)ur* Bur. *hur*, and also Yeneseian **xur*). However, the case of southern *-n-* (*Sindes/Sindhu* river, Bur. *sende*) is more complex. NEC **š^Tenc'o-* "river" can be compared with the rivers *Sind-* in the Caucasian Kuban,

¹⁹⁰ Witzel 1980: 112 n. 76 and Eilers 1982: 30 with literature on place names.

¹⁹¹ Incidentally, the Greek name of the capital of this area, *Poura*, does not seem to be connected with IA *pura* "city", cf. modern *Bampur*; see Mayrhofer 1979.

¹⁹² In the SW, the towns *Kuganakā*, *Tāravā*, *Māru*, and *Raxā* and the mountains *Arakadi* and *Parga*; *Paṭiś(h)uvari* (Gr. *Pateiskhorēis*), a Persian tribe Eilers 1987: 49, "gegen die Sonne gerichtet" in Bartholomae's *Wörterbuch*.

Turkmenian Tedzhen, and Indus areas, where it is fairly early, as Mesop. *sinda* "wood from Meluhha" (Baluchistan/Sindh) attests.¹⁹³ A Macro-Caucasian word ***sin(d-)* apparently was taken over early on into Indo-Iranian and given a popular etymology from IE/Ir. **sidh/sindh* "to divide", which is what border rivers indeed do; consequently Ir. **sindh-u* designates (border) rivers and the rim of the world, the ocean surrounding the world (Avest. *həndu* "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/Həndu/Hāndu*.

With some reservations one may add the name of the province of Gorgān (Hyrkania, *Vəhrkāna*) in N. Iran, Avest. *Xnənta*, found next to an original **Xrənta* (Humbach 1984, 1991). The name certainly is non-Ir. and as the Avestan texts were redacted in Arachosia, this would represent a split between the Macro-Caucasian areas close to the Caucasus, in Turkmenistan-Bactria and the Pamirs, versus a southeastern O. Iranian dialect close to the Vedic area, with *-n-*, (see above on *-šer/*sing'ha*).

The interchange between ***sinšap* "mustard" > pre-Iran. **sinšapa*, Khot. *śśaśvāna*, Parth. *šyš-d'n*, Sodg. *šywšp-δn*, M.P. *span-dān* "mustard seed" vs. Vedic *saṣarpa* "mustard" may also belong here.

In sum, much of the old Macro-Caucasian belt, stretching from the Pamirs to the Caucasus (and further west), has a predilection for forms with *-r-*, where others have *-n-*.¹⁹⁴ The actual distribution has been influenced and confused by the immigration of Ir. speakers.

It probably goes too far to see in the interchange of *r/n* two representations of a prehistoric retroflex **ṇ*, though Pinault (2003) thinks that the combined evidence of *Paṇi* and *āṇi* points in that direction: BMAC **āni* (*āṇi?*) : Ved. *āṇi* "lynch pin" : Toch. **āni-en* "hip"; BMAC **Paṇi/Parna* (Gr. *Parnoi*) : Ved. *Paṇi* : Toch. **paniyā* "that which belongs to wealthy people" where the Greek form *Parnoi* would be a local variant with "intrusive" *-r-*, cf. Kuiper 1991:70-81).

If this were the case and the BMAC language indeed had forms like **āni*, **parṇa/parṇi*, one must wonder why the regional Hindukush-Pamir feature of retroflexation (cf. Tikkanen 1988, 1999, Witzel 1999a,b) would have affected the BMAC area and Nuristani/Vedic but not the neighboring O.Iranian: Avestan was spoken in the presently retroflexing Pashto territory, for example in Kandahar/Arachosia. One way out of this dilemma would be to assume an earlier "retroflex Avesta", imported from Arachosia (K. Hoffmann 1975/6, 1992) into the Persis around 500 BCE, where it was transmitted orally under Old Persian phonetic influences for the next thousand years or so. However, there are no materials, so far, to sustain the assumption of an older, "retroflex Avesta".

Furthermore, the supposition is contradicted by inner-OIA and E. Iran. evidence: the word **mani* "jewel" turns up as *maṇi* in Vedic but without retroflex in the other OIA dialect, Mitanni-IA, as *mani-nnu* and also in Avestan *-maⁱni*. At best, we can assume local variation in the Hindukush-Pamir belt itself and with all

¹⁹³ 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.

¹⁹⁴ Note also the more widely spread IE apophony of *r/n* (and *r/i*, named after W. Caland) in Indo-European of the type nom. **wedō-r*, gen. **wede-n-es* (innovative from the point of view of Nostratic, Witzel 1992).

those who have migrated through this "retroflexing belt" (Nuristani, E. Iranians, Saka, Burusho, Vedic Indians, Dravidians, etc.); note however, that NWC also has retroflexation of its affricates and fricatives.

(2) k : g Avestan may supply another feature, an interchange of *k : g* as in **kaiša* > *keša* (~ Caesar?) :: Avest. *gaēsa*; cf. above, western (Hitt., Semitic) *kant/hant* "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* > **būja* > 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, kh : ka may be added; note Ved. *khā*, Avest. *xan* "source, well" : Ved. *khan* "to dig", *khara* "dug out mound" :: Iran. **kan* "to dig" (Avest. *us.kənti*, *ni-kaiṇti*, *auua-kaṇta*, O.P. *ni-kaⁿtuv*, *kaⁿtanaiy* etc.), cf. EWA I 446.

(4) t : dh

is seen early on in IE *med(h)-* (Skt. *madhu*, Engl. *mead*, etc.) :: western IE ***melit*, Gr. *mélit-*, Hitt. *milit*, Lat. *mel*, *mell-*, Gothic *miliþ*.

(5) c : j?

may perhaps be deduced from Armen. *inc*, *inj* "lion" < ***sinj* ~ pre-Ved. **sinjh* < **sing'h* :: Toch. A *śisäk*, B *śecake* "lion".

(6) syllable structure CəCəCə,

as discovered for the Central Asian substrate by A. Lubotsky (2001: 303, 305, see above, n. 113).

In sum, the still expandable list of the names, etyma, phonetic and grammatical details, discussed so far,¹⁹⁵ allows us to posit a pervasive substrate in Old Indo-Iranian that has certain phonetic peculiarities and that does not go back to the Sumerian, Akkadian, Hurrite, Elamite, Dravidian, Indus or Tocharian 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.

In order to reach a still greater understanding of the BMAC civilization, it will be necessary to carry out further linguistic investigations -- such as a detailed study of (Old) Iranian and Tocharian words-- and then comparisons with the archaeological record. Cooperation with geneticists, physical anthropologists, and other specialists will also yield valuable clues to the identity of the people(s) who inhabited the area of the BMAC. By all the evidence available to us, it is very clear that they interacted with their neighbors (and indeed, more distant communities) in all four directions. The strong influence they exerted on the late Indus Civilization is becoming increasingly obvious now.

¹⁹⁵ 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; see now above, n. 158.

On the other hand, the Bactria-Margiana area is the immediate contact zone for steppe populations coming from the north.¹⁹⁶

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

There is no need to underline that the establishment of a BMAC substrate belt has grave implications for the theory of the immigration of speakers of Indo-Iranian languages into Greater Iran and then into the Panjab.

By and large, the body of words taken over into the Indo-Iranian languages in the BMAC area, necessarily by bilingualism, closes the linguistic gap between the Urals and the languages of Greater Iran and India. Uralic and Yeneseian were situated, as many IIr. loan words indicate, to the north of the steppe/taiga boundary of the (Proto-)IIr. speaking territories (§2.1.1). The individual IIr. languages are firmly attested in Greater Iran (Avestan, O.Persian, Median) as well as in the northwestern Indian subcontinent (R̥gvedic, Middle Vedic).

