•§ Michael Witzel   Feb. 17, 2000   135 k

The Languages of Harappa

§ 1. The riddle of the Indus script and language

Each year, or as it seems, by now every other month, we see a new decipherment of the Indus script. Beyond G. Possehl's fifty-odd examples (Possehl 1996), there must be some dozens more, by now frequently found on the internet. However, as is well known, neither is the script itself well understood nor do we know on which language it is based and for which languages it has been used. In this paper, I am not going to add myself to the long list of failures of decipherment of the script. Instead, I think we should first prepare the ground for such undertaking by investigating the peculiarities of the script itself, a task that has not been carried systematically enough (Wells 1998). Secondly, we should take a much closer look at the various sources which may indicate which languages were spoken in the core areas of the Indus civilization, in the Greater Panjab and in Sindh-Baluchistan.

During the past few decades, the language underlying the Indus script has almost universally been taken as Proto-Dravidian (leaving aside such unlikely explanations as Sanskrit, Sumerian, etc., see Possehl 1996). This means a form of reconstructed, early Dravidian that precedes --by some two thousand years or more-- that of the Sangam (*Caṅkam*) texts which were composed in archaic Tamil around the beginning of our era. The tacit presupposition has been that the by and large South Indian Dravidian languages preceded the various dialects of Old Indo-Aryan (OIA) in the Panjab and Sindh. Of OIA we basically have only texts in Vedic Sanskrit which is commonly believed to have been introduced into the Greater Panjab by c. 1500 BCE. In support of the Dravidian theory one usually pointed to the remnant North Dravidian Brahui language, spoken in Baluchistan; however, its presence has now been explained by a late immigration that took place within this millennium (Elfenbein 1987).[1] Regardless of the fact whether the Brahuis are remnants of an early N. Dravidian settlement or not, the map of Drav. languages (fig.1@@) shows the typical feature of small "islands" submerged in a sea of newcomers, the speakers of Indo-Aryan. Dravidian, in this view, should have been the language of the Indus area.

However, in spite of efforts spanning more than 30 years by now, relatively little of the Indus inscriptions has been 'read', either by the Russian or the Finnish teams, even after all these years of concerted, computer-aided attempts, -- and not yet in a fashion that can be verified independently (cf. the summary of criticism by Zvelebil 1990). Perhaps a decipherment is not even attainable in that way, due to the brief nature of the inscriptions (7 signs on average and hardly more than 20).

This raises the question whether there may be a fundamental flaw in the whole endeavor. Apart from the inherent difficulties of deciphering a script whose signs have unknown meaning there is the additional difficulty of the unknown language(s) it is

---

[1] Note that Brahui has eastern *b* instead of western *v*; the other N. Drav. (Kurukh, Malto) languages are latecomers in their present habitat in S. Bihar as well, as is seen by the strong Munda influence they have undergone.

written in. All decipherers of the Indus script made so far have faced this question by a leap of faith: a language was assumed and texts were read in a (proto-)form of that language, with little possibility to check on the accuracy of such translations, except for the countercheck to be undertaken against the general cultural background of the area and time. Obviously, even without going into any details, such 'translations' as "the (aquatic) birds have covered all the waterways" (Sankarananda in Possehl 1996: 109) do not make any sense on a document, nor are such sentences likely ever to be used on (business) seals -- at best, one could expect some theophoric names of a similar structure as found in the ancient Near East.

Some scholars (Erdosy 1995, Possehl 1996) have indeed raised the question whether the Indus inscriptions are written in (an) unknown language(s). A good candidate could be Masica's "Language X" that is found at the bottom, at the lowest level of Hindi and neighboring New Indo-Aryan languages (Masica 1979). Some 30% of Hindi agricultural vocabulary are neither IA nor Drav. nor Munda, in short, they stem from the unknown substrate language "X".

There are several other possibilities which shall not be explored here in detail: the ancestor of the modern Burushaski, the substrate languages of Kashmir, of Kusunda and Tharu in Nepal, of the Central Indian Nahali[2], etc. These and other substrate languages they have not yet been explored well: so far, we miss etymological dictionaries of all Indian languages, except Sanskrit. Yet, we have older and better, so far unused sources that help us to determine which language(s) was/were spoken in the area of the Indus civilization.

So far, linguists have concentrated on finding Dravidian and Munda reflexes. These studies are summed up conveniently in the etymological dictionaries by M. Mayrhofer (Indo-Aryan; *KEWA, EWA*), Th. Burrow - M.B. Emeneau (Dravidian; *DED, DEDR*), and in the work of F.B.J. Kuiper (Munda/Austro-Asiatic; 1948, 1955, 1991, Pinnow 1959). In addition, it has especially been F. Southworth (1979, 1988, 1990, 1995) who has done comparative work on the linguistic history of India (IA, Drav., Munda) during the past few decades; his book on the subject is eagerly awaited.

§ 2. A neglected source, the RV

The best way to explore this question is, by definition, to investigate the oldest text of South Asia next to the Indus inscriptions, but strangely enough, it has not yet been used to study the Indus language. This is the Rigveda, which was by and large composed in an area that overlaps with that of the Indus civilization, the Greater Panjab. It is a lively testimony of the poetry, religion, society and the language(s) of the northern half of the Indus area during the immediate post-Harappan period, the localization period of the second millennium BCE.

The language of the text, Vedic Sanskrit, is a form of Indo-European and not "Harappan". The 1028 Ṛgvedic hymns are addressed to the gods and mostly used in ritual.

---

[2] The Nihāl or Nāhal, first described as unique by R. Shaffer in 1940, are found in medieval texts such as Hemacandra's Grammar, often together with the Bhils (*bhilla*), as *lāhala, nāhalaka, ṇāhala* and as a mountain/jungle tribe on the Narmadā; details in W. Koppers 1948: 23, Shafer 1954: 349 as the original language of the Bhil, S. Bhattacharya 1957, Berger 1959, Kuiper 1962, Witzel 1999.

They were orally composed and strictly preserved by exact repetition through by rote learning (in some areas of India, until today). Vedic Sanskrit is an archaic Indo-European language, belonging to the Indo-Iranian branch (which comprises Indo-Aryan, Iranian, and Nuristani). But some 4% of the words in this sacred text are clearly of non-IE, non-Indo-Aryan origin, in other words, from a pre-IA substrate. This is remarkable: to apply this to the Near Eastern context, it would mean that an ancient Jerusalem temple ritual would contain Philistine, Lebanese or other "heathen" words ... (differently from the situation in the neighboring Hittite empire, where the preceding non-IE language, Hattic, was actually used as ritual language).

§ 3. Loan words and substrate languages

At this stage, a few words about linguistic substrates are in order. Defined here as words from substrate languages are all words in early Vedic that do not conform to IE/IIr word structure (including sounds, root structure and word formation) and have no clear IE/IIr etymology. (Lubotsky, forthc., adds also some less indicative features: limited geographical distribution, specific semantics, i.e. a category which is particularly liable to borrowing).

The situation is easily understandable when compared to that of English. We know that the early form of English, an Old Saxon dialect (a part of the Germanic branch of IE) has overlaid, in the middle of the first mill. CE, the Celtic (and Latin) languages of Britain; both Celtic and Latin have left a number of loan words in Old English as substrate words, such as London < Celtic *Lugdunum* 'town of the god Lug', *-chester* < Latin *-castrum* 'fortified settlement'. Later on, English saw the superimposed (superstrate) influences of the Viking language (N. Germanic, with words such as *egg*, place names in *-vik, -ay*), of Norman French with a large number of loans, an equally a huge amount of learned, newly formed Graeco-Latin words, and various influences from the neighboring languages such as Dutch (adstrate words such as *dike, boss,* etc.). Interestingly, Old Saxon (and German in general also have a large percentage of non-IE words (such as *sheep*) of a long-lost prehistoric North Sea language). The situation in the Panjab is similar, as will be seen below. A brief, simplifiued summary looks like this:

```
-------------------------------------------------------------------------------
GREATER PANJAB                          GREATER IRAN
-------------------------------------------------------------------------------
English loan words
Urdu/Perso-Arabic superstr. loans
influx of learned (Sanskrit) words

(dev. to New Indo-Aryan)                (dev. to modern Iranian)
influx of learned (Sanskrit) words      Arabic superstrate loans

(dev. of various Prākr̥ts)               (dev. to Middle Iranian)
Old Greek loan words                    some Old Greek loans
Old Persian/Iranian loan words      Old Persian /Later Avestan

(development from Vedic to MIA)
Later (R̥g)Vedic / OIA  dialects      Old  Avestan
```

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

Harappan language (see below)   Indo-Iranian in C. Asia
unknown local language(s)   <------Indo-European

                         N. Iraq/Syria
                         <---- OIA adstrate/superstrate remnants
                         in the Hurrite of the Mitanni realm
                       Akkadian (& Sumerian,
                       etc., local languages)

Many loan words from foreign languages can be easily detected by linguistic means, even if the source of the loan remains unknown, or if it belongs to a long disappeared language. The reason is that all languages follow certain patterns, allow certain sounds or groups of sounds while others are difficult to pronounce, or 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 now a few African names have been adopted in America, such as *Mfume*).[3]

The occurrence of uncharacteristic words therefore indicates a certain cultural influence, even if the native speaker (or a latter day scholar) may not know where these words had come from originally.

§ 4. The structure of IA and IIr words

The structure of IE words is well known (Szemerenyi 1996, Beekes 1995) and can be briefly summarized as follows. IE words usually consist of three parts: the root which carries the lexical meaning, prefixes and/or suffixes and grammatical endings. The prefixes and suffixes modify the meaning, e.g. *to write, writ-er*) and are part and parcel of the word as such (i.e. the stem); the suffixes are followed by endings which specify the grammatical (syntactical) relationships (*he write-s, the writ-er-s*).[4] IE roots ordinarily have three consonants (including the laryngeals h_1, h_2, h_3), and can only have the structure given below (Szemerényi 1970, Engl. version 1996)

---

[3] Whorf's structural formula of English monosyllabic words { 0, (s+/-) C-ŋ + V + 0, C-h } tells the same story. However, not all loans are as easily discernible as the Amerindian loan words *tipi, squaw, papoose, Manitou*, etc.; cf. however, moose < *moosu*, chipmunk < *ösitomu,* or woodchuck (marmota monax) which is derived with folk etymology from Algonkian *otchek, ochig, odjik* 'fisher, weasel'; nevertheless, the Engl. folk etymology gives the word away.

[4] The form of primary suffixes is limited to certain types, usually *Ce, CR, CRe, R, Re*, es (e.g., Skt. *-ta, -ti, -tra, -i, -ya, -as*, etc.) Secondary suffixes build up on the primary ones, thus Skt. *-u-mant, -a-tāt, -a-māna,* etc. On the other hand, suffixes such as *-āśa, -ṭa, -an-da/-a-nda-, -būth-a/-bū-th-a* (see below) do not exist in IE and IIr.

+/- prefixes   {(s) (C) (R) (e) (R) (C/s)} +/- suffixes +/- endings)[5]

While grammatical endings are not always transmitted when words are borrowed from other languages, this formula often allows to classify non-IE roots and words at a glance.[6] Not allowed are the types RCe- or Rse- (Skt. *$\dot{r}$ka, *usa*, etc.), and the types: *bed, *bhet, *tebh, *pep, *teurk/tekt.* The very structure of many of the 'foreign' and loan words in the RV simply does not fit the IE one of those properly belonging to Ved. Sanskrit (just as *Nkrumah, Mfume* must be foreign words viz. recent loans in English). Consequently, RV words such as *kīnāśa, Kīkaṭa, Pramaganda, Balbūtha,* or *Bṛbu, Bṛsaya* (Burrow 1976) are simply not explainable in terms of IE or IIr (Witzel 1999). Differently from what is sometimes maintained, we do not need a *satisfying etymology* to recognize such words as foreign. Words such as *busa, Pramaganda, Balbūtha* give themselves away regardless, simply due to their structure. This point is not always readily understood (e.g, Das 1991, Oberlies 1994).[7] In the heavily Anglicized *Massachusetts* area, for example, one does not need to know the local native American language to notice that place names such as *Massatoit, Massachusetts, Wachusetts, Montachusetts, Cohasset, Neponset, Mattapoisett, Mattapan,* or *Mashpee, Chicopee,* or *Nantucket, Pawtucket* are related and without English etymology.

In addition, it also has to be noted that IA etymologies now are (or should be) at a comparatively high level of linguistic sophistication; they must include the explanation not just of individual words but also of their constituent parts, of related roots and suffixes. The same cannot yet be said for Dravidian and Munda: *DED* and *DEDR* still consist of lists of related words only, with no explanation of their structure and the interrelation between related roots or expanded roots (roots plus certain suffixes); worse, a Munda etymological dictionary still is only in the planning and collection stage D. Stampe, N. Zide et al.). We do not have reconstructions for older Burushaski, Nahali, etc. (cf. however, *MT* 2, 3, 4). Great caution thus is to be exercised when comparing modern Munda words or even old Tamil forms (from c. the beginning of our era) with data from the RV (c. 1500 BCE).

Regrettably, etymological dictionaries for Panjabi, Sindhi, Hindi etc., do not exist either, so that the modern words used in the area of the Vedic language and culture cannot

---

[5] ( ) indicates possible appearance; *b* is very rare in IE; C = consonant (includes the laryngeal sounds, H = h₁, h₂, h₃); *e* = standard IE vowel (> Skt. *a*); it can change to *o* (> Skt. *a),* *ē, ō* (> Skt. *ā*) or disappear (*zero* forms); R = resonants, the "semi-vowels" *i, ṛ, ḷ, u, m, n,* which can also appear as *y, r, l, v, m, n*; the consonant *s,* when found at the beginning of roots, is unstable and can disappear (as in *spaś* 'spy' : *paś-ya-ti* 'he sees').

[6] Possible thus are, e.g., Skt. *ad* (eC), *pat* (CeC), *śrath* (CReC), *bandh* (CeRC), *kṛ* (CR), *śru* (CRR), *kram* (CReR), *krand* (CReRC), *i* (R), *iṣ* (RC), *man* (ReR), *manth* (ReRC), *tras* (CRes), *tvakṣ* (CReKs), *stambh* (sCeRC), *svap* (sReC), *sas* (ses), etc.; with laryngeals: *bhū* (CRH), *brū* (CRRH), *īkṣ* (HRCs), *as* (Hes), etc. Sounds inside a root are arranged according to the following order of preference: C/s-R-e..., thus : CRe-(Skt. *śram*), sRe- (Skt. *srav*).

[7] Some Indo-Europeanists and Indologists still try to find IE etymologies for as many Vedic words as possible, i.e. Werba *WZKS* 36, 1992, 14 for *Pra-maganda,* from an unlikely *pra* + **mṛgāda* 'dear eater', involving, like many other similar attempts, hypothetical MIA developments (**ā > an*??; cf. Thieme on *pūj ZDMG* 93, 1939, 105 < *pṛñc,* but see Witzel, *WZKS* 24, 1980: 24). Nevertheless, a few MIA developments are found already in the RV, e.g., *jyotiṣ* RV < **dyotiṣ,* perhaps *muhu(r) < mṛhu(r),* but cf. Kuiper 1991: 79); and cf. later on: *nāpita* ŚB 'barber' < **nhāpita* from **snāpaya-* 'to bath someone'.

immediately be compared with their older forms. Instead, they will have to be sifted out laboriously by specialists of Persian, Arabic, Turkic, MIA and OIA before we can pronounce that a certain Panjabi word is derived from IA, Drav., Munda, or from an unknown local language. The only reliable help in this undertaking we have so far is Turner's CDIAL which, for the most part, lists only IA words in their OIA form.