These materials, mentioned above (§2.1.) and some more materials relating to religion (Witzel forthc. b) indicate an early habitat of Proto-IIr. in the steppes south of the Russian/Siberian taiga belt. The most obvious linguistic proofs of this location are the FU words corresponding to IIr. *Arya* "self-designation of the IIr. tribes": Pre-Saami **orja* > *oarji* "southwest" (Koivulehto 2001: 248), *ārjel* "Southerner", and Finnish *orja*, Mordvin *ufe/uřä*, Votyak *var*, Syry. *ver* "slave" (Rédei 1986: 54). In other words, the IIr. speaking area may have included the S. Ural "country of towns" (Petrovka, Sintashta, Arkhaim) dated at c. 2100/2000 BCE (see the archaeological and linguistic summary in Witzel 2000a, Lamberg-Karlovsky 2002). This, however, is not the place to engage in a detailed discussion of all of the relevant archaeological materials.

It is a truism that "Linguists too often assign languages to archaeological cultures, while archaeologists are often too quick to assign their sherds a language" (Lamberg-Karlovsky 2002: 74), but Mallory (in Lamberg-Karlovsky 2002: 79) is equally right in asserting that "there are still degrees of geo-linguistic plausibility".

Indeed, we cannot be sure that (Proto-)IIr. was actually spoken at Sintashta-Arkheim around 2100/2000 BCE (Witzel 2000a), but it must be pointed out that the archaeological assemblage and the geographical position of these sites close to the taiga makes this quite likely: the Sintashta-Arkheim complex has the newly developed spoked (proto-)chariot and many other items (horse sacrifice, grave structure, Dadhyañc style replaced horse head in a grave at Potapovka, *pur*-style forts, etc.) overlapping with the early IA and Old Iranian cultures and texts (Witzel 2000a, Anthony in Lamberg-Karlovsky 2002: 75). The discussion of all such relevant IIr. words and concepts is unfortunately missing in Lamberg-Karlovsky (2002) and with most of his interlocutors in that issue of *Current Anthropology* (with the partial exception of Anthony and Mallory); instead they operate with rather vague, bloodless notions of IIr., hardly progressing beyond Benveniste's IE(!) linguistic reconstructions of the social sphere (Benveniste 1973).

¹⁹⁶ It should be noted that the late BMAC shows an intrusion of steppe pottery (Hiebert, 1998, Shishlina and Hiebert 1998, Lamberg-Karlovsky 2002).

	Eurasian Steppes	Central & South Asia	Mesopotamia
Year BC 1000			
1500	Andronovo Almashevo Fedorov	Bishkent Yashk	Kassite Hittite
2000	SAP Catacomb Graves	Indus Civ. BMAC	Ur III Akkadian Empire
2500	Almashevo	Namazga	Early Dynastic
3000			Unik
3500			
4000			

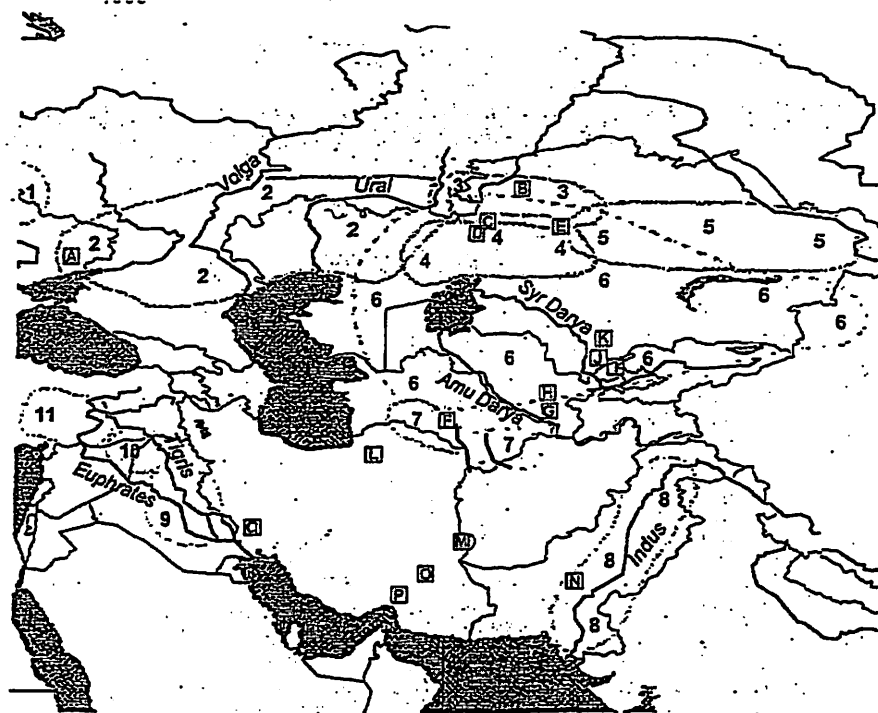


FIG. 2. Principal archaeological sites and cultures mentioned in text. Sites: A, Mikhailovka; B, Petrovka; C, Arkaim; D, Sintashta; E, Bqtai; F, Namazga; G, Gonur; H, Togolok; I, Dashly Oasis; J, Sapelli; K, Djarkutan; L, Hissar; M, Shahr-i-Sokhta; N, Sibri; O, Shahdad; P, Yahya; Q, Susa. Cultures: 1, Tripolye; 2, Pit Grave/Catacomb; 3, Sintashta/Arkaim; 4, Abashevo; 5, Afanasevo; 6, Andronovo; 7, Bactrian Margiana archaeological complex; 8, Indus; 9, Akkadian; 10, Hurrian; 11, Hittite.

(From Lamberg-Karlovsky 2002)

That the oldest IIr. texts (Rgveda, Avesta) are about 1000 years later than the date of the Sintashta-Arkhaïm complex (Lamberg-Karlovsky 2002) is *not* of as great relevance as thought. First, the relevant words from the two very closely related languages can easily be reconstructed from the extant texts for the P-IIr. period. In addition, both texts are notoriously archaic in their language, culture, and religion, and actually contain some reminiscences of Central Asia (Gr. *Rhā* "Volga" ~ N. Iran. *Rahā*, Ved. *Rasā*, *Parna* ~ Ved. *Pañi*; N. Iran. *Daha*, *Dahā-ka*, Ved. *Dāsa*, *Dasyu*; *Sarayu* = *Harōiuu-m/Harē* = Herat R., Ved. *Sarayu*; **Sindh-* ~ *Sindēs* River (Tedzhen) ~ Iran. *Həndu*, Ved. *Sindhu*, etc., (see above, Witzel 1984, 1995, 1999c).

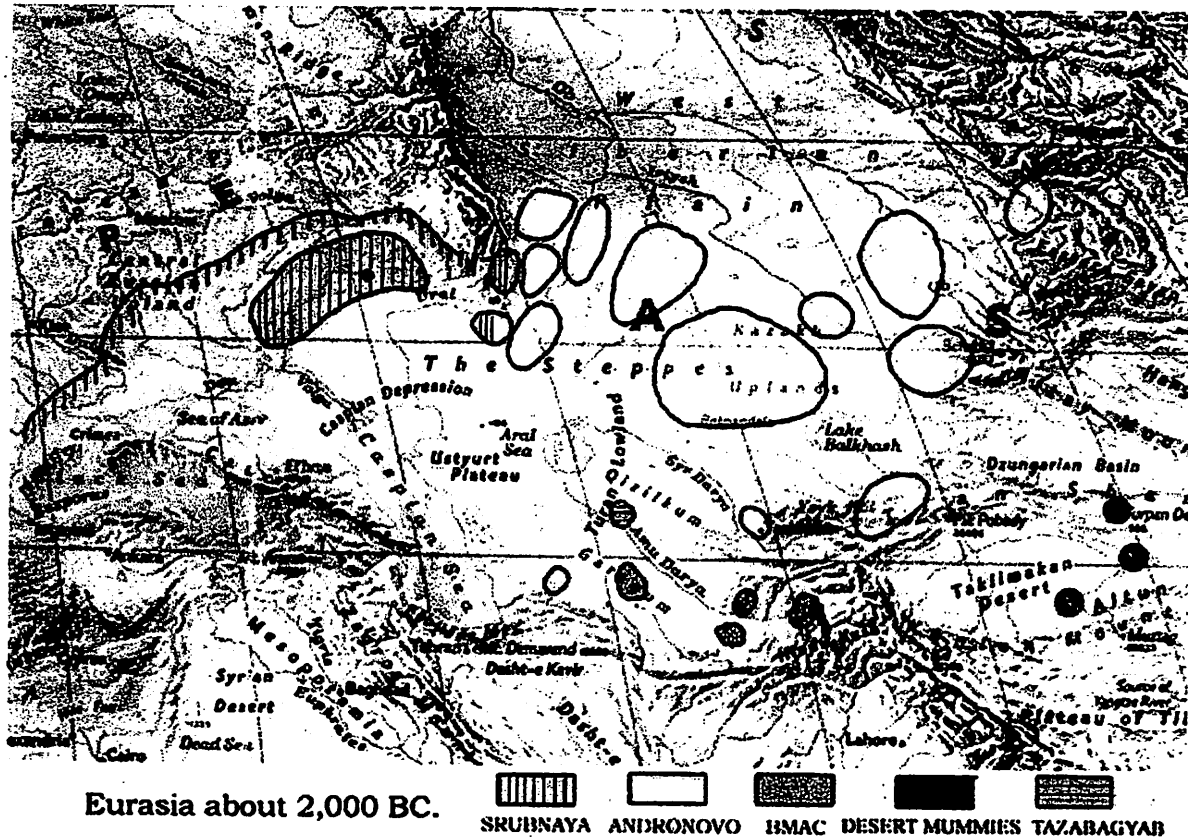
The *older* forms of IIr. words have been taken over into Uralic and Proto-Yeneseian, as has been discussed above (see Kott *art'a* §2.1.2; see n. 151 for *asura* > Mordwin *azoro* not, e.g., from the later, Iran. *ahura*). This again underlines the early age of contact, before and around 2000 BCE. In this light, the geographical location and spread of the eastern Catacomb, Sintashta-Arkhaïm, Afanasievo and finally the early (northern) Andronovo cultures make for a more or less widespread overlap with speakers of (P)IIr., though occupation by some other languages (also lost ones) cannot be ruled out altogether, at least for part of the area: i.e., Uralic and Yeneseian at the northern borders, while Altaic is excluded (perhaps except for some Proto-Turkic in the extreme East, Róna-Tas in Lamberg-Karlovsky 2002: 82 sq.).

It is likely that, like in Turkic and Mongolian times, there was use of a *lingua franca* in the wide steppe (and desert) belt. This cannot have been Uralic, Yeneseian, Altaic or another unknown language as we do not have any indication of any respective influence on the southern languages (BMAC, Elamite, or later, on attested OIA, OIr.) This *lingua franca* most likely was an IIr. *koine* (cf. Kohl in Lamberg-Karlovsky 2002: 77-78), a form of P-IIr. (and later on, of pre-OIA, then of pre-Ir.), as is witnessed in the various levels of IIr. loans into Uralic and Yeneseian.

The clearly defined situation described above contradicts Mallory's assertion, in spite of his principle of "degrees of geo-linguistic plausibility", that "there are clear instances, the Indo-Iranians being a case in point, in which there is no hint of the distribution of any archaeological assemblage that might correlate with the target language group" (loc.cit., p. 80). The use of an IIr. *koine* also does *not* contradict, as Kohl seems to think, the model of a tree-like linguistic divergence model: the IIr. "mythical homeland" is indicated by the correlation of linguistic and zoological/botanical evidence, and as the various stages and branches of the IE/IIr. tree model are visible in the "quasi-archaeological" layers of loans words taken over from the IIr. languages into the Uralic and Yeneseian languages. A *koine* (Hellenic Greek, Latin, French, Russian, English) simply does not imply "fusion" of languages à la Trubetskoy (Kohl in Lamberg-Karlovsky 2002: 77, cf. Makkay p. 78). Such fusion is rarely if at all visible even in the developments of Pidgin and Creole languages. They always have a strong basis in one extant language but have taken over some grammatical traits and words from others (not unlike medieval English!).

In sum, the agnosticism of Lamberg-Karlovsky and other archaeologists with regard to a correlation between IIr. languages and the steppe archaeological cultures is repudiated by the increasing wealth of "archaeologically" stratified linguistic data, generally neglected, that locate PIIr. in the steppe belt *just south* of the Uralic/Yeneseian *taiga*, in other words, in the very archaeological areas discussed above (eastern Catacomb to northern Andronovo).

Finally, as outlined elsewhere (Witzel forthc. b), there is an additional number of words from the religious sphere (*anču* ~ Soma, etc., Lubotsky 2001) that again indicate a gradual spread of IIr. speaking tribes southwards from the "quickly filling steppes" (Kohl) of the Catacomb - S. Ural - Afanasievo areas, all of which is not unlike the attested eastwards and southwards spread of the Andronovo culture that has created well documented overlaps with the BMAC in the Merw delta, on the Zerafshan River and at Kangartut in S. Tajikistan (see Lamberg-Karlovsky 2002: 71, 73).



(From D. Anthony: Samara Project, <http://users.hartwick.edu/iaes/newsletter/newsletter.html>)

Against this background of a (partial) overlap of the steppe archaeological cultures and the location of tribes speaking various forms of IIr., a scenario of cultural and linguistic interactions and actual movements can be drawn up. In the form of a brief summary, this would include the following steps.

- Gradual immigration of the cattle herding speakers of common Proto-Indo-Iranian (or of pre-Old Indo-Aryan) from the steppe belt into the general BMAC area

- Gradual immigration of the cattle herding speakers of common Proto-Indo-Iranian (or of pre-Old Indo-Aryan) from the steppe belt into the general BMAC area (cf. Mallory in Lamberg-Karlovsky 2002: 80). This general, seasonal migrational pattern was continued, just as in Afghanistan transhumance, well into our time. (Meridional migrations of Kazakhs took place down to 1929 CE, Olsen in Lamberg-Karlovsky 2002: 81). Again, the IIr. languages *must* have come from the northern steppe areas as the *early* (Proto-IIr.) loans into Proto-Uralic (*asura*, Koivulehto 2001: 247) and Yeneseian (*art'a*) clearly indicate. This contact persisted for several millennia as the virtually "archaeological" layers of loans indicate.

- Amalgamation of BMAC/Central Asian words into the (late) common IIr., pre-Vedic and pre-OIr. languages then took place, along with their underlying concepts (**bhiš*, **kapauta*, etc.), religion (the **sauma* drink, **-rwa* beings), animals (**uštra*, **khara*) and plants (**bhanga*, **anču*). The non-IE BMAC religion, as depicted in its seals and other art (Francfort 1994, 2001, Anthony in Lamberg-Karlovsky), seems to have directly influenced the Avestan and Vedic form on certain IIr. beliefs, such as the Avestan version of the hero fighting the dragon of drought (*Aži/Ahi*/ **Vərəθra* / *Vr̥tra*), transforming the IE (and Eurasian, Witzel 2001b) myth of the killing of the dragon into one of releasing the waters by the late spring snow melt in Afghanistan (Avesta) and in the northwestern Indian subcontinent (RV). The prominence of the BMAC Goddess of waters and fertility has influenced, to some extent, the character of the Avestan river Goddess *Anāhitā* and of the Vedic *Sarasvatī*.