While there are some reconstructions for Proto-Drav. and Proto-Munda that fall within the time frame covering the Vedic period, similar reconstructions are still impossible for Burushaski, a remnant language of the Hunza Pamirs, unconnected with other languages unless we invoke the still controversial Macro-Caucasian family and comparisons with Basque, Caucasian and other Asian languages. The same applies to the other remnant languages of S. Asia such as the Central Indian (NIA) Nahali[8], Kusunda in the hills of Central Nepal, (Toba 1971; Reinhard 1969, see the recent discussions in *MT* I-IV), or the little studied (NIA) Tharu in the foothills of the Himalayas, the substrate of the South Indian (Drav.) Nilgiri languages, the Sri Lankan (NIA) Vedda, etc.

Still another problem is posed by the form of certain words which seem to allow for multiple, competing etymologies, for example from IA/IE, Drav. and Munda. A typical case where no solution is in sight involves Ved. *kalaśa* 'mug, beaker, pot' which has been variously explained from IE (Mayrhofer *EWA* 321), Dravidian (Kuiper 1955: 150, *DEDR* 1305), or Munda (Berger 1959: 58). All these etymologies have inherent problems. The IE etymology does not explain the suffix (-*śa* < *k'a* : Gk., Lat. -*ks*), and it is divergent in its vowels (Lat. *calix,* Gk. *kaluks*); Drav. *kalam* etc. do not explain the suffix -*śa* either; PMunda *ka-la(ñ)ja-* (as seen Nahali *lēñjo* 'to scoop up', Skt. *aliñjara*, Pali *alañjara* < *a-leñjo-ḍa*) is a purely hypothetical reconstruction based on Munda patterns.

Or, Ved. *kuliśa* 'axe' which *EWA* I 374 declares as not securely etymologized, has been connected with Drav. (Tam. *kuḻir* 'battle axe', Kan. *kuṭṭu* 'to beat, strike, pound', Kuiper 1955: 163), but also with Munda (in Skt. *kuṭhara, kuddāla* 'hoe', Sant., Mundari *kutam* 'to beat, hammer', Mundari, Ho *kutasi* 'hammer', Kuiper 1955: 163); Berger 1963: 419 derives *kuḍiśa* from *kodeś,* Kharia, Mundari *khoṇḍe'j* 'axe', with prefix *kon-* from Kharia *te'j* 'break'.

At this stage, there is little help, in such cases, in deciding which etymology is better than the other, unless we take refuge in the saying "etymologies are either obvious or wrong".

§ 5. The typology of the 300 RV loan words and their sources

In spite of all these caveats, the application of the rules relating to word structure means that a comparatively large number of RV words, more than 300, are open to an investigation of the Indus language of the Panjab. Kuiper (1991) lists 383, to which some

---

[8] The language of the people called Nihāl or Nahāl were first detected as unique by R. Shaffer 1940; they are found in medieval texts (such as Hemacandra's Grammar), often together with the Bhils (*bhilla*), as *lāhala, nāhalaka, ṇāhala* as a mountain/jungle tribe on the Narmadā. Details in Koppers 1948: 23, Shafer 1954: 349 (as the original language of the Bhil), S. Bhattacharya 1957, Berger 1959, Witzel 1999.

may be added,[9] while a few others must be subtracted (Oberlies 1994). Even if we retain, as Th. Oberlies wants to do, "only" 344-358 'secure' non-IE words, and even if we subtract all local non-IA names of persons and places from Kuiper's list, we still retain some 211-250 'foreign' words,[10] -- still a surprising percentage of c. 2% 'foreign words' in this strictly hieratic text, composed in the very traditional IA, IIr, IE poetic style that deals with equally traditional matters of ritual and myth. It is important to notice, at the outset, that the range of these 'foreign' words (Kuiper 1955) is limited to local flora and fauna, agriculture and artisanship, to terms of toilette, clothing and household; however, dancing and music are particularly prominent, and there are some items of religion and beliefs (Kuiper 1955, 1991). Importantly, these loan words *only* reflect village life, and *not* the intricate civilization of the Indus cities, which fits very well with their post-Harappan time frame.

In evaluating these loan words, we should first look out for a possible source in the language(s) actually found in the area since ancient times. For the Greater Panjab and the Indus civilization in general, not only Dravidian and Munda are likely candidates (Kuiper 1991: 39sq.)[11] but also one or more unknown languages, such as Masica's Gangetic "language X" (cf. Kuiper 1991: 1,4; 1955: 137, 1962; Emeneau 1956). In fact, all the four great language families present in modern South Asia: Indo-European, Dravidian, Austro-Asiatic, and Tibeto-Burmese,[12] have left traces in the Vedic texts; and there are others that point to remnant languages, often isolates among the world's 5000-odd languages. They have been mentioned above: Burushaski, Kusunda, and the substrates in Tharu, Nahali, Nilgiri and Vedda languages. Nahali is of particular interest, as we find, below its present form Indo-Aryan appearance, at successively "lower" levels, traces of Dravidian (9%), and Munda words (36%), and at the lowest substrate level some 24% that do not have any cognates in India (Kuiper 1962, 50; 1966, 96-192) and must belong to the oldest level of

---

[9] For example, *akhkhalī-kṛ* RV, an onomatopoetic which is not attested again until today's Nahali *akkal-*(*kāyni*) '(to cry out) loudly in anguish' see *MT* II 17; *ambu* 'a water plant' Up.; *āraṇa* 'a shallow', *ṛjīṣa* 'Soma residue'; also 'an epithet of Indra, Soma, Marut, eagle'; *kakāṭa* 'backside of the head', *lalāṭa* 'front of the head', *kośa* 'pail, box', *camriṣ* 'ladle', *jīri* 'flowing', *tandrayu* 'lazy', *petva* 'ram', *plakṣa* 'fig tree', *badhira(?)* 'deaf', *bīriṭa* 'troop, host', *makṣījā* 'net to catch birds?', *su-maṅgala* 'auspicious', *manu ~ maṇu* AV 'fertile land', *marīci* 'beam of light', *mukha* 'mouth, head', *musala* 'pestle', *mūla* 'root', *yāduri* '? (in conn. with hip, buttock), *valg(?)* 'to jump, boil up', *vikaṭa(?)* 'huge, monstrous, *vi-śikha* 'having no tuft of hair', *śepa* 'penis', *sītā(?)* 'furrow', *sīrī* 'weaver?' *sthāṇu(?)* 'trunk of a tree', and probably more.

[10] Oberlies' criticism is written from an IE-centered point of view similar to that of Mayrhofer (*EWA*); however, due to the clear attestation of cultural, ethnical and religious amalgamation visible already in the RV, the existence of such a large number of 'foreign' words must not be minimized in its importance. R.P. Das' nit-picking review of Kuiper's book, 'The hunt for foreign words in the Ṛgveda' (*IIJ* 38, 1995, 207-238) has received a well-deserved, rather scathing reply by Kuiper, "On a Hunt for 'Possible' Objections", *IIJ* 38, 1995, 239-247. Whiek indiviudual words, including te ones listed in thew presnet paper, can and must be viewed critically, it is indeed difficult to understand how the language of the Ṛgvedic Arya can be called 'relatively free from foreign influences' (Oberlies 1994: 347). "Pristine" languages and cultures do not exist.

[11] "...prefixes. They are unknown in Dravidian but were common in Austro-Asiatic. They may also have been characteristic of other Indian languages that have disappeared." (In mod. Munda only some petrified relics remain).

[12] For example, the name of the *Kirāta*, AV+ (see below), *Kosala*, the River *Kosi* (below), words for cooked rice in NIA (*cāmal, cāval,* CDIAL 4749, but cf. PAustro-Thai *Cṣamaq*, Benedict 1990: 175), note also *pipīla* 'ant' and Tib. *p'yi, EWA* II 133; see Witzel 1993.

languages traceable in South Asia (Kuiper finds similarities with Ainu, for further connections see now *MT* 3).

Apart from these languages that have actually survived into modern times, there must have been a score of dialects and languages that have "not made it". A recent case seems to be Kusunda which seems to have died out only in the Seventies, and which has been recorded but very little in the early 19th and in the later part of the past century. But for one paper by Hodgson in the first part of the 19th cent., and some recent small articles by Reinhard and Toba (1969, 1971) we would not have known about this important remnant language.

These facts suggest successive levels of immigration by speakers of the several large language families involved, the spread of their languages (not always and not necessarily involving actual movement of people), and a gradual retreat of the older languages and their speakers into the inaccessible hills and jungles. On the other hand, there also is the successive taking over of the newly immigrant languages by populations which stayed in their old habitat.

Not all of the languages mentioned so far are attested early on, and we can only make *reliable* comparisons with the loans in the RV when we have established a reconstructed pre-form of the words of these languages. To give an obvious example, the modern tribal name of the Munda speaking in Orissa *Saüra* (*Saora)* :: Greek *Sabarai* :: Ved. *Śabara* (AB 7.18). The unattested pre-forms of Munda (**šqawar,* Pinnow 1959) allow comparison with Skt. *Śabara* (AB), while there is no *immediate* one between mod. Oriya *Saora* and Ved. *Śabara,* Skt. *śabara* 'hunter'.

The sections below will indicate that we have to reckon, in addition to the substrate languages mentioned earlier, with some unknown languages in the Greater Panjab as well. The inhabitants of a city such as Harappa may very well have been bi- or tri-lingual.

The next question to be answered, then, is how to decide between the languages and language families that were present in S. Asia in Ṛgvedic times. Just as in the case of IE/IA words (see above), it is root and word structure that comes to our aid.

Dravidian word structure is fairly well known now (Krishnamurty 1998; in the sequel ə = long or short vowel).

Drav. root structure: (C)ə(C), thus:

ə	*ē* 'increase'
əC	*oy* 'to drag'
Cə	*kā* 'to preserve'
CəC	*pal* 'tooth', *kāl* 'leg', *nāy* 'dog',

Suffixes have the structure: -C, -Cə, -CCə, -CCCə;
after a root -C the vowels *-a-, -i-,* or *-u* are inserted , thus əC-a-C etc., and CəC-a-C etc.;
(base final -C is followed by *-u,* thus CəC-a-C-u): examples include:

CəC	*kār* 'to be salty' :: CəC-i *kar*-i 'salty to the taste'
CəC-əC	*kar-il* 'pungency', *kāra-am* 'pungency'
Cə-CC	*kapp-u* 'to overspread', (*kap-i* > kavi 'to cover'), *meṭṭ-u* 'to smell' : CəC *mēṭṭ-u* 'height'
	(Tamil also *peyar* > *pēr* 'name', *muka* > *mō* 'to smell'), etc.

Dravidian words thus can take forms quite different from Vedic ones, cf. Ved. CCə-CCəC-Cə *pra-stab-dha-* 'stiff' (from *stambh*) :: Drav. CəC-a-C-u. Still, words such as *pra-mag-a-nd-a* are not covered by Dravidian structure as Drav. lacks prefixes and as *pra-* and similar

initial clusters (CC) are not allowed either. (As will be seen below, the word can only be Munda if it is a Vedic adaptation of *pər-ma-gand- as double consonants are not allowed anywhere).

The present syllabary structure of Munda words (C) ə (C) (Pinnow 1959: 449 sqq.) is similar to the Dravidian one, thus:

ə	*i*	'was'
əC	*aṅ*	'to open the mouth'
Cə	*ma*	'mother'
Cəə	*hai*	'fish'
CəC	*loṅ*	'pungent'

(in Sora also: CCəC, CCəCC = C'ə'C or C'əNC)

Munda word structure:

CəCə	*gara*	'river'	CəəC	*duar*	'door'
CəCəC	*merom*	'goat'	əCCə	*imta*	'then'
əVVəC	*apro'b*	'wing'	CəCCə	*siṅgi*	'sun'
CəCCəC	*sengel*	'fire'	(rarely, larger words)		

However, the oldest (and thus, more or less contemporaneous with RV) word structure of Munda was:

(C)ə(C)
Cə-CəC   (= CəCə'C) and   CəC-   Cə'C   (*pər*- etc.)
CəC-əC   CəC-   Cə'C-əC

Thus, while Drav. and Munda share some of the same structures, e.g. *CəC, CəCəC,* Munda words can be clearly distinguished as *Cə*- in *Cə-Cəc* is a prefix, something that does not exist in Drav.; and while *CəCəc* may exist in IE/IA (even with a prefix Cə-), normally, *CəC*- will be the root and -*əC* a suffix. Word structure, especially when combined with etymology, thus allows to analyze many of the 'foreign' words in the RV at a glance.

This procedure also allows to identify and to separate words such as the following both from Drav. and Munda: *akhkhalī-kṛ* 'to speak haltingly' or 'in syllables?', cf. Nahali *akkal-(kāyni)* '(to cry) loudly in anguish' (*kāyni* < Skt. *kathayati* 'to tell'), *Ikṣvāku, pippala* 'fig tree', *kṣviṅka* 'vulture?', *kṣumpa* 'mushroom', *plakṣa* 'fig tree', *vrīhi* 'rice', *stega* 'reed?' (cf. Masica 1979).

§ 6. Three levels of loans in the RV

To begin an evaluation of the languages that have preceded Vedic Sanskrit the Greater Panjab, we could take the oldest text, the RV, at face level as a unified text of more or less the same time period. This has indeed been done so far; one just distinguished between its older layers and a very young layer (book 10). However, the various layers in the older parts of the RV must be investigated carefully before we can use these materials cogently. For, this pre-iron age text could have been composed *any time* between the end of the Indus civilization (c. 1900 BCE) and the beginning of the iron age (c. 1200 BCE): Indus cities are *not* mentioned in the hymns which know only of small forts; however, (Indus) ruins appear in the RV (*armaka, vailasthāna,* Falk 1981) and they are clearly attested and localized for the Ghaggar-Hakra area (PB 25). The first iron age texts are post-Ṛgvedic (AV, YV mantras, Witzel 1995, 1997).

In recent years, it has become clear that the various lineages of chieftains and poets mentioned in the RV constitute a period of only some 5 generations as the center point of the RV; this is preceded by an (unclear) number of chiefs, poets and their hymns, and it is followed by some more generations[13] (Witzel 1995, 1997). According to still more recent investigations, the later RV period may comprise only some three generations while the fourth one (that of the great Kuru king Parikṣit and his priest Tura Kāvaṣeya) is already of iron age time. This would bring down the bulk of the RV to only c. 1450-1300 BCE at the earliest, which incidentally fits well the Old Indo-Aryan data from the Mitanni area and their agreement with the Hittites, at c. 1380 BCE. (The preservation of IIr *azd* for RV *ed(h)* indicates that the Mitanni form of IA slightly preceded the language of the RV. However, it may have kept some conservative elements from the time of initial contact between the Mitanni Hurrites and the early Indo-Aryans, somewhere in N.W. Iran, an acculturation that preceded Mitanni settlement in N. Iraq/Syria).