While such interaction can be deduced from linguistic analysis and comparative religion, it is very difficult to indicate, by archaeological means alone, the actual "form of symbiosis" of the two antithetical and dissimilar cultures, the agro-pastoral Andronovo and the settled BMAC culture with its irrigation agriculture (Lamberg-Karlovsky 2002: 74). However, there are many steppe type sites near the BMAC settlements (Lamberg-Karlovsky: 71, 73).¹⁹⁷ While there is some indication of steppe materials in actual BMAC sites, the opposite is not true. Some degree of avoidance (Lamberg-Karlovsky 2002: 73) between the bearers of both distinctly different cultures seems likely. However, some details of the BMAC culture *must* have been taken over, at some time in the second mill. BCE, by the speakers of IIr. (note the list of BMAC words of agriculture, settlement religion, above § 3.3-4, and see below).

The incoming steppe people with Andronovo cultural traits must have shed many of these characteristics in the Greater BMAC area (Mallory 1998, in Lamberg-Karlovsky 2002: 80, cf. Kohl, p. 78) before moving on, as "not a single artifact of Andronovo type has been identified in Iran or in northern India" (Lamberg-Karlovsky 2002: 74), all while keeping their IIr. language - and, somewhat differently from Mallory, also much of their spiritual culture.

Mallory thus is right (in Lamberg-Karlovsky 2002: 80) in pointing out that "this would require far more intimate relationships between the Andronovo and the Bactrian Margiana complex than the existing distribution of "mutually exclusive" material culture would permit." However, the question that has not been put yet is: exactly when should the extensive exchange as seen in the BMAC loan words in Vedic

¹⁹⁷ For early steppe-Bactria/Margiana contacts see Francfort 2001: 153 about Kelteminar pottery and a Afanas'ev's funary stone circle found at Sarazm II, i.e. before 2500 BCE. For late steppe pottery see the preceding note.

and OIran. have taken place? The steppe pottery found in the BMAC (see n. 196, 197) may just reflect the *forerunners* (no horses!) of a more massive IA influx at the end of the BMAC, around 1600 BCE. While Lamberg-Karlovsky (2002) is still looking for a model of such cultural change, the actual state of affairs may be still have been remembered in and is reflected by the conservative poetry of the RV: the Paṇi (wealthy, "stingy", rich in cattle) are depicted as holed up in their forts (*pur*) while the R̥gvedic Aryans are depicted as being *outside* and desiring to *get in* and acquire the cattle (Elizarenkova 1995). As has been pointed out above (cf. §1.1) this *topos* may very well be a reminiscence of the situation in the BMAC area where the steppe tribes opposed the Parna (*Parnoi*, *Parni*) on the Sindes (Tedzhen/Sindhu river).

Incidentally, a tradition of avoidance similar to the one in the BMAC area is still seen, much later, in the Sistan/Arachosian area (Falk 1997) and in the R̥gvedic Panjab (Witzel 1995, 1997b): while, conveniently, many agricultural, musical, and a few religious terms of the *small tradition* were taken over (Kuiper 1955, Witzel 1999a,b,c), the local settled Dasyu populations as such were avoided and were despised (note, e.g., RV 3.53.14 about the Kikaṭa and the "misuse" of their cows). What else may one expect of proud, semi-nomadic cattle herders with their habitual disdain for farmers?

The obvious solution to look for, out of Lamberg-Karlovsky's and Mallory's dilemma of contact/avoidance of the steppe and BMAC cultures, is the one indicated just now: some trade and exchange, but also occasional friction and warfare (fortresses of the BMAC!), perhaps even including some steppe mercenaries(?), existed between the impoverished pastoralists at the fringes of BMAC settlements (cf. Kohl in Lamberg-Karlovsky 2002: 78) and the occupants of the BMAC, perhaps not unlike the relationship arising between the nomads and the occupants of fixed settlements in later history.

Some sort of contact is clearly in evidence in the borrowed vocabulary found in the IIr. languages, and just as in the RV later on, it is restricted to agriculture, village life, small tradition religion, but it also included a few more prominent terms for priests (*atharwan*, *uṇij*), ritual (*anáu*, *yātu*) and deities (*éarwa*, *g(h)andharw/b(h)a*). Even then, the IE and IIr. pattern (Father Heaven, drink of immortality, the hero killing the dragon, the IIr. Asura deities, etc.) is clearly maintained in the early Iranian and Vedic texts (Witzel forthc. b), and little influence seen of the prominence of the BMAC goddess or the anthropomorphic dragon and eagle (Frankfort 1994, 2001: 154). Equally so, the Dumézilian three-level IE social structure (poet/priests, nobility, commoners) was maintained but it was enlarged, both in Iran and in the Panjab (or, e.g., in Greece, the *pan-hellenes*), by a fourth class (*Śūdra*) that made room for persons from the local populations that had joined the *arya/ariya*.

Such adjustments will be difficult to detect by archaeology. If they have indeed been looked for, then in the wrong direction: we cannot expect Zoroastrian rituals in the BMAC in 2000 BCE but only around 1000 BCE, not every hearth is an IIr. "fire altar", and the findings of Ephedra ("Soma") in the BMAC have not been substantiated (see the discussion in *EJVS* 9). The occurrence of certain steppe vessels in BMAC contexts could point in that direction -- if they had indeed been found with Soma presses and filters. Most notable is the absence, so far, of horse remains, horse furniture, chariots (invented around 2000 BCE) and clear depictions of horses in stratified BMAC layers. One can hardly imagine the IIr.s without their favorite prestige animal, the horse. The archaeological picture of avoidance/contact by the forerunners

of the massive IA move onto the Iranian plateau so far remains sketchy. Perhaps it can be explained if the main period of major contacts was as late as c. 1500 BCE.

Once the successor settlements of the BMAC were abandoned around 1500 BCE, a partially changed Iir. speaking, entirely pastoral culture (Anthony, *op. cit.* p. 76), probably swelled by some of the Bactria-Margiana populations, spread all over Greater Iran. This is accompanied by a clear cultural change, with the appearance of painted handmade pottery in the former BMAC area (Frankfort 2001: 154) at 1500 BCE and the accompanying disappearance of tomb and grave structures in Central Asia (reflecting some Vedic and Zoroastrian customs). The proposed comparatively late date of the onward migration towards Mesopotamia and the Panjab at c. 1500/1200 BCE fits this scenario better than an early influx into, and cohabitation with, the late Indus civilization, as some have assumed (e.g., Allchin 1995: 47, at 2200-2000 BCE *sqq.*).

- This new, amalgamated, *late* Iir./pre-OIA speaking entity moved -- *Kulturkugel* fashion (Mallory 1998, 2001: 360 *sqq.*)-- into Iran and towards the Panjab. By this term, Mallory means a culture that has kept its Iir. language but has taken over (much of) BMAC cultural and societal structures. Conversely to the situation during the BMAC period, this expansion can only sparsely be substantiated, so far, by linguistic data as the relevant spade work in (Old) Iranian has not yet been done.¹⁹⁸

It is probable that this move was preceded by successive spearheading forays of (non-Iir. speaking) mountain peoples into Mesopotamia, such as the Guti, Lullubi, and Kassites¹⁹⁹ (c. 2250-1750 BCE), who were as yet only marginally influenced by Iir. languages and customs. Some of them are perhaps represented by the sudden expansion of BMAC materials into Susa, Shahdad, Tepe Yahya, Hissar, the Gulf, Baluchistan, the S. Indus area (Lamberg-Karlovsky 2002: 72, 74, 84) and Harappa (R. Meadow, *pers. comm.*).²⁰⁰ Lamberg-Karlovsky (2002: 84), however, thinks of this spread as "the prime candidate for Indo-Iranian arrival on the Iranian plateau," which

¹⁹⁸ Only some initial guesses are possible, for example about the ethnic nature of the *Tukriš* (see above n. 102) which might be connected with Ved. *tugra*, *tugrya* (both personal names), Iran. *tuyr*-. If true, we would have continuing RV (and later Vedic, BŚS) links with Bolan, Aratta, and Shahdad -- recalling the more northern trail that lead the Mitanni-Indo-Aryans westward into N. Mesopotamia. However, note the pre-OIA words in Kassite (c. 1740 BCE-), and cf. now Blažek (1999, 2002a) on early Elamite connections with Vedic.

¹⁹⁹ Only a few Kassite words seem to come from Iir., e.g. *Šuriišaš* "sun god", *Maruttaš* "divine Marut comrades of Indra", *Bugaš* "god Bhaga?"; see Balkan 1954, for horse names such as *akriyaš* = *agriya-s* "(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. later on, Finnish *kuningas* "king" < P.Germanic **kuningaz*, as seen in Dutch *koning*).

²⁰⁰ In this context, a remarkable overlap between BMAC and Indus shamanistic concepts has not been noticed, as far as I see: a cylinder seal (Sarianidi 1992: 25, fig. 33) and a terracotta tablet from Mohenjodaro (Kenoyer 1998: 83, fig. 5.6) show remarkably similar scenes of processions of flag and standard bearers (cf. Avestan *ərəδβō.drafsa* V. 1.6), the latter involving carrying animals on a pole and being accompanied by a figure beating a typical shamanic circular drum (still found with in Kalasha ritual, in the eastern Hindukush). Sarianidi (1992: 24, 26) takes the scene as one depicting jumping athletes or acrobats. There is, however, comparatively little shamanism in the Veda, and the use of the circular drum is not attested so far.

in the light of the above discussion is too early, but he (correctly) suggests that "the indigenous people, although in the majority, adopted their language," -- later on, that is (cf. below, § 6, end). A similar move may have brought speakers of PDrav. to Bolan and Sindh.

- Later, apparently after the abandonment of the BMAC and successor settlements around 1650/1500 BCE and the spread of pastoralism all over Iran (Anthony, in Lamberg-Karlovsky 2002: 76), the actual spread of speakers of pre-Vedic IA took place, that is of Mitanni-OIA, into N. Iraq/Syria (c. 1400 BCE), an area settled by the Caucasian-speaking Hurrites. The speakers of the linguistically slightly later, though still pre-Iron Age R̥gvedic then moved into Arachosia (**Sarasvatī* > Avest. *Harax̥a'ti*), Swat (*Suvāstu*) and Panjab (*Sapta Sindhu*), before c.1200/1000 BCE -- depending on the local date of the introduction of iron (Possehl and Gullapalli 1999), which still is missing in the R̥gveda but found in the next level of Vedic texts.

- The intermediate Hindukush area has been largely neglected in scenarios of this kind. However, the R̥gveda does not only take note of some its geographical features (*Kubha* = Kabul River, *Suvāstu* = Swat, the opposition *giri: ajra* "mountains: flat valley pastures"), it also is influenced by certain religious ideas of the Hindukush area, such as the concept of Yakṣ(iṇ)ī/Apsaras (**Śucī* "pure" > Kalash *sūci*) and Rudra/Gandharva as inhabitants of the pure snow mountains, snow/ice dragons engulfing the flowing waters (the later Kashmirian Nāgas), and the like (Witzel, forthc. b: §1.5.6.). The RV also contains a number of words that can be linked with the local Pamir language, Burushaski (Witzel 1999 a,b), such as Bur. *kīlay*, RV *kīlala*- "biestings, a sweet drink". Indeed, the Hindukush/Pamir area is one of transhumance that was well suited for the Indo-Aryan pastoralists (Witzel 2000a). Movements between the mountain pastures of the Hindukush highlands and the Panjab/Sindh lowlands and the continue to this day, including that of cattle (Meadow, oral comm. based on personal observance).

Furthermore, it is precisely in this area that the phonetic feature of retroflexation, so typical of Vedic (and of South Asian languages in general), must have set in (Witzel 1999 a,b). This feature is *missing* in Mitanni-IA and Old Iranian but typical for all languages of the Hindukush/Pamir areas, whether they be Burushaski, E. Iranian, N. Iranian (Saka), Nuristani, or IA (from RV to modern Dardic); retroflexation even has affected the eastern (i.e. S. Asian) dialects of the newcomer, Baluchi, a *West* Iranian language.

- The move toward the Panjab may have been independent of and may actually have been preceded by that of the speakers of the third group of IIr. languages, now called Nuristani, whose speakers, originally called Kafirs by their Muslim neighbors, live in the Hindukush mountains of NE Afghanistan. They have preserved some archaic features until today (Nur. *ç* is older than RV *ś* or Avestan *s*, all from IIr. **ć*). Such movements may also have included that of the speakers of the non-IIr., western-IE group now represented in the substrate of Bangani, a NIA language in the high Himalayas of Uttarkhand, on the border to Himachal Pradesh. However, the people who spoke that substrate language may just as well have come, as potential IE neighbors of the "western-IE" Tocharians, *across the mountains* from the general area of modern Xinjiang. People often establish their alpine grazing grounds (and settlements) *across* the mountain range they border on: German speakers in Wallis/Valois and S. Tyrol, Slovenian in Carynthia, Ossete north and south of the Caucasus range, Iranian Yidgha

in the NIA speaking Chitral, Kafiri in westernmost Chitral, Tibetans (Sherpa, Bhutanese, etc.) on the southern side of the Himalayas.

It might be added that the general path of immigration of the speakers of Indo-Aryan from the north into the Panjab, via the general BMAC/Hindukush area, is also indicated by an early loan from Nuristani. This is Nur. **kat's'a* > Ved. *kāca* "shining piece of jewelry" (K. Hoffmann 1976, EWA I 33),²⁰¹ also taken over into O.P. as *kāsa-ka* "semi-precious stone."

• All of this is followed by the spread into Greater Iran of the earliest Iranians (c. 1000 BCE, Hintze 1998, cf. K. Hoffmann 1976-92 [= 1941], for some pre-Ir. names in the RV), with the introduction of E. Iranian (Avestan) into E. Iran (1200/1000 BCE -- note the overlap with AV *Balhika* "Bactria", Witzel 1980). The movement of the West Iranian tribes, Median and Persian, into W. Iran, is later still, c. 900-700 BCE.²⁰² Lamberg-Karlovsky (2002: 74) stresses the fact that the spread of BMAC materials cannot be linked to the later archaeological developments on the Iranian plateau in the later 2nd and 1st millennium as would be required by the spread of the Iranian speaking groups.²⁰³

In sum, as far as South Asia is concerned, it can now be stated more securely that speakers of an IE language, early OIA (pre-Rgvedic) entered the Greater Panjab from Afghanistan, acquired local words from the Northern Indus dialect (such as *śaṇa*, *laṅgala*, *vṛīhi*, *godhūma*, *kaṅgu*, *Gandhāra*, Witzel 1999a,b). About the same time(?) speakers of Proto-Dravidian entered Sindh, acquired related words from the southern

²⁰¹ 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 O.P. *kāsaka* "semi-precious stone", *kāsaka kapauta* "lapis lazuli," and *siṁkabru* "carnelian" described as brought from Sogdia, and *kāsaka axšaina* "from Choresmia" (DSf 37-40). One would expect Bactria/Badakhshan.