We have to reckon, thus, with three periods of RV composition. The beginning of period I is difficult to establish. If we think of it as being after the demise of the Indus civilization at c. 1900 we would have a long period of some 350 years for the older RV. It is more likely that the appearance of the first horses in South Asia, at Pirak and the Kachi Plain at c. 1700 BCE, signals -- if not the arrival of tribes speaking Indo-Aryan -- a change to a pastoral society which used horses. Their IA contemporaries are pastoral, use horses (and some camels), horse drawn spoked-wheel chariots, and have a material and spiritual culture that is reflected in the RV. At the present state of our knowledge, 1700 BCE is a more likely figure for the import of the IA language and all the cultural items that come with it (in Ehret's terms, the IA 'status kit', Ehret 1988).

The three stages in the development of the texts of the RV are reproduced below with some rough absolute dating. This will have to be revised occasionally as archaeological knowledge of this period expands, and as the linguistic and philological database is refined, especially within the oldest books, RV 4-6, and with regard to the exact position of book 2 (western, middle RV period) and book 8 (southwestern: Baluchistan?). Especially, we need a new grid of multiple axes: time, place, kings, tribes, poets, clans, grammar, retroflexes, etc. which is under preparation with the help of revolutionary software (J.R. Gardner/Witzel, forthc.). Even now, however, three RV periods can be established, as follows.

1. early Ṛgvedic period[14]: c. 1700-1450 BCE: RV books 4, 5, 6;

---

[13] This account was originally based on the genealogies of Ṛgvedic 'kings' or chieftains (Witzel 1995), and poets (Witzel 1997, not yet published); the two grids overlap and support each other; further evidence comes from the development of RV ritual (Proferes 1999) whose thesis allows to follow its changes from the older RV, through RV 3 - RV 1/RV9 to RV 10; this development coincides with the analysis given above: books 3 & 7 belong to the middle RV; book 1.1-50 is later, and RV 10 latest. The same is found in individual analyses, such as J.R. Gardner's study of the various terms for the self (Gardner 1998), or grammatical study: absolutives and original retroflexes are rare in the "Iranian" book 7 of the newcomer Vasiṣṭha (Witzel 1997, n. 145).

[14] With Indo-Aryan settlement mainly in Gandhāra/Panjab, but occasionally extending up to Yamunā/Gaṅgā, e.g. Atri poem 5.52.17; the relatively old poem 6.45.13 has *gāṅgya,* next to chieftain *Bṛbu* and a few references to W. Afghanistan (Sarayu = Herat River).

2. middle, main *Ṛgvedic period*, c. 1450-1300 BCE: books 3, 7, 8. 1-47, 8.60-66 and 1. 51-191, most probably also 2; prominent: Pūru chieftain Trasadasyu and Bharata chieftain Sudās and their ancestors, and

3. late *Ṛgvedic period*, c. 1300-1200 BCE:

books 1.1-50, 8.48-59 (the late Vālakhilya hymns), 8.67-103, large sections of 9, and finally, 10.1-854, 10.85-191; emergence of the Kuru tribe, fully developed by the time of Parikṣit,[15] a descendant of Trasadasyu.

If the 'foreign words' in the RV are investigated by paying close attention to these three textual levels, some new and rather surprising results are reached (Witzel 1999a,b). Kuiper has studied these loans in some detail in his book of 1991. He has already pointed out that these loans have certain phonetical features unusual in Sanskrit, and that they have certain typical prefixes and suffixes.

§ 7. Indus words from the Panjab

A large group of them share, as many scholars have agreed, the characteristic of having unusual prefixes, that is prefixes not known from IA.

It is of course, difficult to establish what is a prefix, what is the root and what are suffixes if one does not know the language in question. In English, for example, we may list a number of words that have the old, now unproductive prefix *for(e)-* or *be-*: *for-get, fore-go, for-feit, fore-close,* etc. or *be-dazzle, be-friend, be-get, be-have, be-hold, be-lie, be-long, be-moan, be-seech, be-tray, be-ware*, etc. To a modern speaker the meaning of *be-* is unclear, but as we know the roots (words) *get, have, hold,* etc. we can easily divide up these words as done just now. Someone not knowing English may add some wrong word divisions such as *be-amy, be-anie, be-ad, be-auty, be-gum, be-gonia, be-luga, Be-nares, be-neficient, be-nign, be-ryl, be-vy.* If the number of words is large enough, however, it will be come apparent, even if the language in question is unknown, that the words of the second list do not contain a root/word as second part, and that the analysis must be wrong.

This condition is more difficult to meet with when studying the 300 odd loan words in the RV. We have only incidental words, not their roots or their roots with other prefixes. Cases such as *tila* AV 'sesame' and *jar-tila* KS 'wild sesame' are rare, but they indicate that we are on the right track with dividing words into prefixes and roots. Interestingly, the word is linked to *taila* 'sesame oil' AV, *til-vila* 'fertile' RV 10.62.7, *til-piñja* AV, *til-piñjī* AV 'infertile sesame', *til-vaka* ŚB 'name of a tree, Symplocos racemosa', *tailvaka* MS 'stemming from that tree', and Munda: Santali, Mundari *i-til* 'grease, fat, rich', Sant. *til-min* 'oil seed' (Kuiper 1955: 157).[16]

Even the skeptical Indo-Europeanist M. Mayrhofer (1986-1996), who in his *EWA* now is always looking for IE/IIr/IA etymologies, therefore speaks of prefixing language(s) in the RV substrate, while Kuiper (1948, 1955, 1991) identifies the donor language directly

---

[15] Strictly speaking, Parikṣit does not occur in the RV but only in an appendix, RVKh; however, the Kuru make at least one appearance, in the name *Kuru-śrávaṇa* "glory of the Kuru."

[16] Further connections: Sumer. *ili* 'sesame', Akkad. *ellu/ūlu* 'sesame oil', and among the Drav. languages only South Drav. with *el, eḷḷu* 'Sesamum indicum' (D. Bedigian 1985); cf. also Ossetic *tillaeg, tyllaeg* 'corn, grain, seed'. The question remains how *tila* and **il/el* are related. Should we think of a Proto-Indian **t-il-* ?

as Proto-Munda. As willl be explained below, I rather prefer the provisional name, *Para-Munda* (or simply, *Harappan*). Indeed, a number of prefixes in the loan words of the RV look exactly like those prominent in Munda (and also in the rest of the Austro-Asiatic family, which includes Khasi in Assam, Mon in S. Burma, Khmer in Cambodia, Nicobar, and some other smaller S.E. Asian languages). Furthermore, even the meaning of these prefixes seems to agree: of interest in the RV substrate are especially the prefixes *ka-, ki, kī-, ku-, ke-*, which relate to persons and animals (Pinnow 1959: 11; cf. p. 265 §341a). In the rest of Austro-Asiatic, the 'article' of Khasi (masc. *u-*, fem. *ka-*, pl. *ki-,* cf. Pinnow 1959: 14), is of interest.

Consequently, e.g., the following prefixes found in the RV are important (arranged historically, according to the three levels of the RV; for a detailed discussion see Witzel 1999a, b).

• ka- : *kabandh-in, kavandha* 'barrel' 5.54.8, 9.74.7; *kākambīra* 'a certain tree' 6.48.17; *kavaṣa* 'straddle-legged' 7.18.12; *kakardu* 'wooden stick', 10.102.6; *kapard-in* 'with hair knot' 10.114.3;

•ki-: *kimīd-in* 'a demon', 10.87.24; 7.104.2, 23 (late); cf. *śimida, śimidā* 'a demoness'; *kilāsa* 'spotted, leprous', 5. 53.1; *kilbiṣa* 'evil action', 5.34.4, 10.71.10; *kīsta* 'praiser, poet' 6.67.10, 1.127.7, to be read as [*kisətāsaḥ*]; to be compared with RV *śīṣṭa* 8.53.4 with var. lect. *śīṣṭeṣu, śīrṣṭeṣu, śīrṣṭrēsa*, Sanskritization of **k'īsəteṣu*; *kīkaṭa* 'a tribe' 3.53.14, cf. Sant. *kaṭ-* 'fierce, cruel', or common totemic tribal name (like *Mara-ṭa* PS : Munda *mara'* 'peacock' IA *Matsya* 'fish', *Kunti* 'bird') ~ Sant. *kaṭkom* 'crab'?; *kīja* 'implement, spur?', 8.66.3; *su-kiṃśu-ka* 'a tree, 'Butea frondosa' ' 10.85.20; *kiyāmbu* 'a water plant' 10.116.13, cf. Sant. *um* 'to bathe', Khasi *ūm* 'water'; *kīkasā* (dual) 'vertebra, rib bone' 10.163.2; *kīnārā* dual, 'two ploughmen' 10.106.10; *kīnāśa* 'plough man' 4.57.8 (late), see *kīnara*; *kīlāla* 'biestings, a sweet drink' 10.91.14;

•ku-: *kumāra* 'boy, young man', 4.15.7 etc.; note *śi(ṃ)śu-māra* 'Gangetic dolphin', Munda *məndra, mər* 'man'; *kulāya* 'nest' 6.15.16; *kuliśa* 'ax' 3.2.1, 1.32.5, cf. Sant., Mundari *kutam* 'to beat, hammer', Mundari, Ho *kutasi* 'hammer', from **kodeś* in Kharia *khoṇḍe'j* 'ax', Mundari *koṇḍe'j* 'smaller kind of wood ax', with prefix *kon-* and Kharia *te'j* 'to break'; *kuṇāru* 'lame in the arm?' 3.30.8; *kuruṅga* 8.4.19, name of a chieftain of the Turvaśa; cf. *kuluṅga* 'antelope', and the frequent totemistic names of the Munda; *kuśika* name of a poets' clan, RV 2 etc.; *kupaya* 'shimmering?' 1.140.3; *kuṣumbhaka* 'poison gland of an insect' 1.191.15-16; if not one of the common IA animal names in *-bha* (*śara-bha* etc.), then: **ku-śumb(h); kurīra* 'women's hair dress', 10.92.8.

Further, in Munda and in general Austro-Asiatic there also are the so-called 'double prefixes' which are composed of a prefix (e.g. *k-*) followed by a second prefix (mostly *-n-*, Pinnow 1959: 11). The double prefixes, such as *pə-r-, kə-r-/šə-r-* etc.[17] are quite archaic. RV *pra/pər-* etc. stands for modern Munda *po-*, but modern Munda has only a few remnants of them while Ṛgvedic Munda is prolific. They also seem to be more rare in Munda than in Eastern Austro-Asiatic (see Kuiper 1991: 94 *śar-varī, śa-bala*, Pinnow 1959: 143 § 285, Witzel 1999 a,b).

In Ṛgvedic loans, formations containing consonant (C) - vowel (ə) - *r,* such as *Cər-* (and also *Cən-, Cəm-*); the substrate vowel (probably *schwa, ə*) were adapted into Vedic by various vowels such as *ṛ, ur,* etc., (see Kuiper 1991: 40 sqq., 23); due to the interchange of *k*[k']/*ś,* the prefix *śar-/śal-* belongs here as well, as seen is in: *kar-koṭa-ka* RVKh ~ *śar-koṭa* AV. The clearest Vedic case is, perhaps, *jar-tila* 'wild sesame' AV : *tila* 'sesame' AV (cf. *tilvila* 'fertile' RV, Kuiper 1955: 157, *tilpiñja, -ī* 'infertile sesame' AV, *tilvaka* 'a tree'; on Sumer. connections see below). The Ṛgvedic cases include the following:

---

17 Note *karpāsa* Suśruta+, *kārpasa* 'cotton' in late Vedic (Śrautasūtras) > Greek *kárpasos*, Bur. *ɣupas,* etc.*; for -r-* see Kuiper 1948 *ku-r-p-aṛ : kaphoḍa : kūbeḍ < *kurpar, *kapoḍ; kir-bed, bid*, etc; details in Witzel 1999a,b.

śarvarī 'night' 5.52.3, *api-śarvara* 3.9.7, 8.1.29; *karambha* 'gruel' 6.57.2; *sṛñjaya* a name of a person 6.27.7, 4.15.4, *sārñjaya* 'descendent of S.' 6.47.25; *sṛjaya* 'a certain bird' KS and the totemistic names in Munda; *śalmali* name of a tree, 'Salmalia malabarica', cf. RV 3.53.22 *śimbala,* dissimilations of **śamma/śimmal; a-kharva* 'mutilated' 7.32.13; *sṛ-binda* name of a demon 8.32.2; cf. *Ku-suru-binda* TS, PB, ṢB, *Ku-sur-binda* JB and *Bainda* VS 'member of the tribe of the Binds' (probably also the name of the Mountain range, post-Vedic *Vindh-ya*), *Vi-bhindu* RV 8.2.41, 1.116.20, *Vi-bhindu-ka, Vi-bhindu-kīya* JB' (cf. also Vi-darbha, Vi-paś); *karkari* 'lute' 2.43.3 (late), onomatopoetic, cf. Sant. *karkur, gargaṛ, gargor,* etc.; *kalmalīk-in* 'shining' 2.33.8; *khṛgala* meaning unclear: 'staff, crutch, amulet, armor, brush?' 2.39.4; *karañja* name of a demon, 1.53.8, *karañja-ha* 10.48.4; *kṛkadāśū* 1.29.7, unclear meaning, personal name?; *kṛṣana* 'pearl' 1.35.1, 10.68.11, 10.144.2 *ūrdhva-kṛṣana,* 10.144.2, *kṛśanā-vat* 1.126.4; *śaryāta* name of a person, 1.112.17, *śāryāta* 1.51.12, 3.51.7; *karkandhu* later, a tree name 'Zizyphus Jujuba', but personal name in RV 1.112.6; *kārotara* 'sieve, filter' 1.116.7; *kṛpīṭa* 'bush, brush' 10.28.8; *khargalā* 'owl' 7.104.17 (late).

In the same way, other prefixes can be discussed, such as *jar, tar, nar, par, bar, śar, sṛ* = *[jər, tər],* etc.: *jarāyu, jarūtha* (cf. also *jar-tila : tila*); *taranta, tarukṣa, tṛkṣi, tṛtsu, nārmiṇī,* epithet of a fort; *nārmara* 'area of or the chief of *Ūrjayantī*'; *parṇaya, parpharī-ka, parśāna; prakaṅkaṭa* (cf. *kaṅkata, katkaṭa, kakkaṭa*), *prakala, parpharvī, pramaganda, pra-skaṇva, pharva-ra, phāriva; pṛthi, pṛthī, pṛ-dāku, barjaha;* finally, adaptations of [-*ər*] as above in [*kər*]: *tirindi-ra, turīpa, turpharī*[18]*, turva/turvaśa, turvīti, tūrṇāśa, sūrmī;* (cf. also *Nār-ṣada* RV, *Nār-vidāla, Nār-kavinda* PS and **ku-bind* in: Ved. *ku-sur(u)-binda, bainda, vi-bhindu, vi-bhindu-kī-ya*).