²⁰² It remains to be investigated whether the Persians (*Pārsa* < **pārćva*-) are related to the Parśu (< **parću*) of the Vedic texts (RV, BSS), where they are located next to the Arattas (*āratta*, *aratta*), thus in Afghanistan. These are likely to be the ancestors of the Pashto (*paštō* < *-rǵ/**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 **parćva* is connected with Pashto or not, Old Persian -s- (as in < *asa* "horse") < **śś* < *śv* < *ćv* < IE *k'w* shares the development of Ilr. *ćv* > *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*.

²⁰³ The question of the location and spread of early Iranian is not discussed here. It is likely (see above) that this form of Ilr. developed further north in the steppes and spread both westwards (Scythians) and eastwards (Saka) as well as southwards (E. Iranian), and still later, also south-westwards (W. Iranian: Median, Persian). This took place only after an early southward move of the (pre-)OIAs from the northern steppes, as suggested by Burrow in 1973; cf. Lubotsky 2001: 308 sq. and Chlenova (1984) who "shows a correspondence between Iranian place names and the distribution of the Timber Grave, Andronovo, and related cultural groups. Place names of Indo-Aryan character are scattered or absent in that area" (Makkay in Lamberg-Karlovsky 2002: 79).

Indus dialect (*gōnu*, *ñāñcil*, *variñci*, *godī*, *kañku/kampu*), and perhaps it was they who brought the first horses to South Asia (Pirak, Eastern Baluchistan near the Bolān Pass, c. 1800 BCE, see Allchin 1995: 31, Kenoyer 1998: 78, Witzel 1999a,b), rather than the IA(?) Bhalānas (RV 7.18), whose name seems to be reflected by the modern Iranian place name.

A similar scenario for Greater Iran cannot yet be written as the relevant linguistic investigations have not yet been carried out: we do not have a comprehensive study of loan words in early Iranian (and Hurrite/Urartian, Elamite, etc.). Instead, it has often been alleged that Old Iranian has fewer loan words from the local substrates than Rgvedic, all in spite of the well attested pre-IIr. archaeological cultures of Greater Iran, from Tepe Hissar to Mundigak. The assumption is a fallacy, as a closer look at the Avestan vocabulary will indicate (see n. 158 for the direction to be taken.) Scholars apparently have been misled by the glaring archaisms of Zoroaster's IE poetic language (cf. Kuiper 1979) as to assume a "pure" IIr. language.

The whole process of "Aryanization" in Iran and India, progressing with a large degree of intervening bilingualism, may be summed up in the words of Polomé (1990: 337). He 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, of customs, beliefs, rituals, religion,²⁰⁴ and material culture.

²⁰⁴ For an initial discussion see Witzel (forthc. b); to be added is the comparison of a shamanistic BMAC seal and its Hindukush and Vedic relationships, see n. 200.

ABBREVIATIONS

AB	Aitareya Brāhmaṇa
Akkad.	Akkadian
Armen.	Armenian
Austro-As.	Austro-Asiatic
AV	Atharvaveda Samhitā
Avest.	Avestan
Brah.	Brahui
BŚS	Baudhāyana Śrautasūtra
Bur.	Burushaski
Cauc.	Caucasian
CDIAL	Turner 1966-69
Circ.	Circassian
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
GS	Gr̥hyasūtra(s)
Guj.	Gujarati
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
JB	Jaiminiya Brāhmaṇa
Jpn.	Japanese
Kab.	Kabardian
Kan.	Kannada, Canarese
Kartv.	Kartvelian
Kaśm.	Kashmiri
KEWA	Mayrhofer 1986-96
Khot.	Khotanese Saka
KZ	Zeitschrift für Vergleichende Sprachwissenschaft
Lith.	Lithuanian
M	Middle-
Mal.	Malayalam
Mar.	Marathi
Mbh.	Mahābhārata
MIA	Middle Indo-Aryan
Mong.	Mongolian
M.P.	Middle Persian
MS	Maitrāyaṇi Samhitā

MT	Mother Tongue (Boston)
N	New-
NEC	Northeast Caucasian
Nep.	Nepali
N.P.	New Persian
NIA	New Indo-Aryan
Nur.	Nuristani (Kafiri)
NWC	Northwest Caucasian
O	Old-
O.Avest	Old Avestan
OHG	Old High German
OIA	Old Indo-Aryan
O.P.	Old Persian
Osset.	Ossete
P	Proto-
Panj.	Panjabi
PEC	Proto-East Caucasian
Pkt.	Prakrit
PNC	Proto-North Caucasian
PPerm.	Proto-Permian
PS	Paippalāda Saṃhitā
RV	Ṛgveda Saṃhitā
RVKh	Ṛgveda Khila
Samh.	Saṃhitā(s)
Sant.	Santali
ŚB	Śatapatha Brāhmaṇa
ŚS	Śrautasūtra
Skt.	Sanskrit
Sum(er).	Sumerian
Sū.	Sūtra(s)
StII	Studien zur Indologie und Iranistik
Tam.	Tamil
Tel.	Telugu
Tib.	Tibetan
Tib.-Burm.	Tibeto-Burmese
Toch.	Tocharian
TS	Taittirīya Saṃhitā
Up.	Upaniṣad(s)
V.	Videvdād
Ved.	Vedic
VS	Vājasaneyi Saṃhitā
Y.	Yasna
Y.Avest.	Young Avestan
Yt.	Yašt
Yen.	Yeneseian
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
- Allchin, F. R. *The Archaeology of Early Historic South Asia. The Emergence of Cities and States*. With Contributions from G. Erdosy, R. A. E. Coningham, D. K. Chakrabarti and B. Allchin. Cambridge: Cambridge University Press 1995
- 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.
- Behr, W. Hic sunt leones. Two ancient Eurasian migratory terms in Chinese revisited. Paper at the 31st International Conference on Sino-Tibetan Languages & Linguistics (Lund, 1-4 Oct. 1998). Preprint, n.d.
- 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
- Benveniste, E. *Indo-European language and society*. Coral Gables: University of Miami Press 1973
- Berger, H. Die Burušaski-Lehnwörter in der Zigeunersprache. *IJ 3*, 1959, 17- 43
- , Bericht über sprachliche und volkskundliche Forschungen im Hunzatal. *Anthropos* 55, 1960, 657-664
- Blažek, V. The Sino-Tibetan etymology of the Tocharian A *mkow* - B *moko* 'monkey'. *Archív Orientalní* 52, 1984, 390-392
- , The new Dravidian-Afroasiatic Parallels. In: V. Shevoroshkin (ed.). *Nostratic, Sino-Caucasian, Austric and Amerind*. Bochum: Brockmeyer 1992: 150-165
- , Is Indo-European *H₁ek'wo- 'horse' really of Indo-European origin? *Studia Indogermanica Lodziensia* 2, 1998, 21-31
- , Elam: a bridge between Ancient Near East and Dravidian India? In: Blench, R. and M. Spriggs 1999, 48-77
- , Elam: A Bridge between the Ancient Near East and Dravidian India? *Mother Tongue VII*, 2002a, 123-143
- , Elamo-Arica, *JIES* 30, 2002b, 215-241
- , Hic erant leones. Indo-Iranian "lion" et alii. (preprint, nd. np.)
- , 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