From among these words, the (designations of) persons in *k-, kər-/ ś-* are attested in, or are likely of (Para-)Munda origin: *kimīd-in/śimida, Kīkaṭa, kīnāra, kīnāśa, Kīsta/śīṣṭā, kumāra, Kuruṅga, Kuśika, karañja, kṛkadāśū* (1.29.7, note *kuṇḍṛṇācī* 1.29.6), *Śaryāta, Sṛñjaya.* It would mean that some of the Ṛgvedic "aboriginal" tribes such as the Kīkaṭa living in the Kurukṣetra area west of Delhi would have been speaking Para-Munda.

Of special interest are the snake demon *Kar-koṭa* RVKh */Śar-koṭa* AV and the name of *Pra-maganda* RV 3.53.14, a chieftain of the clearly non-Indo-Aryan *Kīkaṭa* (RV 3.53.14), who lived south of Kurukṣetra. The prefix *pər-* may mean 'son of' (Kuiper 1991: 43), *ma-ganda* may contain the old, now unproductive Munda prefix *ma-* that indicates possession, and *gand* may belong to Munda **gad/gaḍ, ga-n-d/gaṇḍ* 'water' (Pinnow 1959: 351), thus 'son of the water/river/canal(?) posessor, water-man'. It appears in the river names *Gaṇḍa-kī, Gaṅgā* (Witzel 1999), in W. Nepali *gāḍ* (Witzel 1993), the *Gandhina* people (Mbh), and apparently also in the country of *Ma-gadha* 'having water/rivers' (with Sanskritization > *dh;* also seen in **ganda > Gandhā-ri,* the tribe/area on the upper Indus/Kabul R.) The *Kīkaṭa* have either the typical 'tribal' suffix *-ṭa* or the old Austro-As. plural prefix *ki-*.

Many substrate names of persons, tribes and rivers come exactly from the areas where Indus people are to be expected, from their new settlements (J. Shaffer and D.A. Lichtenstein 1995: 139) in the eastern Panjab, in Haryana (Kurukṣetra), and east of there. Surprisingly in this area, the eastern Panjab and Haryana, even the middle-Ṛgvedic rivers are still designated by non-Indo-Aryan names, such as *Śutudrī* and *Vipāś* (RV 3.33).

---

[18] Later, in medieval Mīmāṃsā and Buddhist texts, the poem "*jarbharī turpharī*" RV 10.106 is regarded as non-sensical; it has many strange and foreign words: *pharvara, śakuna, vaṃsaga, pūṣarya, śimbata, śātapanta, saparya(?), purīṣa, jarbharī, turpharītu, turpharī, parpharīka, maderu, jarāyu, marāyu(?), kharamajra, kharajru, parpharant, jaṭhara, saneru, turphari, phāriva, āraṅgara, mana-ṛṅga?, kīnāra;* see now Kuiper (forthc.) on its 'bilingual' poet.

The cluster of local names in the Sutlej-Sarasvati area (Witzel 1999) includes the following: *Śutudrī* 'Satlej', (note the later popular etymology *Śatadru* 'running with a hundred streams'), *Vipāś* 'Beas' < **vipāž/*vibāl,* (cf. *Vibālī* RV 4.30.11-12), and even the Sarasvatī, which is called *Vaiśambhalyā* and by many variants, always a sign of foreign origin, (cf. also *Viśpalā*?) < **višambaž, *višambāl. Both words, *vi-pāž, *vi-šam-baž,* probably are formed with the prefix *śam/k'am-* as seen in *Śam-bara, Kam-boja,* from **(vi)-šam-bāž* (cf. *Vi-darbha, Vi-bhindu,* and note also the popular etymology *vi-śambala* 'having widespread blankets'). This cluster is further strengthened by the local names of the land of *Tūrghna,* north of this region, and *Khāṇḍava* with its suspicious cluster *-ṇḍ-* (K. Hoffmann 1941), south it. This 'southern' land is inhabited by the *Kī-kaṭa* under their chieftain *Pra-maganda.* Note also the appearance of Pinnow's *u-*suffixes in 'foreign' names of the same region, Kurukṣetra, e.g. *Khāṇḍava, Kārapacava, Naitandhava* (Pinnow 1953-4).

Another cluster is found further west in the Greater Panjab, with place/river names such as *Gandhāri, Kubhā, Krumu, Kamboja, Gandhāri* RV, OPers. *Ga^{n}dāra.* The last two forms can be compared with Munda **ga(n)d* 'river' (see above). The *Kamboja* (AV, PS) settled in S.E. Afghanistan, cf. OP *Ka^{m}bujīya* (or *Kambaujīya*?) 'Cambyses', and Ambautai (Ptolemy, Geography 6.18.3) without the prefix, a change typical in Munda names (*Aṅga : Vaṅga, Kaliṅga : Teliṅga; Kulūṭa : Ulūṭa,* etc.).

While Mundas that far west cannot be excluded (Kuiper 1991: 39), it may be asked, how far Para-Munda or Austro-Asiatic speakers indeed extended westwards during and before the RV period. Based on the *present* distribution of the Munda languages in Eastern India (Bihar, Orissa, W. Bengal) and in Central India on the River Tapti (NW Maharastra and Madhya Pradesh) some regard it as impossible that the Munda could ever have settled in the Panjab. Kuiper (1991: 39, cf. 1955: 140, 1948: 8) thinks differently. "The occurrence of Munda borrowings in the Rigveda raises some questions. According to some scholars Munda was never spoken west of Orissa, Bihar, Madhya Pradesh and eastern Maharashtra (Burrow 1958, quoted by Southworth 1979, 200). The obvious occurrence of Old Munda names in the Rigveda points to the conclusion that this statement should be revised or that some parts of the Rigveda (e.g., book VIII, see p. 16) stem from eastern parts of North India" (Kuiper 1991).

Indeed, the cases listed above indicate a strong Para-Munda or Austro-Asiatic substrate in the Panjab, and there are some indications that point to Munda influence in E. Afghanistan (*Śambara, Kamboja*). Kuiper's question should therefore be answered in the following way: There are clear indications of (Para-) Munda in the Greater Panjab. As for RV 8, it rather points to Greater Baluchistan (see below).

§ 8. Prefixing languages in South Asia

Is the Indus language, then, a kind of Proto-Munda? Against an identification of the loan words discussed above with Munda or the larger Austro-Asiatic family may indeed speak, first of all (Kuiper 1991), that the RV substrate does not have infixes like Munda. However, *-n-*infixes can perhaps be adduced in *ka-bandha/ka-vandha* 'headless rump', *kar-kandhu* 'name of a tree, Zizyphus jujuba', *gandhā-ri* 'name of a tribe in N. Pakistan', *pra-ma-ganda* 'name of a chieftain of the Kīkaṭa non-Aryans', *śa-kunti* 'bird' < PMunda **ša-*

konti'd (cf. Epic *Śa-kuntalā, Ved. Kunti, ~* Munda *kon-ti'd* 'bird'), *sṛ-binda, sṛjaya* KS 'a bird' / *sṛñjaya* RV 'name of a man, tribe' and in post-RV, e.g., *ku-sur(u)-binda, bainda, vi-bhindu, vi-bhindu-kī-ya* 'name of a tribe'. Unfortunately, we do not have many words related to each other such as *tila, jar-tila* which could be the basis for *n*-infixes.

Languages with prefixes stand out in S. Asia. Apart from IA, only modern Burushaski and archaic Tibeto-Burmese have a few of them. However, they occur in Burushaski only in a few secondary verbal formations (Berger 1999: 103sqq.: pronominal, negation, absolutive and intransitive), and in Tib.-Burm. the six prefixes (*s-, r-, b-, g-, d-, m-*) which were still seperable in PTib.-Burm., have a number of secondary verbal functions (*s-* for causative, intensives, directives, *b-* perfect, *m-* middle voice, *a-* pronominal, etc.); they also function as class signs (*s-* for body parts and animals, etc.); others (including also some of the preceding prefixes) and *r-* have still unclear functions (Benedict 1972: 103sqq) in some archaic nouns. In short, these prefixes are neither as numerous as the (Para)-Munda ones nor are many of them to be expected in the nominal and adjectives of the 300 odd RV loan words. In addition, neither language seems to have been prominent in the prehistorical Panjab as reflected by Vedic texts; further, the RV loan words with prefixes do not correspond to the Tib.-Burm. and Bur. ones in structure.

Remains (next to the use of prefixes in Vedic) one *known* typical prefix language: Austro-Asiatic or its Indian sub-family, Munda. Munda speakers are concentrated now in S. Bihar and Orissa, with one outlier, Korku ore Kurku, on the border of Maharashtra and Madhya Pradesh, on the Tapti river. The substrates of Nahali spoken in that region (Kuiper 1962, *MT* 3) show that Munda has been spoken in the area for a long time: Nahali, now an IA language, has below this deceptive surface, a Dravidian, and below this a Munda level (and, at the lowest level, some 25% of an unknown substrate). This clearly points to the possibility of early (Para-)Munda settlements further west than Munda speakers are found now.

The Ṛgvedic substrate words from a prefixing language may be a very early form of Munda (or another variety of Austro-Asiatic) which still used prefixes actively, such as the eastern Austro-As. languages (Mon, Khmer) . Further, the infixes may have developed from prefixes which had found their way into the root (Pinnow 1959: 15). Among these, one can include 'double' prefixes such as *kə-r-*, *šə-r-*, *pə-r-* etc. (Pinnow 1959: 11). Consequently, if the Ṛgvedic Para-Munda is a somewhat deviant form of Austro-Asiatic, it represents a *very* old stage of this language family indeed. In that case, this Ṛgvedic western Austro-Asiatic would stand next to Munda and eastern Austro-Asiatic (Mon-Khmer).

Finally, it should also be taken into account that Northern and Southern Munda differ from each other in many respects, the southern version usually being more archaic (Zide 1969: 414 sq., 423); unfortunately it is much less known. This difference, as well as the shift of Munda from a prefixing language with mono-syllabic roots to one working, in typical South Asian fashion, with suffixes, may have been influenced by or it may even have been due to a unknown north Indian substrate such as Masica's "Language X". (Once this has been explored beyond Masica's materials, one should investigate its possible influences on E. Himalayan Tibeto-Burmese and on Northern Munda.)

In other words, the western Austro-Asiatic visible in the RV loans, may have been another type of Austro-As. (therefore, I chose the term Para-Munda, Witzel 1999), -- another

sub-family of the great Austro-Asiatic family which stretches from the Greater Panjab to Vietnam and from N. Burma to the outskirts of Singapore.

While it has been argued by some that speakers of Munda (or Austro-Asiatic) could not have lived as far west as the Panjab, we must take the RV loan words seriously (cf. Kuiper 1991: 39sq, 1955: 140, 1962: 14, 51, 1991: 39sq; Hock 1975: 86 sqq). The c. *300 words* in the RV constitute, after all, *the oldest recorded language in the Panjab*. It must be underlined that, just like an ancient inscription, these words have not changed since the composition of these hymns c. 1500 BCE, as the RV has been transmitted almost without any change, i.e. we know exactly in which limited cases certain sounds - but not words, tonal accents, sentences-- have been changed. The modern oral recitation of the RV is a *tape recording of c. 1700-1200* BCE, and as that, of the oldest Austro-Asiatic that we have.

If, on the other hand, a relationship with Munda and Austro-Asiatic could not be confirmed by a larger number of obvious etymologies, a minimal position would be to define the c. 300 non-Dravidian loan words in the RV as coming from an unknown, prefixing language of the Greater Panjab, which might be called, for lack of a self-designation, after its prominent geographical features, the Gandhāra-Khāṇḍava or perhaps better, Kubhā-Vipāś (Kabul-Beas), or simply, the Harappan language.

§ 9. Other local substrates

Apart from the Para-Munda *substrate* and the new Dravidian *adstrate* (see below), we can observe that there is some evidence for the existence other ancient South Asian language in the early Ṛgvedic period. A few words which seem to agree more with the U.P. "Language X", notably the words with double consonants (geminates) (fig.13). These appear in some of the Vedic texts with different combinations of consonants: In the RV, we find *pippala* 'fig tree' RV, but -*pp*- is replaced by -*ṣp*- in *piṣpala* AV, *su-piṣpala* MS (Kuiper 1991: 61); cf. also *guggulu* AV, PS : *gulgulu* 'bdellium' KS, TS, *katkaṭa* TS, *kakkaṭa* 'a certain bird' PS 20.51.6, KSAśv (cf. also (*pra-*)*kaṅkata*); *akhkhalī-kṛ* 'to cry out, to bleat' (with unusual double consonant *khkh* for normal *kkh*), which can now be compared with Nahali *akkal(-kayni)* '(to cry) loudly in anguish' (*kāyni* < Skt. *kathayati* 'to tell'); cf. also *ciccika* 'a bird', *aśvattha* : *aśvatha* 'fig tree, name of a person'. Indeed, Language "X", an ancient substrate in Hindi (Masica 1979: 135), indicates several such geminates: Hindi *kaith* < Skt. *kapittha* CDIAL 2749 (Mbh), *piplī/pīplā* < *pippala* (RV), etc. (details in Witzel 1999 a,b).

In short, the Panjab is an area of a Pre-Ṛgvedic, largely Para-Munda substrate that apparently overlays a still older local level. Since no traces of the supposedly Dravidian language of the Indus civilization (Parpola 1994) are visible in the early RV (see below), the people who spoke this language must either have disappeared without a trace, or, more likely, the language of the Panjab was Para-Munda already during the Indus period (2600-1900 BCE). Therefore, the most commonly used language among *the languages of the Indus people*, at least of those in the Panjab, must have been *Para-Munda or a western form of Austro-Asiatic*.

The caution expressed here will be further qualified by two features to be dealt with in the next sections: the absence of Dravidian loans in the *early* parts of the RV (level I) and a still older substrate in the same level of the RV. When I began, in the summer of 1998, a new

survey of the Ṛgvedic loans as found distributed in the three levels of the RV I did not expect this result. It came as a total surprise as it goes against the general assumption of a strong Drav. substrate in northern South Asia.

§ 10. Dravidian loan words in the Panjab

The absence of old Dravidian loans (in RV stage I), alluded to several times in the preceding sections, comes as a real surprise, as it always has been assumed by many scholars, not to speak of the general public, that the Panjab was settled by speakers of Dravidian language(s) during the Indus period. However, a point of caution could have been raised as early as 1955 when Kuiper undertook a detailed stayed of the loan words in the RV. Though not expressly stated by Kuiper, the large number of agricultural words that have no Dravidian explanation could have indicated even then that the language of the Indus people cannot have been Dravidian (cf. also Southworth 1988: 663). Their successors, the Indo-Aryans, preferred to tend their cattle and they spoke, like their brethren in spirit, the Maasai, about their sedentary non-Indo-Aryan neighbors in southern Kurukṣetra in this fashion: "what is the use of cattle among the Kīkaṭa?" (*kím te kṛṇvanti Kíkaṭeṣu gávaḥ*, RV 3.53.14).