Michael Witzel, "Linguistic Evidence for Cultural Exchange in Prehistoric Western Central Asia,"
Sino-Platonic Papers, 129 (December, 2003)

---, *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, È.M. Murzaev, V.A. Nikonov, V.V. Tsybul'ski (eds.). *Toponimika Vostoka. Novye
Issledovaniia*. Moskva: Nauka 1964

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

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

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

C. Carpelan et al., *Early contacts between Uralic and Indo-European*, Helsinki, 2001 [Memoires de la
Societie Finno-Ougrienne 242]

Carpelan, C. and A. Parpola. Emergence, Contacts and dispersal of Proto-Indo-European, Proto-Uralic
and Proto-Aryan in archaeological perspective. In: Carpelan et al. 2001: 55-150

Chlenova, N. Archaeological materials of the Iranians of pre-Scythian times and the Indo-Iranians.
Sovetskaia Arkheologiia, 1, 1984, 88-103

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 Kiadó 1975

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. *Pamirskaiia 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 indoevropskogo proiskhozheniia*. In: Bondaruk et al., 1964,
14-17

---, *Ketskie imena i prozvischa*. In: Murzaev et al., 1980: 5-9

Èdel'man, D. I. *Osnovnye voprosy lingvisticheskoi geografii. Na materiale indoiranskikh iazykov*. Moskva:
Nauka 1968

---, *K substratnomu naslediiu tsentral'noaziatskogo iazykovogo soiuzu. Voprosy iazykoznaniiia*, 5, 1980,
21-32

---, *Geograficheskie nazvaniia Pamira. Strany i narody Vostoka*. Vyp. 16. Pamir. Moskva 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 Onomastics. 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.

EJVS 9: see Houben, J., et al. [Soma Workshop, Leiden 1999]

Elizarenkova, T. *Language and style of the Vedic R̥sis*, Albany : SUNI 1995.

Erdosy, G. (ed.). *The Indo-Aryans of Ancient South Asia*. (Indian Philology and South Asian Studies, 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 R̥gvedic 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.]. [Mémoires de la Mission archéologique française en Asie centrale] Paris: Mission archéologique française en Asie centrale: de Boccard 1989.

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

---, The Central Asian dimension of the symbolic system in Bactria and Margiana. *Antiquity* 68, 1994, 406-418

--- 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

---, The Archaeology of Protohistoric Central Asia. In Carpelan et al. 2001: 151-168

Gamkrelidze, V. T. and V.V. Ivanov. Pervye Indoevropeitsy v istorii: predki Tokhar v drevnei Perednei Azii. *Vestnik Drevnei Istorii*, 1989, 14-39 [Engl. Summary: "The first Indo-Europeans in history: ancestors of the Tokharians in ancient Near East", p. 38-39]

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 Archaeology*, 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 nazvaniikh reki Kuban. In: Murzaev et al., 1960, 135-139

Hänsel, B., S. Zimmer et al. (eds.) *Die Indogermanen und das Pferd. [Akten des Internationalen interdisziplinären Kolloquiums. Freie Universität Berlin, 1.-3. Juli 1992, Bernfried Schlerath zum 70. Geburtstag gewidmet]*. Budapest 1995

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

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

Michael Witzel, "Linguistic Evidence for Cultural Exchange in Prehistoric Western Central Asia,"
Sino-Platonic Papers, 129 (December, 2003)

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

Houben, J.E.M. (ed.) *EJVS* 9: 1. The Soma-Haoma problem: Introductory overview and observations on the discussion (J.E.M. Houben); 2. Report of the Workshop (J.E.M. Houben); 3. Report concerning the contents of a ceramic vessel found in the 'white room' of the Gonur Temenos, Merv Oasis, Turkmenistan (C.C. Bakels); 4. Margiana and Soma-Haoma (Victor I. Sarianidi); 5. Soma and Ecstasy in the Rgveda (G. Thompson). [Proceedings of a Workshop on the Soma-Haoma problem organized by the Research school CNWS, Leiden University, 3-4 July 1999] *Electronic Journal of Vedic Studies* 9, 2003. <http://users.primushost.com/~india/ejvs/>

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 sravneniia nostraticheskikh iazykov*. II, Moskva: Nauka 1976

Ivanov, V.V. Tocharian and Ugrian. *Studia Linguistica Diachronica et Synchronica*. Berlin-New York-Amsterdam 1985: 411-419

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

---, Uralic and Ural-Altaic, in Carpelan et al. 2001, 207-220

Jettmar, K. *Die Religionen des Hindukusch*. Stuttgart: Kohlhammer 1975

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. M. 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

---, *Ancient Cities of the Indus Valley Civilization*. Oxford: Oxford University Press/American Institute of Pakistan Studies 1998

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

A. L. Khromov, Izuchenie geograficheskikh nazbanii Tadzhikistana v SSSR i za rubezhom. *Izvestiia otdeleniia obshchestvennykh nauk AN Tadzhikskoi SSR*. No. 2 (36) Duzhanbe 1959

---, Sogdiiskaia toponimiia verkhov'ev Zeravshana. In: Murzaev et al., 1960, 87-98;

---, Nekotorye osobennosti formirovaniia toponimii Srednei Azii v IX-XIII vv. In: Murzaev 1980: 132-139

Kibrik, A. E. and S.V. Kodzasov. *Sopostavitel'noe izuchenie dagestanskikh iazykov: imia, fonetika*. Moskva: Izdatel'stvo Moskovskogo universiteta 1990.

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

Koivulehto, J. The earliest contacts between Indo-European and Uralic speakers. In: C. Carpelan et al., 2001, 235-263

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

---, Rigvedic loan-words. In: O. Spies (ed.) *Studia Indologica. Festschrift für Willibald Kirfel zur Vollendung seines 70. Lebensjahres*. Bonn: Orientalisches Seminar 1955.

---, *On Zarathustra's Language*. Amsterdam: Noord-Hollandsche Uitgeversmaatschappij 1979

---, *Aryans in the Rigveda*. Amsterdam-Atlanta 1991.

Kuz'mina, E. E. Contacts between Finno-Ugric and Indo-Iranian Speakers. In: Carpelan et al., 2001: 289-300

Lamberg-Karlovsky, C.C. Language and Archaeology. The Indo-Iranians. *Current Anthropology* 43, 2002, 63-88

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 al. 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. In: C. Carpelan et al. 2001, 301-317

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/Philadelphia: 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

---, (ed.) *Contact and Exchange in the Ancient World*. [Proceedings of a conference held at Philadelphia on May 5-6, 2001.] Honolulu: University of Hawaii Press (forthc.)

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

---, Uralics and Indo-Europeans: Problems of Time and Space. In: Carpelan et. al. 2001: 345- 365

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* (S. Deger-Jalkotzy u. R. Schmitt, eds.): Wiesbaden: Reichert 1979

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.'

[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

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: Harrassowitz 1973

Murzaev, È.M. *Tsentral'noasiatskie toponimicheskie miniatiury*. In: Bandaruk 1964, 3- 13

---, B.A. Nikonov, V.V. Tsybul'ski (eds.). *Toponimika vostoka. Issledovaniia 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 proistoskhzhdenii toponimov Ishkashim, Yazguliām i Vakhān. In: *Iranskoe Iazykoznanie*. Moskva 1976

Parkes, P. Livestock symbolism and pastoral ideology among the Kafirs of the Hindu Kush. *Man* (N.S.) 22, 1987, 637-660

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
---, Tokh. B *k_ucaññe*, A *k_ucim* et Skt. *tokharika*, *IJ* 45, 2002, 311-345
---, Further links between the Indo-Iranian substratum and the BMAC language. Paper given at the 12th World Sanskrit Conference, Helsinki, 14-18 July 2003.