For comparisons of the RV loan words with the (Proto-)Dravidian spoken at the time of the RV, we are limited to Burrow-Emeneau's *DEDR*, and a few lists from old Tamil texts, but scholars usually work directly with Tamil, Kannada, Telugu (etc.) comparisons; a reconstruction of Proto-Drav. forms is but rarely given. This is a serious hindrance in making connections between PDrav., and RV words transparent, though the present reconstruction of PDrav. looks very much like the forms found in (old) Tamil. A detailed study of the supposedly Dravidian words in the Ṛgveda has been undertaken elsewhere (Witzel 1999a, b). However, many of the 26 words attested in the RV that Burrow (1945, 1946, 1947-48, 1955, cf. Southworth 1979 sqq.) originally listed as Drav., as well as those added by Southworth (1979) and Zvelebil (1990) cannot be regarded either as *early* Dravidian loans or as Dravidian at all. Even if one would take *all* of them, for argument's sake, as Dravidian, only the following occur in *early* Ṛgvedic: *kulāya-* 'nest' 6.15.16, *phalgu/phalgva* 'minute', 4.5.14, *āṇi* 'lynch pin' 5.43.8, *ukha(-cchid)* '(lame in the) hip' 4.19.9, *bala* 'force' 5.57.6, 5.30.9. Even these words can, however, no longer securely be explained as Dravidian.

> • *kulāy-in* 'nest-like' 6.15.16, cf. *kulāyayat-* 7.50.1; from Drav.: CDIAL 3340, cf. *DEDR* 1884 Tam. *kuṭai, DEDR* 1883 Tel. *gūḍa* 'basket', but Drav. *-ḍ- >* Ved. -l-?; -- *EWA* I 373 'not clear', comparing N.Pers *kunāṃ,* E. Baluchi *kuδām < kudāman,* with the same kind of problems; -- 'foreign word', Kuiper 1991: 14. -- Nevertheless, Palaniappan (pers. comm. Sept. 1999) derives it from Drav. *DEDR* 1821 Tam. *kuḻāvu, kulavu* 'to stay close together', *kuḻumu* 'to collect in large number', esp. from the participle *kulāya.*
> • *ukha* 'pan, hip' in *ukha-chid* 'breaking the hip, lame' 4.19.9, cf. MS 4, p. 4.9 *ukhá* (dual) 'hips'; *DEDR* 564 'particular part of upper leg' : *ukkam* 'waist' Tulu *okka* 'hip'; for sound change Drav. *k:* Ved. *kh,* s. Kuiper 1991: 36, cf. 1995: 243; however, *EWA* I 210 compares Latin *auxilla* 'small pot', Lat. *aulla* 'pot' (Pokorny 88), but declares 'not sufficiently explained'. If indeed Dravidian, and as RV 4.19 is not a late hymn, this word and *kulay-* might be the oldest Drav. loans in Vedic (RV I).
> • *aṇi* 'axle pin' is not Drav.; even in *DEDR* it is listed as a loan from IA, which can be the case (as it occurs as early as RV), but the word itself definitely is local and not from IA or IIr.

• *bala* 'strength, force', which is found virtually all over the RV (RV 1,3,5,6,7,9,10), may be derived from IE **belo-*, cf. Latin *de-bilis* etc.; it is otherwise found inside IIr. only in Ossetic/Sarmatian. However, as Kuiper 1990: 90, points out, IE (initial) *b-* is rare; -- now, against Drav. origin Burrow, see *EWA* II 215; indeed, while one may point to *DEDR* 5276 Tam. *val* 'strong', PDrav. *v-* should not turn up as Vedic *b*; Kurukh *balē* 'with the help of', Brahui *balun* 'big' both have *eastern* North Indian *b-* as opposed to the more western dialects (and indication that Brahui comes for a more eastern location, see above, n.1)

In short, from the few so far unassailed Drav. etymologies which remain with some degree of probability, *ukha* and *bala* have also be explained as IE, and *āṇi* seems to be a local word that the *DEDR* explains as a loan from IA(!). The Dravidian presence in the older RV is rather *phalgu* 'minute'. Indeed, these few village type words would indeed constitute a strange legacy of the c. 700 years of the great Indus civilization, had it been speaking Dravidian.

In sum, as has been repeatedly mentioned, there are no traces of Dravidian language in the Panjab until the middle period of the RV, around c. 1450 BCE, -- not even of the supposedly Dravidian speaking traders and rulers of the Indus civilization; however, a few Drav. loan words suddenly appear in the RV texts of level II (books 3, 7, 8.1-66 and 1.51-191) and of level III (books RV 1.1-50, 8.67-103, 10.1-854; 10.85-191). Even these are fairly rare (and not all of the following (note those with question marks) have stood the test of being questioned as Dravidian (see detailed discussion in Witzel 1999a,b).

From the middle period (RV II: books 2,3, 7, 8) stem:
> *kuṇāru* 'lame in the arm?' or name of a person 3.30.8; *mayūra(?)* 3.45.1, *phala(?)* 'fruit' 3.45.4?; *ukhā(?)* 3.53 'pan, hip' (late);
> *kāṇa(?)* 'one-eyed' 7.50.1, *kulāya-* 'nest' 7.50.1, *kulpha* 'ankle' 7.50.2;
> *kavaṣa* 'straddle legged' 7.18.12, *daṇḍa* 'stick', 7.33.6 (late); *kuṇḍa-(?)* 'vessel' 8.17.13, *mayūra(?)* 'peacock', 8.1.25, *naḍa/naḷa(??)* 'reed' 8.1.33, *kūla* 'slope, bank' 8.47.11.

From the later RV (RV III) come:
> *ulūkhala?* 'mortar' 1.1.28; *bila* 'hollow' 1.11.5, 1.32.11; *khala* 'treshing floor' 10.48.7; *kaṭu(ka)* 'pungent' 10.85.34, *a-phala-(?)* 'without fruit', 10.71.5, *phalinī(?)* 'having fruits', 10.97.15, *phala(?)* 'fruit' 10.146.5, *kūṭa(?)* 'hammer' 10.102.4, *kāṇa(?)* 'one-eyed' 10.155.1, *piṇḍa?* 'ball, dumpling' 1.162.19, *ukhā(?)* 'pan, hip' 1.162.13,15; *phāla?* 'plough share' 10.117.07, *piśāci-* 'demon' R V 1.133.5, *budbuda(?)* 'bubbling' 10.155.4 *vriś(??)* 'finger' 1.144.5; *bala(??)* 'strength' $1^4,3^2,5^1,6^1,7^1,9^1,10^{10}$, *garda-bha(??)* 'donkey' RV 1.23.5, 3.53.23 (late, next to the artificial *rāsa-bha* 'donkey'!), RV Vālakhilya 8.56.3 (late) *vāra(??)* RV 1, 8, 10, etc.'tail, hair'.

Even if we accept all the doubtful cases listed above, the relatively small number of Drav. words, compared to those from Para-Munda, surprises. Zvelebil, echoing Emeneau writes: "We end, then with a small, but precious handful of Vedic forms for which Dr. etymologies are certain and acceptable as may be expected in this field of areal linguistics, adding, though that no chronology of the borrowings is possible" (Zvelebil 1990: 81; similarly Parpola 1994: 168). This is correct only for the (middle and) later RV which can now be localized in the Greater Panjab and dated to the immediate pre-iron age period (c. 1450-1200 BCE). At best, one can speak of a few isolated words which have been taken over into the RV, which indicates an adstrate rather than a substrate.

Apparently, speakers of Indo-Aryan *did not* get into contact with speakers of (Northern) Drav. as soon as they entered South Asia (as is commonly thought, see Burrow 1955, 1958, cf. Zvelebil 1990: 46.), but only much later, when IA speaking IA tribes were already living in the Panjab and on the Sarasvatī (Ghaggar-Hakra) and Yamunā (Jamna).

Apparently, Dravidian speakers began influencing the Panjab *only at this moment in time* (cf. Allchin 1995: 31sqq.). *Consequently, all linguistic and cultural deliberations based on the early presence of the Drav. in the area of speakers of IA, are void* or they have to be reinvestigated. One may echo Thieme here who answered to Emeneau's statement "the assumption that the language of the Indus valley documents was Dravidian clearly is not fantastic" (*Proceed. Amer. Phil. Soc.* 98, 283) with "that the assumption that it was NOT is clearly not fantastic either" (*Language* 31, 439).

An alternative solution could be to suppose that only the Indus upper class such as traders (cf. *vaṇij* 'trader?' RV 1.112.11, 5.45.6, AV, (*pra-*)*vāṇa* 'trade?' 4.24.9, see Kuiper 1955: 168) and administrators of the Indus Civilization was composed of Dravidian speakers (Parpola 1994, Fairservis in: Southworth, 1979: 208, 228; contra, Hock 1975: 87f., Southworth 1988: 663), and that in consequence, the Indus documents should be read as Dravidian. However, one would then expect, after some 700 years of a flourishing Indus civilization, obvious cases of bilingualism and its reflection in the RV loan words. They should constitute, especially for the oldest RV, a large bulk among the c. 380 'foreign' words. Yet, such loan words from the fields of trade, handicraft or state organization do *not* appear. For example, *Paṇi* '(rich) foreigner, demon' cannot be connected with the meaning 'trader' inside the RV, and the possibly Drav . root *paṇ* 'to barter' appears first only in (post-Ṛgvedic) KS, *pra-paṇa* 'trade' AV, *prati-paṇa* 'exchange' (see *EWA* II 69, *DEDR* 3884, details in Witzel 1999 a,b). In short, even if Drav. had been the traders' language, it remains unexplainable why Drav. influence is only seen in the middle and late RV as well as later one (AV+). The reason cannot be, as van Driem (1999, appendix p. 2, quoting agreement with Parpola) supposes, that the oldest RV hymns were still "composed in more northerly areas, perhaps as far north as modern Afghanistan." (Parpola forthc.)

On the contrary, even the oldest books of the RV (4-6) contain data covering all of the Greater Panjab: note the rivers *Sindhu* 4.54.6, 4.55.3, 5.53.9 'Indus'; *Asiknī* 4.17.5 'Chenab', *Paruṣṇī* 4.22.3, 5.52.9 'Ravi'; *Vipāś* 4.30.11 (~Vibālī) 'Beas'; *Yamunā* 5.52.17; *Gaṅgā* 6.45.31 with *gāṅgya* 'belonging to the Ganges', but also the *Sarayu* 5.53.9, 4.30.18 'Herat River', and the *Kubhā* 5.53.9 'Kabul River'. (Interestingly, one and the same poet, Śyāvāśva, refers to both the river Yamunā on the one hand (5.52.17)and the *Rasā* (on the upper Indus), *Anitabhā, Kubhā, Krumu* (Kurram), *Sindhu* and even the *Sarayu* on the other (5.53.9); this covers all of the Greater Panjab and Afghanistan and indicates a large network of intertribal relations.

In sum, *early* Dravidian influence on the Ṛgvedic Panjab can be excluded, but the relatively small, then increasing influence on the middle and later RV must be *explained* (cf. Kuiper 1991: 7 sq.). This could point to the middle/later Ṛgvedic immigration of Drav. speakers, for example, from Sindh. Such a scenario would agree with the reconstruction of Fairservis (1995), Southworth (1979, 1988: 663, 1990 and McAlpin (1979) of early Dravidian, that is an originally pastoral society that acquired agriculture only in South Asia. Be that as it may, a closer look at the regions of the Indus civilization bordering the Panjab in the South, especially Sindh, it advisable.

§ 11. The situation in Sindh

Unfortunately, the evidence of the RV is, by and large, limited to the Greater Panjab, and Sindh and Baluchistan appear only very marginally even in the post-Ṛgvedic texts. Details are very sketchy and cannot be discussed here in detail (for this, see Witzel 1999a, b) and results must remain very tentative for the moment.

Nevertheless, a number of instances in RV 8 may point to the Baluchi and Sindh areas. The southernmost tribe mentioned in the RV are the *Bhalānas* took part in the Ten Kings' Battle (RV 7.18) and are certainly to be located near the present *Bolān* pass and river near Quetta, but this, like all other southern rivers, is not mentioned anywhere in the RV (south of the Gomatī, the Gomal River).

Indeed, Book 8 has long been connected with Eastern Iran: (K. Hoffmann 1940 = 1975: 1 sqq.) as it has (pre-)Iranian looking names such as *Kaśu* ~ Avest. *Kasu, Kaśu Caidya*, *Kanīta* ~ Scythian *Kanitēs, Tirindira* ~ *Tiridatēs* ~ Avest. *Tīrō.nakaϑßa, Kṛśa* 8.59.3 ~ *Kərəsāspa, Parśu* ~ Old Pers. *Pārsa* 'Persian', *Paktha* 8.22.10 ~ mod. Pashto, Paktho, *Arśasāna, Anarśani* ~ Iran. *əršan*, cf. also the unusual name *Varo Suṣāman*. Their dialect was close to the one that later developed to E. Iranian. Book 8 also knows of camels (*uṣṭra* 8.4.21-24, 31, 46-48, O. Iran. *uštra*, as in Zaraϑ-*uštra*), that are first attested archaeologically in S. Asia in the Bolān area, at Pirak, c.1700 BCE.

In addition to these East Iranian connections, some of the first Drav. words, as also many words with original retroflexes (Kuiper 1991: 17, Hoffmann 1941, 1975:16) occur first in the Middle RV book 8, more specifically in its Kāṇva section, RV 8.1-48, 8.49-59, 60-66. The name of this clan of poets, which means 'sorcerer', has been connected by Kuiper (1991: 80) with *Pra-skaṇva*, with the common Para-Munda prefix *pra-* *[*pər-*] (cf. above). This *may* indicate that Para-Munda influence extended to South-Eastern Iran, to Baluchistan, and to Makran with its many Indus settlements. Book 8 would then represent an amalgam of Dravidian and Para-Munda influences (including some pre-Iranian).

It is probably is from this Southern basis that speakers of Dravidian suddenly appear in the mid-level RV, with names such as *Kavaṣa* 'straddle legged' (*K. Ailūṣa* RV), an important priest on the side of the Pūru, Bhalānas and the other tribes opposing of the Bharata; cf. *Śailūṣa* "dancer, singer". Kuiper 1991: 24 explains that initial *c-* is often dropped in South(!) Dravidian (further: *Śirimbīṭha* : *Irimbiṭhi, śiriṇā* 'hiding place, night?' : *irīṇa* 'salt pan, hiding place (for gambling; Witzel 1999). Ailūṣa's great-grandson Tura Kāvaṣeya is an important priest of the Kuru realm who developed the 'classical' Agnicayana ritual (Th. Proferes 1999). This case shows the inclusion of a Dravidian into the fold, and underlines the important role a new 'convert' to Ārya religion could play in its very development, all of which is indicative of the high degree of language acquisition and cultural amalgamation during the middle and late Ṛgveda period.

Early Dravidian influence can indeed be traced back to the areas from the Iranian plateau (Arachosia?) to Sindh, where Drav. place names are assumed to appear first (L.V. Ramaswamy Iyer 1929-30). These names (showing MIA development $p > v$) extend from Sindh via Gujarat and Maharastra to the South: Sindhi *-vali*, Gujarati *-wārī/warī* (Sankalia 1949), Mar. *-oli*, all from a Drav. word for 'village' (Tam. *paḷḷi* 'hamlet', Kan. *paḷḷi, haḷḷi*, Tel. *palli* 'village', Kur. *pallī*, Parpola 1984, 1994: 170 sqq., 1997; Southworth 1995: 271).

A similar view has been proposed, on the basis of linguistic and archaeological observations, by Zvelebil (1972, 1990: 48, 123), Southworth and McAlpin,[19] and Fairservis (1992: 17, 21). Fairservis (1992, 1997) and Zvelebil think of an immigration by Drav. speaking tribes at c. 4000/3500 BCE, from the mountainous lands of East Iran into the Indus valley. Both underline data that characterize the Dravida as originally pastoral hill tribes. We may reckon with them in Baluchistan and only later on, after a period of acculturation with the Indus people, with Drav. farmers (Southworth 1979, 1990, 1995) who practiced intensive rice (Kenoyer 1998: 178, Jarrige 1985) and millet cultivation in Sindh.

Several passages in the 'Baluchi' book 8 of the RV indeed seem to point to the earliest references of rice in S. Asian texts. It is precisely here, at Pirak, that the newly introduced rice was first found towards the end of the Indus civilization (Kenoyer 1998: 178). It is referred to, in all likelihood, as *odana* 'rice gruel' (from Munda *oḍi(kā)* 'wild rice' in Santali *hoṛo, huṛu* 'rice plant', Berger 1963: 420, Kuiper 1950: 179) in the myth of the boar Emuṣa, whose suffix *-uṣa* indicates Para-Munda origin and which was "apparently not domesticated, not used in Indus economy" (Kenoyer 1998: 165) but is ancient among the Munda Nicobars with their boar cult. Emuṣa occurs at RV 8.69.14, 8.77.6-11, 8.77.10, 8.96.2, 1.61.7, *vy-odana* 8.63.9 (Kuiper 1991: 16 sqq., cf. other words with a non-IA prefix *vi-*, below). The myth, which clearly smacks of 'foreign' origin, is a mixture of an IA, Austro-Asiatic and possibly Drav. elements; it imitates the well known Ṛgvedic Vala myth, and its hero is a divine bow shooter (probably seen on an Indus copper plate, only at Mohenjo Daro, in Sindh, see Parpola 1997: 39; cf. also Avesta, Yt. 8.6,37 *ərəxša*, Kṛśānu RV 4.27.3, Rudra, and Murukan in S. India). Together with the introduction of rice, its *charter myth* (Malinowski) may have been taken over as well.

This kind of scanty evidence suggests that we may have to reckon, in the southern Indus area, with a similar, maybe even more complicated combination of factors than in the Panjab: with the Para-Munda Indus language, with some Munda influences, with immigration from E. Iran (e.g., Vasiṣṭha (RV 7), and of (pre-)Iranian and some Indo-Aryan tribes into Baluchistan and the neighboring Kachi plain (e.g. at Pirak, 1700 BCE), as well as with Dravidian immigration.

It is indeed possible that the Dravida constituted a first wave of central Asian tribes that came to Iran before the IA, just as the (only very marginally IA) Kassites entered Mesopotamia before the Mitanni Indo-Aryans. The Dravida knew the horse already in Central Asia, as indicated its non-Indo-Aryan word in Brahui *(h)ullī*, O.Tam. *ivuḷi* 'horse', etc., different from IIr. *aȧva*). The early introduction of horse and camel from the Iranian plateau into Sindh (Pirak and Kachi plain in western Sindh, (c. 1700 BCE, Kenoyer 1998: 178; Allchin 1995: 31) may have been due to the Dravidian speakers. But they apparently did not preserve a word for 'camel' (Tam. *oṭṭai* < IA *uṣṭra*).

---

[19] McAlpin 1981 which is based on the lexico-statistic calculations of P. Gardner, distinguishes:
Proto-Drav.: South Drav./Central Drav. - Brahui 4100-3000 BC
PDr-1 : SDr/CDr - Kurukh-Malto 2800-1900 BC
PDr-2 : SDr - CDr (Kolami, Naiki, Parji) 1500-1100 BC
PDr-3 : SDr I - SDr II (Tamil, Telugu) 1000-900 BC.

They, or a large section of the population of Sindh left this area towards the end of the Indus period (Allchin 1995, Shaffer and Lichtenstein 1999: 254), moving further east, to Gujarat, where we find a late, local phase of the Indus civilization (Rangpur phase IIb, IIc, see Allchin 1995: 32 sqq., Kenoyer 1998: 173 sqq.), and, again, Drav. place names.

There are a number of facts which point to a dialect difference between the northern and southern Indus language. Wheat, the staple of the Indus civilization, is called in Dravidian by an adaptation of a local word: *gō-di* 'low red plant' (Southworth 1988, 1979, 1990) which is quite different from the Panjab word *go-dum* > Vedic *godhūma* 'cow smoke' (but note Avestan *gantuma*!). Similarly, the Southern word is preserved by PDrav. *variñci* 'rice' and must have reached Iran (> M.Pers. *brinj*); this form differs from the northern one, *vrijhi* > Ved. *vrīhi,* which moved, via the northwestern Khaiber Pass, into Pashto *wrizē,* etc.

As seen in Avestan *gantuma* : Ved. *godhūma*, in some Vedic loan words of the northwest and the Panjab, original *-an-* is opposed to southern *-o-*. The same relationship is also found in northwestern *ś* : subcontinental *k,* northwestern *-ñ- :* subcontinental *zero* in the word for 'rice'. We can discern a clear difference between the Panjab (-->Vedic) and Sindh/Gujarat (--> Dravidian) forms of the Indus language (details in Witzel 1999 a,b).

Dialect differences between Panjab and Sindh seem even to be indicated in the Indus inscriptions themselves. Seals and plates from Harappa (Panjab) differ in a number of items from those found at Mohenjo Daro (Sindh), for example in the probable sign for 'container, quantity' which looks like a V; this is almost only found at Harappa (B. Wells 1998). The same applies to some 'suffixes' in the inscriptions (Wells, by letter 1999).

It can be concluded that the southern, 'Meluḥḥan' variety of the Indus language (after the name the Sumerians gave to the country, Meluḥḥa), the 'original' language of Sindh, differed slightly from that of the Panjab, Para-Munda Harappan. Its traces might still be found in modern Sindhi -- a type of research that has not been undertaken. There is no etymological dictionary of Sindhi.

There are some indications that this Meluhhan language was prevalent in Sindh before the immigration of the Dravida. The trade of the Indus civilization with Sumeria and later Mesopotamia has left us a number of words that are *not* Dravidian. Meluhhan was also sufficiently different from Elamite or Sumerian to require a 'translator from Meluḥḥa' (Possehl 1996a: no. 2), whose name is *Šu-ilišu* (Parpola 1994: 132). In fact, "the language of Marhaši [Bampur area, just west of Iranian Baluchistan] is different from that of the Simaškians [Tepe Yahya in southern Central Iran], and only very partially Elamite-related." (Vallat 1985: 52). Apparently, there was a language boundary, somewhere to the west of the present Iran-Pakistan border. Possehl identifies the area of Meluḥḥa (1996a, 1997) as having a center in the hills and mountains of Baluchistan, closer to the population center of the early Indus civilization. This allows for a hypothetical identification of the Marhaši language with that of Meluḥḥa, and makes a thorough investigation of the data of RV 8 even more important.

In Mesopotamia, there are men with *Meluḥḥa* as a personal name, thus apparently, 'the Meluḥḥan'; several persons, among them *Urkal* and *Ur-ᵈlama,* are called 'the son of *Meluḥḥa'.* There also is a 'village of Meluḥḥa', from where a person called *Nin-ana* comes.

The products of Meluḫḫa include *giš-ab-ba-me-lu-ḫḫa* (*abba* wood, a thorn tree), *mêsu* wood ('of the plains'), *magilum* boats of Meluḫḫan style (Possehl 1996a). In total, there are some 40 "Indian" words transmitted to ancient Mesopotamia, some of which may have been coined by Dilmun (Bahrain) traders. They include: Sindh wood *sinda* (*si-in-da-a, si-in-du*), date palm, the 'red dog of Meluḫḫa', *zaza* cattle (zebu?), elephants, etc. Coming from Dilmun (Bahrain), we may add the Meluhhan(?) trees *giš-ḫa-lu-ub* or *ḫaluppu* wood, *giš-mes-makan* or *mêsu* wood of Magan, and the *gišgišimmar* wood (cf. *śimmal in Ved. *śimbala, śalmali* 'Salmalia malabarica').

A slightly later(?) loan-word relationship is seen in Sumer. *ili* 'sesame', Akkad. *ellu/ūlu* 'sesame oil', a corresponding word is found not in North or Central Drav. but only in South Drav. with *eḷ, eḷḷu* 'Sesamum indicum' (cf. Blažek and Boisson 1992 on Sumer. loans in Dravidian); the word must be compared, however, also with Ved. *tila* and *jar-tila* 'sesame' which shows the typical Para-Munda prefix *Cər-* (cf. Kuiper 1955: 157). The ultimate source, ***(t)il*, however, is unclear (cf. above).

The word *meluḫḫa* is of special interest. It occurs as a verb in a different form (*mlecha-ti*) in Vedic only in ŚB 3.2.1, an eastern text of N. Bihar where it indicates 'to speak in barbarian fashion'. But it has a form closer to *Meluḫḫa* in Middle Indian (MIA): Pali, the church language of S. Buddhism which originated as a *western* N. Indian dialect (roughly, between Mathura, Gujarat and the Vindhya) has *milakkha, milakkhu*. Other forms, closer to ŚB *mleccha* are found in mod. Sindhi, Panjabi, Kashmiri, W. Pahari. It seems that, just as in other cases mentioned above, the original local form **m(e)luḫ* was preserved only in the South (> Pali), while the North has *mlecch. The meaning of *Mleccha* must have evolved from 'self-designation' > 'name of foreigners', cf. those of the Franks > Arab *Farinjī* 'foreigner.' Its introduction into Vedic must have begun in Meluḫḫa, in Baluchistan-Sindh, long before surfacing in eastern North India in Middle/Late Vedic as *Mleccha* (for details, see Witzel 1999 a,b).

§ 12. An older, Central Asian level.

The Para-Munda (Harappan/Meluhhan) layer is, however, not the oldest we can discern in OIA/Vedic. There are clear indications of a still older levels of substrates that can be traced, not unexpectedly for the speakers of IA, to Central Asia. This feature is of little bearing for the Indus area, unless further study should indicate the influx of Central Asian words, together with the occasional presence of BMAC objects during the later Indus period.

There is a host of unstudied words found both in OIA and Old Iranian that do not have an IE etymology and must represent old Central Asian substrate(s), most likely that of Bactria-Margiana (BMAC culture, 2100-1900 BCE). They include plants, animals, and material culture. This substrate must be of considerable age, as it also occurs in the linguistically closely related Old Iranian (Avestan and Old Persian). That automatically means that some of it is shared by the common ancestral tongue of both languages, Indo-Iranian. Even more importantly, the loans in Old Indo-Aryan (Vedic) and in Old Iranian (or Irano-Aryan) often appear in forms which often diverge slightly from each other. This situation is typical for loan words that stem from different dialects of a source language or

from two closely related sources languages, e.g. Engl. *tea*, French *thé* :: Russian and North Indian *chai,* coming from S. and N. China respectively. It is clear that many of the following words go back to such slightly divergent Central Asian origins.

First, it has to be established why we should think of central Asian origins. For, in theory, such loans could also have originated in the Panjab and have traveled to Iran (as is the case with the words for rice, discussed above), or, conversely, from some Iranian area to the Panjab. However, a selection of the most typical loans tends to center on the Bactria-Margiana area (cf. Witzel 1995, 1997a n. 54, 1999a,b), perhaps with the inclusion of Sistan for some words. The following ones can be plotted both in time and place: **uštr* 'camel', **khar* 'donkey', **išt* 'brick'.

• *uštr > Ved. *uṣṭra* / Avest. *uštra* 'camel'. The middle and new Akkadian word *udru* 'Bactrian camel' is a loan from Iran, see *EWA* I 238, *KEWA* III 652, cf. Diakonoff in *JAOS* 105, 1985, 600. The domesticated camel was introduced into the BMAC area from Central Asia only in the late 3rd mill. BCE.

• *khar* > Ved. *khara* / Avest. *xara* 'donkey', cf. also Toch. B *ker-ca-po* < **karca-bha*?, with the common Indian animal suffix -*bha* (as seen in *garda-bha* 'donkey', *rāsa-bha* 'donkey', *śara-bha* 'a horned animal', *ṛṣa-bha* 'bull'). The word may be a late 3rd mill. Near Eastern loan that came to the BMAC area together with the domesticated donkey, cf. Akkadian (Mari) *ḥārum, ajarum* 'male donkey', *EWA* I 447. The apparent overlap with Dravidian (denied by *EWA* 473) in Drav. **garda-* > Tamil *kaḻutai* 'donkey' is of importance as it may represent one of the few possible links of a Central Asian substrate that is found both in Dravidian and Vedic.

• **išt* > Ved. *iṣṭi, iṣṭikā* / Avest. *ištiia* 'brick', *zəmō-ištuua* 'clay brick'; OPers. *išti*, MP., NPers. *xišt*; cf. also Toch. *iścem* 'clay'? The regions north of the BMAC do not have any bricks in the period before c. 2000 BCE (except, perhaps for the Caucasus area and at Sintashta, east of the Urals). Again, the item and the word would have been encountered by tribes speaking Indo-Iranian or the already separated Old Indo-Aryan and Old Iranian dialects as soon as they expanded into the one area where these three loans are possible around c. 2000 BCE (or later): the BMAC area.

These so far neglected words in Indo-Iranian without Indo-European etymologies (see now Witzel 1999a,b, Lubotsky, forthc.) include, notably, the local words for the newly introduced domesticated camel and donkey, and for wheat, Ved. *godhūma*, Avest. *gantuma* < **gant-um* < *gant + um* (cf. Berger's reconstruction, Burushaski ***ɣund-um*). They are joined by a number of words from agriculture and settled life: **parš* 'sheaf', **bīj* 'seed, semen', **ya(u)v(ī)yā* 'stream, irrigation channel', **sthūnā* 'pillar', **bhiš, bhiš-aj'* 'to heal, healer', **vīnā* 'lute', **ling* 'mark', **kapaut* 'blue', **kadru* 'brown'. Of interest are also the words that have a wider range of spread in central Asia and beyond (for details see Witzel 1999a,b): ***kana/k'ana* 'hemp', **bhang* 'hemp', ***sinšap* 'mustard', **kaåyap* 'tortoise', ***pard/pandh* 'spotted animal, panther' (often found in early and also in later C. Asian art, see Dani 1992), ***kart-ka* 'rhinoceros'. The last few words which indicate animals that need a wetter regions than usually found in Central Asia may be connected with the area of the (then well watered) Sistan. (Of still older, E. European or N. Central Asian origin may be: ***medh/melit* 'sweet, honey', ***sengha/singha* 'lion').

Further, there are a number of old, little studied local pre-Iranian names. They include, among others, Avest. *Xnəṇta* < **khnanta* = Avest. *vəhrkāna,* Gorgān < **khranta*); *Suɣδa* 'Sogdia', and in O.Pers., the Median *Kampanda, Sikaya(h)u-vati*; the Arachosian *Aršādā, Gandutava*; the Southeast Iranian name Cambyses (*Kamb(a)ū/ujiya*) ~ Ved. *Kamboja, Karmāna* and *Maka* (modern Makran), *Maciya* 'a person from Maka' (details in Witzel 1999a,b).