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-European 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 tsentraloi Azii. In: Murzaev 1980, 118-123

Pulleyblank, E.G., The Consonantal 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

Raulwing, P. *Horse, Chariots and Indo-Europeans. Foundations and Methods of Charioty Research from the Viewpoint of Comparative Indo-European Linguistics*. Budapest : Archaeolingua [Series Minor 13] 2000

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

Reinhart, J. Die tocharischen Entlehnungen im Altaischen und die Chronologie der tocharischen Lautgesetze. In: *Tocharische Akten der Fachtagung der Indogermanischen Gesellschaft*. Berlin 1990 : 73-92

Róna-Tas, A. *Tibeto-Mongolica*. The Hague: Mouton 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 textkritischen Noten*. Rom, IsMEO, 1971.
- Rozenfel'd, A.Z. *Materialy po ètnografii i toponimike Vancha*. *Izvestiia VGO*. 85: 4, 1953
- , Zametki po gidronomii iugo-vostochnogo Tadzhikistana. In: Bondaruk et al., 1964: 175-182
- , Tadzhikskaja toponimiia Badakhshana. (Gorno-Badakhshanskaja oblast' Tadzhikskoi SSR). In: Murzaev 1980: 156-164
- Sapir, E. Tibetan influences on Tocharian. Mandelstam, D.G. (ed.) *Selected writings of E. Sapir in language, culture, and personality*. Los Angeles 1951, 273-284
- Sarianidi, V. I. *Drevneishii Merv. [In' gradymy Merv. The oldest Merv]*. Moskva: Izdatel'skaia 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 1998a.
- , *Myths of Ancient Bactria and Margiana on its Seals and Amulets*. Moscow 1998b
- Savina, V. I. O tipakh slovoobrazovaniia toponimov Irana. In: Bondaruk 1964 : 149-174
- , Ètnonimy v toponimii Irana. In: Murzaev (ed.). *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: V. Mair (ed., forthc.)
- 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
- Skjærvø, P.O. The Avesta as a source for the early history of the Iranians. In: G. Erdosy (ed.), *The Indo-Aryans of Ancient South Asia*, (Indian Philology and South Asian Studies 1), Berlin/New York: de Gruyter 1995, 155-173.
- Staal, J. F. How a psychoactive substance becomes a ritual: the case of Soma. *Social Research* 68, 2001, 745-778
- Starostin, S.A. *Praeniseiskaia rekonstruktsiia i vneshnie sviazi eniseiskikh iazykov. Ketskii sbornik: antropologiia, etnografiia, mifologiia, lingvistika*. Leningrad: Nauka 1982: 144-237
- , *Rekonstruktsiia drevnekitaikoi fonologicheskoi sistemy*. Moskva: Nauka 1989.

Steinkeller, P. The question of Marḥaši: A Contribution to the Historical Geography of Iran in the Third Millennium B.C. *Zeitschrift für Assyriologie und Vorderasiatische Archäologie* 72, 1982, 237-65
---, Marḥaši. 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 Akademie der Wissenschaften, 386, Bd. 9, 1980

Thieme, P. *Kleine Schriften*. K. L. Janert (ed.). 2 vols. Wiesbaden. 2nd ed. Wiesbaden 1984.

---, *Kleine Schriften 1970-1987*. R. Söhnen (ed.). 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>)

---, Soma and Ecstasy in the Rgveda. In: Houben, J.E.M. (ed.) *EJVS* 9 (2003).

<http://users.primushost.com/~india/ejvs>

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

---, *Éléments de géographie élamite (résumé)* PO 11, 1985, 49-54

---, *Les noms géographiques des sources suso-élamites*. Wiesbaden: Reichert 1993

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

Vasil'kov, Ya. V. and N.V. Gurov. Strana Aratta po drevnim pis'mennym istochnikam. *Vestnik Vostochnogo Instituta/Acta Institutionis Orientalis* (Sankt Peterburg) 1, 1995, 12-66

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*

Michael Witzel, "Linguistic Evidence for Cultural Exchange in Prehistoric Western Central Asia,"
Sino-Platonic Papers, 129 (December, 2003)

Gesellschaft, Innsbruck, 22.-28. September 1996. Innsbruck: Institut für Sprachwissenschaft, 1998: 119--138

--- and A. Bammesberger, *Germania Semitica*. Heidelberg 2001

Voigt, W. *Die Wertung des Tieres in der zarathustrischen Religion*. München: Reinhardt 1937

Wells, R.S. et al. The Eurasian Heartland: A continental perspective on Y-chromosome diversity.

Proceedings of the National Academy of Sciences USA 98, 2001, 10244-10249

<http://www.pnas.org/content/vol98/issue18/index.shtml>

pdf: <http://www.pnas.org/cgi/reprint/98/18/10244>

Werner, H. Reconstructing Proto-Yeneseian, *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. Jungavestisch *apāxəδra-* im System der avestischen Himmelsrichtungsbezeichnungen.

Münchener Studien zur Sprachwissenschaft 30, 1972, pp. 163-191

---, 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 1997a, 257-345

---, Introduction, In: F.B.J. Kuiper, *Selected Writings on Indian Linguistics and Philology*, ed. by A. Lubotsky, M.S. Oort, M. Witzel, Amsterdam- Atlanta: Rodopi 1997b, pp. ix-xxvi

---, Substrate Languages in Old Indo-Aryan (Ṛgvedic, Middle and Late Vedic), *EJVS* Vol. 5:1 1999a, 1-67 (<http://www1.shore.net/~india/ejvs> = <http://users.primushost.com/~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. Röhl 2000a, 283-338

---, A first link between the Ṛgvedic Panjab and Mesopotamia: *śimbala/śalmali*, and *GIŠgišimmar?* In: Karttunen, K. and P. Koskikallio (eds.) *Vidyāṛṇavavandanam. Essays in Honour of Asko Parpola*. 2000b [*Studia Orientalia*, published by the Finnish Or. Soc. 94]: 497-508

---, Die sprachliche Situation Nordindiens in vedischer Zeit. *Indoarisch, Iranisch, und die Indogermanistik, Arbeitstagung der Indogermanischen Gesellschaft 1997 in Erlangen*; hg. B. Forssman & Robert Plath. Wiesbaden: L. Reichert 2000c, 543-579

---, Autochthonous Aryans? The Evidence from Old Indian and Iranian Texts. *EJVS* 7-3 May 25, 2001a. <http://users.primushost.com/~india/ejvs/>

---, Comparison and Reconstruction : Language and Mythology. *MT VI*, 2001b: 45- 62

Michael Witzel, "Linguistic Evidence for Cultural Exchange in Prehistoric Western Central Asia,"
Sino-Platonic Papers, 129 (December, 2003)

---, 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. The R̥gvedic Religious System and its Central Asian and Hindukush Antecedents. In: *The Vedas: Texts, Language and Ritual (Proceedings of the Third International Vedic Workshop, Leiden 2002)*. Edited by A. Griffiths & J.E.M. Houben. [Groningen Oriental Studies XIX] Groningen: Egbert Forsten. [2004]

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

Since June 2006, all new issues of *Sino-Platonic Papers* have been published electronically on the Web and are accessible to readers at no charge. Back issues are also being released periodically in e-editions, also free. For a complete catalog of *Sino-Platonic Papers*, with links to free issues, visit the *SPP* Web site.

www.sino-platonic.org