Such data could be amplified by quite a number of verbal roots, (some of which, however, may be IE and have been preserved only in IIr.): *kan 'to find pleasure, please', *krap 'to whimper, plead', *kram 'to stride', *kroå 'to shout', etc. or the socially important *bhiš 'to heal'. Among the Iranian words without IE etymology, there is *kaiš 'to apportion magically, to teach', or *kaufa 'mountain, hump (of a camel)'.

The easily expandable list of such items allows to posit a so far unrecognized, strong substratum in Indo-Iranian that does not go back to Sumerian, Akkadian, Hurrite or Elamite but is to be located somewhere in Iran and Central Asia, most likely in the BMAC area, the immediate contact zone for steppe populations. Apparently, Elamian was spoken up to Simaški (Kerman/Bandar Abbas area), while Aratta (Sistan) and Marhaši (W. Baluchistan, Bampur region) had other language(s) (Vallat 1980). The loan word links between Sumerian and Drav. need further investigation (see above, Blažek and Boisson 1992). For the BMAC area, we have no sources at all, except for the loans quoted above (cf. however, below). All of these data need to be studied in greater detail, especially the reconstruction of the 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).

A. Lubotsky (forthc.), however, has now taken the next step, and has checked *all* words found in Mayrhofer's *EWA* and in Iranian to get a fuller picture of the Central Asian and Afghani substrate. However, he still misses the typical substrate words found only in Iranian that have no IIr etymology, such as *kaufa* 'hump, mountain'. On the other hand, he includes even the evidence from late Vedic texts which increases the risk of misidentifying local Indian words with loans *into* Iranian and vice versa (*vīṇā, piṇḍa* may be such cases, see *EWA*). However, in addition to the brief, more culturally oriented list given above, and to some of the irregular phonetic correspondences listed above (Witzel 1999, 1999a, b), he has added a number of items of formal analysis.

First, there are a number of words with unusual IIr suffixes, including *-ka, -pa, -sa* which are directly attached to the root: *atka 'cloak', *stuka 'tuft', *wr̥tka 'kidney', *j'aj'ha/uka 'hedgehog'; *pāpa 'evil', *stū/upa 'tuft', *š(w)aipa 'tail', *kac'yapa 'tortoise', *uc'ig 'sacrificing priest', *jharm(i)ya 'firm structure, house', etc. Unfortunately none of them, perhaps excluding the last one, is culturally significant enough in order to point to a particular region.

Second, there is the unusual structure of trisyllabic words with long middle syllable (type CəCəCə). They include *yavīyā 'irrigation canal', *kapāra 'vessel, dish', *mayūkha 'wooden peg', *piyūša 'biestings', *warāj'ha 'wild boar', which seem to point in the direction of an agricultural society such as that of the BMAC (as pointed out in Witzel 1999, 1999a,b).

However, it remains doubtful whether we are allowed to take this lengthening that seriously, as there are quite a number of cases were long and short penultimate syllable seem to vacillate, that is in the suffixes -aśa / -āśa; -iśa / -īśa, -aṣa / -āṣa, -īṣa / -iṣa, -uṣa / -ūṣa, -āsa, such as seen in the Ṛṣi *Kavaṣa,* the Aśvin protegé *Jāhuṣa,* the demon *Ilībiśa,* the mythical boar *Emuṣa,* the occupational designations *kīnāśa/ kīnāra,* the river designations *Kuliśī,* and *tūrṇāśa.* (A fuller list of such words, found in RV 4-6, would include *kalaśa, Turva(śa); Kavaṣa; palāśa, kīnāśa / kīnāra; tūrṇāśa; cāṣa?, jalāṣa; kilāsa, kīkāsā; paḍbīśa; Ilībiśa, kuliśa, Kuliśī; ambarīṣa; kilbiṣa; ṛbīsa; Emuṣa, Jāhuṣa; aṅgūṣa, pīyūṣa, vinaṃgr̥sa.*) Based on such

suffixes, it can also be asked whether the Harappan language distinguished between *ṣ , ś* and *s* (cf. Witzel 1999: 355 n. 65).

However, it is from the set of words with long middle vowels that Lubotsky draws some important conclusions. Since they occur both in the Old Iranian and Old Indo-Aryan (Vedic) vocabulary but are also typical for the Ṛgvedic loans that are not found in Old Iranian (such as *āṅgūṣa* RV 'hymn'), he assumes that these words could have only been taken over *inside* the subcontinent and that, consequently, the language north of the Hindukush and in the Panjab was the *same* as that in the BMAC.

This assumption does not take into account that such words with the suffix *-ə́ša, -ə́c'a, -ə́sa* also occur with short medial vowel, i.e. as the typical suffixes *-əša, əc'a, -əsa* of Para-Munda just mentioned above; this type can, in part, even stem from Central Asia, for example **jajaka/jaju-ka; *kac'ya-pa, *sika-tā/c'ika-tā,* cf. also: **aru-na, *aru-ša, *pra-salavi/fra-haravam.* Further, many of the Ṛgvedic (and later) words of the type *CəCəCə* may already have been acquired in Central Asia but were either not taken over into Iranian or have been lost there. (It must be noted that the Avestan corpus is very small compared to the Vedic one).

Actual proof for a *CəCəCə* type Panjab language could only come from some words typical for that climate, flora, fauna or culture. Indeed, there are at least two such words which may be indicative, *mayūra* 'peacock' and *śārdūla* 'tiger' (both of which, incidentally, do not occur in oldest Vedic, but only in post-RV texts). Nevertheless, the peacock is met with also further west, as an imported luxury animal. It is found already in Old Elamian art, subsequently in the Mitanni realm, and in a legend about the import of peacocks to Babylon, in a Pali text, the Baveru Jātaka. If it was originally borrowed from Munda **mara'k* (Witzel 1999, 1999a,b) this points to a Panjab form (*mayūra*) that was closer to the Central Asian/Para-Munda one (*CəCəCə*) than to the Sindh form that was retained in PDrav. as **mayil,* (or *maññai DEDR* 4642). The tiger, along with the rhinoceros, may have occurred in the swamps of Sistan. (This needs further confirmation). Its absence in the RV, in contrast to its prominence in the Indus civilization, may be due to the fact that it not belong to the traditional imagery of the IA who preferred the lion. In view of these incertainties, it remains therefore at least doubtful whether the languages of the BMAC and of Harappa were related at all.

In general, the observation of the *earliest* attestation of a particular 'foreign' word in the RV is the *only* secure basis for analysis: all words that are attested later may have emerged, at various points in time, from the popular level of speech into the sacred speech of the (post-)Ṛgvedic Mantras and into the High Vedic recorded in our prose texts, that is the educated speech of the class of the Brahmins (cf. Kuiper 1991, Witzel 1989). It therefore remains unclear whether they stem from Central Asia, from the Panjab, or from further east, -- wherever the texts that mention them first are were composed. In this *particular* context, it does not even matter whether such words are attested both in Iranian and Indo-Aryan. We have no way to find out when and where they were taken over into, e.g. early Vedic and or early Avestan.

Further research is needed, preferably by a study of Lubotsky's words as they found in the various strata of Vedic and Iranian texts. Just as in the discussion of Para-Munda and PDrav. words, carried out above, this would allow to determine what is securely reconstructable for Indo-Iranian, or for the various historical levels of *early* Old Iranian

and for *early* Old Indo-Aryan. Only then, we may determine which words may subsequently have emerged from the lower strata of society (or may even have been taken over from the local substrate languages later on).

Lubotsky's data and his discussion, however, support my analysis of Central Asian loans made earlier (Witzel 1999, 1999a,b). He, too, assumes a Central Asian substrate, though he does not localize it beyond pointing to the towns of the BMAC and assuming that the Panjab substrate had many of the same traits. This is echoed by G. van Driem (*MT*, Oct. 1999): on the basis of Lubotsky's still unpublished work, he thinks that these loan words belong to a substratum north *and* south of the Hindukush, but are restricted to Afghanistan. The contention that the geographical horizon of the oldest books of the RV was confined to Afghanistan and excluded the Panjab has already been contradicted: the Panjab rivers (and only twice the Sarayu, once the Kubhā) are prominent in these books, and include even the Gaṅgā (see above).

While there is no proof that (almost all of) these words should be derived from S. Afghanistan, the word *anc'u* 'Soma plant' (probably Ephedra), which is used to prepare the sacred drink of the Iranian and Vedic peoples, may point to the high mountains of Central Asia (Tian Shan) and Afghanistan. It is on the high mountains that the best Soma grows, both according to the Avesta and the Ṛgveda, and that is exactly where the more potent variety of Ephedra is found. There are, indeed, some indications of non-IIr speakers in the high mountains of Afghanistan (cf. Witzel, forthc., on *Airiianəm Vaējah*, and note that one buys the high mountain Soma plants from aboriginals who are then beaten up); such 'foreigners' indeed still survive in the Pamirs (with Burushaski).

One may add Lubotsky's words stemming from the religious sphere, *c'arwa* 'name of Rudra', *indra*, *g(h)andharw/b(h)a* 'a demi-god or demon' (all these may however be late Avestan loans in Vīdēvdād, from OIA, in the late 1st. mill. BCE), and *atharwan* 'priest', *ṛši* 'seer', *uc'ig* 'sacrificing priest', *yātu* 'black magic'. He draws attention to the common 'suffix' *-arwa-* and thinks that this group of words belongs to a *still older* layer of loans in IIr. At any rate, the cluster of words related to religion indicates that the Indo-Iranians were strongly influenced by a Central Asian population with its own peculiar religion which included a cult of the sacred drink *anc'u*, the Soma (Parpola 1995). This must have superseded the older IE sacred drink, *madhu* 'mede', a word that is still retained in the RV in connection with Soma.

A further indication, perhaps, of the trail of those who spoke Indo-Aryan, from Central Asia to the Panjab is seen in an early Nuristani or (or Kafiri) loan in Vedic, Nur. *kat's'a* > Ved. *kāca* (KS+) 'shining piece of jewelry' (K. Hoffmann 1976, *EWA* I 335. However, this may also be a post-Ṛgvedic loan from this mountain language, the archaic third branch of the Indo-Iranians (Morgenstierne 1973) that has survived in the mountains of East Afghanistan and in neighboring Chitral (N.W. Pakistan).

§ 13. The oldest loan words of the RV revisited

After this excursion into Central Asia, Greater Iran, and the borderlands of the Indus, it is useful to revisit the oldest books of the RV, RV 4-6: we obtain the same results

that we have already encountered above in the lowest layers of Ṛgvedic substrates, i.e. Para-Munda and Central Asian.

First of all, it must be mentioned again that even these hymns bear all the characteristics of a text of the Greater Panjab. G. van Driem and A. Parpola (1999) believe that these oldest hymns were still composed in Afghanistan and this is why they would not contain Drav. words, as outlined by Witzel (1999a,b). This is, however, not the case as these books contain references to the major rivers of the Panjab, even the Ganges (see above).

In general, the books of RV level I (RV 4-6) are thoroughly South Asian and have references to local climate, trees and animals. We therefore have to take them seriously, at their word, and cannot claim that they belong just to Afghanistan. Books 4-6 have 219 hymns, a little more than 1/5 of the RV with its 1028 hymns. Books 4-6 contain about 100 'foreign words', c. 2.19 per hymn, which is only slightly less than the ratio for the whole text, 2.68 per hymn. And the languages represented are the same that are encountered in the other books.

A list of the 'foreign words' of books 4-6 follows; it is extracted from Kuiper's general list (1991), but excludes the words found in appendix hymns (Oldenberg 1888).

atasa, avata, aṅgūṣa, āṇi, indu, ukha-, ulkā, Urvaśī, ṛbīsa, kakāṭa, katpaya, kapanā, kabandh-in, kalaśa, kākambīra, kila, kilāsa, kilbiṣa, kīri, kīrin, Kīsta, Kutsa, Kubhā, kulāya, kenipā, kevaṭa, kośa, Kaulitara, krīḍ, khaja, khādi, khārī, khilya, garta, Gāṅgya, Cumuri, jaghana, jajjhatī, jaṭhara, jabāru, taskara, tilvila, Tugra, Tṛkṣi, Namī, Nahuṣa, niṇik, niṣka, nīla, Paṇi, Piṭhinas, pippala, Pipru, pīḍ, Puraya, Puru-kutsa, puṣkara, phaṇ, phaliga, phalgva, baṭ, badhira, bali, bīja, Bṛbu, bhan, mukha, menā, Yadu, Rākā, rāṇḍya, reṇu, vaṃsaga, vaṇij, Vara-śikha, vāṇa, Vipāś, Vetasu, śakuna, Śambara, śarvarī, śāṇḍa, śoṇa, Saramā, Sarayu, Sāpya, sīta, Sṛñjaya.

This constitutes 88 'foreign' words, plus some 17 words in late hymns of RV 4-6, a total of 105 words. From a purely formal point of view we have :

• 2 words with non-IA *-īs-, -ṛs-;*
• 2 words with 'foreign' *-ṇḍ-* (Hoffmann 1941);
• 15 words with other non-IE retroflexes (*ṭ, ṭh, ḍ?, ṇ*);
• 10 words with Lubotsky's non-IE *kh, ph, th;*
• 3 words with Lubotsky's non-IA *-ka, -sa* suffixes;
• 2 words with non-Vedic, "language X" geminates: *pippala;* (*jajjhatī* probably is from IA *has* 'to laugh' (*EWA*); cf. also *katpaya < kappaya*, if with geminate substitution, see Kuiper 1991). These 'foreign' words can be classified as follows.

* **Para-munda is represented by words with typical prefixes:**
• *śa-kuna, śa-kuni*, from PMunda **śa-konti'd* , Munda *kon-ti'd* 'bird' (cf. Epic *Śa-kuntalā, Ved. Kunti,* for details see above);
• *Puru-kutsa*, to be connected with *Kutsa* (prefix *pər-*, Kuiper 1991), like *Pra-maga-n-da* with *Ma-gadha, Ga-n-dhā-ri*, and via popular etymology, with *Gaṅgā;*
• *Śam-bara*, cf. above on *śam-bala, śar-varī, api-śar-vara, śa-bala* with simple prefix, note also: *Śa-bara* 'a tribe', *śa-vala* 'spotted', *śam-bara* 'a demon, a (spotted) deer', *kar-bara* 'spotted' (CDIAL 12296-7, 12314-5, 2882);
• *Vi-pāś, < *vi-pāž/*vi-bāl,* see above on *Vibālī, Vaiśambhalyā < *vi-šambaž, *vi-šambāl* (cf. *Viśpalā?*). Both words, **vi-pāž, *vi-šam-baž,* probably are formed with the prefix *śam/k'am-* as seen in *Śam-bara, Kam-boja,* from **(vi)-šam-bāž* (cf. *Vi-darbha, Vi-bhindu,* cf. popular etymology *vi-śambala* 'having widespread blankets')

Other likely candidates are: *ka-kāṭa, ka-panā, ka-bandh-in, ka-laśa?, kā-kambīra; ki-lāsa, kil-biṣa; kenipā, kevaṭa* (cf. *avata*), *ja-ghana* (if not from Ved. *han*), *ja-ṭhara, ja-bāru, tas-kara, śar-varī, sṛñ-jaya.*
Finally, the compound word *til-vila* is clearly connected with *tila, jar-tila,* perhaps the clearest cases of a word a the typical (Para-) Munda prefix (see above).

* The Central Asian substrate -- if we provisionally follow Lubotsky's specifications -- is not larger than in the other books; it includes only:
- a few words with long middle syllable: *aṅgūṣa, ṛbīsa, kilāsa* (but see above about this claim);
- a few with non-IA word formation: *niṣ-ka* (none in *-pa*!), *ata-sa?, Kut-sa?*;
- words with non-IA sounds: *kh, ph, th*: *ukha-, khaja, khādi, khāri, khilya, mukha, Vara-śikha; phaliga, phalgva, Piṭhinas?*;

* Dravidian is restricted to less than a handful of words, discussed above (*kulāy-in* 'nest-like', *ukha* 'pan, hip'); however, *āṇi* 'axle pin' and *bala* 'strength, force' are unlikely as being of Drav. origin.

In sum, the oldest books of the RV reflect a strong local Para-Munda substrate, with some Central Asian influence but just a few traces of Dravidian and "Language X".

§ 14. The 300 Ṛgvedic loan words as Rosetta Stone

The conclusion to be drawn from the data presented above is that a large amount of the c. 380 loans in the RV (excluding, of course the older, Central Asian loans), stem from the original Panjab language(s) of the pre-IA population. In other words, they represent the Harappan population(s), and therefore can serve as the Rosetta stone for the Indus script, -- even if most of them deal with agriculture, village life, music, popular customs and some religion (Kuiper 1955, 1991) and they contain very little, if anything at all of city life.
However, as has become clear now the first recorded Indus signs, that is pre-firing and post-firing symbols and signs written on pottery, date from the pre-city levels of 3500/3300 BCE (Kenoyer and Meadow at http://www.harappa.com/indus2/index.html; R. Meadow at http://sarasvati.simplenet.com/meadow/meadow2.htm, cf. Newsweek.com May 6, 1999). We may therefore assume that they represent real 'Harappan' words, starting in the later 4th mill. BCE. While we cannot be sure whether the signs have been interpreted and read in the same fashion up to the end of the Indus period (or even the same language), we can begin right there, at c. 1900 BCE, and compare the late Indus documents with the next available evidence, the loan words in the oldest layers of the RV, at c. 1500 BCE.
It is, therefore, the 300-odd 'Panjab' words and the few we can gather from Sindh that should be employed in any attempt to decipher the Indus script (Fairservis 1992: 14, Parpola 1994: 137 sqq., Possehl 1996b).[20] The decision, however, which language to try out in a possible decipherment is just one step in many, and should be a secondary one,   at this stage of our knowledge. For, even now, a proper and complete sign list of all its characters

---

[20] Some initial observations point in this direction, for example the role of the prefixes indicating personal names and tribal names. It is too early, however, to announce details.

(without premature conflation of similar characters) is not yet available (see Wells 1998, 1999), -- not to speak of the even more indispensable catalogues of all possible sign combinations. A lot of preparatory work has to be done still before we can again try to actually decipher the inscriptions.

One may then try to argue that, after all, the language of the Indus documents was Dravidian, if it was *the* dominant, or one of the dominant languages of Sindh at the time. (However, the language of the Indus documents seems to be the same as that in the Panjab, even if there are typical dialect deviations, apparently reflected also in some special signs used only in Harappa and others only in Mohenjo Daro, see Wells 1998). Dravidian might have been the traders' language of the Indus area. However, as has been pointed out above, if this had been the case, one would expect some remnants of Dravidian in the *earliest* level of the - already Panjab centered - RV. At least some words such as 'village chief' or 'trade, trader' should be of Dravidian origin (see above). This is *not* the case[21] in the earliest RV (books 4-6). In addition, such a scheme would involve an elaborate -- and very unlikely -- scenario of dominance and subsequent replacement of languages: the *complete* disappearance of Dravidian influence in the Panjab between c. 1900/1700 BCE and its sudden re-emergence at c. 1450/1300 BCE.

§ 15. Towards a new paradigm

From the preceding paragraphs it transpires that we should strive for a new paradigm of the early linguistic (and ethnic) situation of South Asia. Using the earliest

---

[21] One may investigate, in the oldest RV books 4-6, *vaṇij* 4.43.6 which Geldner translated by '(flying) merchant' / 'der (fliegende) Kaufmann' (??), 5.45.6, 'merchant', *vāṇa* 'trade?' in *(pra-)vāṇá* 4.24.9 'trade' (Kuiper 1955: 168; cf. also *vāṇa,vāṇī* 'voice, music,' *vāṇicī* 'similar to voice'). Mayrhofer *EWA* II 494 derives *van-ij* 'winning goods' from *van* (*vasu ... van*) 'with context free -*ṇ*-'. But does a context free -*ṇ*- really exist in the RV? (Discussion by Mayrhofer 1968, 509-517, contrast Kuiper 1991). Clearly, the meaniung of RV *vaṇij,* in their mythological contexts is not sure (personal name? e.g., with Munda suffix, **pan-i'c*?)

Kuiper 1955: 168 combines [+]*pravāṇam* with Drav. *paṇ* but concedes that these etymologies 'do not remove the difficulties'. Indeed, the phrase is open to discussion as well (*ví duhanti prá vāṇám)*. Hillebrandt 1927-29: I.226 translates *vāṇa* 'reed, Soma stem' (cf. *bāṇa* 6.75.17, Kuiper 1948: 34, but see *EWA* II 541), *paṇ* does not yet mean 'to barter'' and the *Paṇi* are demonic (see immediately). Do we have to compare Munda: Sant. *pañca* 'to lend, to borrow', *pante* 'line, row, to agree together'?

However, RV *Paṇi* (4.25.7, 4.51.3, 4.58.4 late; 5.34.7, 5.61.8 late, 6.13.3, 6.20.4, 6.31.2, 6.39.2, 6.20.4, 6.31.2, 6.39.2, 6.45.31, 6.53.3, 6.53.5, 6.53.7, 6.61.1, 6,) cannot be connected with the words for 'trade', and *paṇ* 'to barter' is found only in post-RV texts, such as KS, *prapaṇa* 'trade' AV, *pratipaṇa* AV 'barter, exchange', see *EWA* II 69; cf. also *DEDR* 3884 *paṇ* 'work, service', *paṇikkan* 'carpenter'. S. Palaniappan wants to connect (by letter 1999) *DEDR* 3891 Old Tam. *paṇ* 'field' (Puṟanāṉūru 16.16), *paṇai, paṇṇai* (cf. Kan. *paṇṇeya, paṇya*), *paṇṇiyam, paṇṇaiyam* 'farming', *paṇai* 'to thrive, flourish'. He reconstructs *paṇi 'flourishing rich farmer' and compares *DEDR* 3884 Mal. *paṇiyan,* a caste of cultivators.

However, the RV *Paṇi* are (rich) 'foreigners', and the prominent Vala myth actually talks only about their cattle, not grain. They are on the boundary line between demons and are half-mythical, human enemies, and may reflect past encounters, if the connection with the *Parna (Greek form *Parnoi*) in the BMAC area is correct; it echoes the ethnic term for foreigners an enemies *Dāsa (N.Iran. *Dāha*, Greek *Daai,* Avestan *Dåŋha,* cf. *Dahāka*, etc., see Witzel forthc.). All these would be Central Asian reminiscences.

available evidence, the loan words in the Ṛgveda (and a few words transmitted in Mesopotamian documents), we can establish the following spheres of influence. (The dates given below are, as always, very approximate).

• **Before c. 1450 BCE.**
In the Greater Panjab, the prefixing Para-Mundic or Para-Austroasiatic Harappan language was spoken, along with a few hints of Masica's more eastern (Haryana/U.P.) "Language X"; the Central Asian substrate, brought into the Panjab by the oldest layer of Ṛgvedic, probably was of little consequence during the Harappan period.

• **From c. 1450-1300 BCE**
 The Greater Panjab saw the first influx of Dravidian words, most likely from Sindh; the situation in Sindh and Baluchistan is less clear: a variety of the Para-Mundic Harappan was amalgamated with Munda influences from the east, with the immigrant(?) Proto-Dravidian, and the immigrant early Old Indo-Aryan and in Baluchistan also some pre-Iranian. Amalgamation of Indo-Aryan, Para-Munda and Dravidian elements in the Greater Panjab and in Haryana.

• **After c. 1300 BCE.**
Continuing amalgamation, evidenced by increase of 'foreign' words in the late Ṛgveda. The trend continues in the post-RV texts (YV, AV Mantras and in later Vedic) with a continuing influx of the *same* types of vocabulary into the educated Vedic speech of the Brahmins. By this time, the increasing amount of textual materials allows to detail the existence of some other languages in the Greater Panjab or on its rim: Proto-Burushaski in the northwest, Tibeto-Burmese in the Himalayas and in Kosala, Dravidian in Sindh, Gujarat and Central India, and predecessors of remnants language groups, now found in isolated pockets of the subcontinent.

Because of the amalgamation of at least three groups (IA, Para-Munda, Drav.) we have to suppose a large degree of bilingualism (cf. Kuiper, A bilingual Ṛṣi, in press) and even trilingualism, and the forming of pidgins. While a Vedic pidgin was the lingua franca and used at home, proper Vedic Sanskrit was learnt 'in school' by members of the Ārya classes. Large pockets of the Para-Munda Indus language, of the newly arrived Dravidian as well as some remnants of the Gangetic Language "X" must have survived as well.

When the linguistic evidence of the immediate post-Harappan period, and by and large the ethnic picture depending on it, are viewed against the evidence found in the texts and in archaeology, a new picture of archaic South Asia emerges.

To begin with, the *details* for the import of IA language and culture still escape us, though the recent linguistic discoveries related to the BMAC and Afghani area clearly show the trail of their import. Further, their word for chariot and chariot driver are new coinage from the point of IE, but already archaic by the time IA appears in South Asia. The earliest IIr *ratha* 'chariot (with two spoked wheels)' (Gening 1977, Pigott 1992, Anthony u. Vinogradov 1995, cf. Littauer u. Crouwel 1996) have been found about 2000 BCE, near the Volga (called, in North Iran. *Rahā* > Greek *Rhā,* preserved as Avest. *Raŋhā,* Ved. *Rasā*) and east of the Urals at Sintashta. The IIr word for 'chariot', however, is old enough to have

resulted in the archaic compounds Ved. *rathe-ṣṭhā*, Avest. *raϑaē-šta-* 'chariot fighter', (cf. Old Avestan *raϑī*, RV *rathī* 'chariot driver); cf. also the many IA names in the Near East containing this word, such as Tuš-ratta.

This is something that cannot be said (yet) from the point of view of South Asian Archaeology. As is well known now, steppe influence is found in the BMAC but does not extend beyond it; instead, there is some BMAC influence that is found from Susa to Harappa. On the other hand, there is, for the period in question, a huge gap in exploration between Khorasan and Kabul for the mid- 2nd millennium BCE.

None of the archaeologically identified post-Harappan cultures so far found, from Cemetery H, Sarai Kala III, the early Gandhara and Gomal Grave Cultures, does make a good fit for the culture of the speakers of Vedic: one would have to look for remains that include horse and chariot, primitive ritual pottery (handmade) and remnants of fire and Soma rituals with wooden implements, burials both with interment and cremation, and the general kit of cattle herders moving about in transhumance fashion on heavy wagons and having some barley cultivation on the side.

Another open question is the presence of speakers of Dravidian. The linguists have not yet decided whether they in fact originally were pastoral people from the mountainous areas of Greater Iran (Zvelebil 1970, 1990: 48, 123); others (Krishnamurti, forthc. @@) would have then inside South Asia at a very early date. In deciding this question, the agricultural loan word links of Dravidian with Sumerian will be of importance (Blažek and Boisson 1992); there also are some vague similarities with Caucasian languages in some loans words such as 'horse' and 'wheat' (Witzel 1999a, b). Such factors may indeed point to a western, Iranian location of Dravidian before immigration into South Asia   (Southworth and McAlpin, 1979, however, think they participated in the Indus civilization, from which they acquired agriculture and the accompanying vocabulary). One thing is clear, though: Dravidian and Indo-Aryan (IE) words for domesticated animals are quite different from each other, for example, Drav. *DEDR* 500 Tam. *ivuḷi*, Brah. *(h)ullī*, 1711 Tam. *kutirai* etc., *DEDR* 3963 Tam. *pari* 'runner', 4780 Tam. *mā* 'animal' (horse, elephant), Tel. *māvu* 'horsecf. Nahali *māv* 'horse' , (cognates mean 'deer' etc. in other Drav. languages); they have no relation with IA *aśva* 'horse' and various words for 'runner' (*arvant, vājin*, etc.). Ved. *garda-bha EWA* I 473, Drav. *kaḻu-tai DEDR* 1364 'donkey') seem to be independent loans from a language of Greater Iran or Central Asia (Witzel 1999a, b).

Early contact between both groups on the Iranian plateau, supposed by Southworth (1979: 203, 228 f., 1990: 222-3, 1995) is to be excluded (Witzel 1999 a,b); *pace* Southworth (1979: 196f.), there also is little secure evidence for *early* loans from IA into Drav.; such words can have been taken over any time between the RV (1200 BCE) and the earliest attestation of Tamil at the begin of our era.

Against this background, it seems possible that (some) bands of speakers of early IA and Drav. entered South Asia independently of each other, perhaps even at about the same time: the IA via the Khyber and maybe some other passes in E. Afghanistan, and the Drav. via Baluchistan, both as pastoral peoples with very little agriculture; both exploited the system collapse of the Indus civilization, and occupied the newly open lands which could now profitably be used for pastoralism. In other words, the whole convoluted, and by now highly emotionalized, question of the so-called Aryan (and Dravidian) 'invasion' or

immigration/trickling in may boil down to the decision of some opportunistic transhumant tribes of Greater Iran who opted to stay on in their winter quaters in the Indus plains, instead of returning to their mountain pastures in the spring...

While this could have provided the initial trigger for the introduction of Indo-Aryan language, spiritual and material culture into Gandhara and the adjoining Panjab, it would also have set the stage for the expansion of the suddenly much more attractive pastoral economy which was quickly copied by the original Para-Munda speakers of the Panjab. In other words, Ehret's scenario of a 'billiard ball like' expansion of a newly synthesized culture and its status kit (Ehret 1988) applies. This model provides all ingredients for the subsequent spread of the Vedic language and eastwards into Haryana and beyond, and it also explains why the term *arya* seems to be so vague in the RV: some local people (with 'foreign' names) had opted to become 'cultural Aryans', others, the *dasyu*, did not.

While Indo-Aryan culture and language may have spread over the Panjab along the lines of this scenario, one also has to consider that the earliest IA of the RV period could not yet be in close contact yet with the immediate agricultural successor cultures of the Indus civilization during the period before c. 1450 BCE, as the post-Harappan cultures predominantly are found expanding further east, in E. Panjab/Haryana (Shaffer 1999). It is only the tribes of the middle RV (RV II), especially the Bharata, who actually 'settled' in Kurukṣetra and came into closer contact with the more numerous remnants or successors of the Indus civilization.

Strikingly, these people are mentioned only in general terms, as *dasyu* 'enemies' with their *pur* 'forts' (Rau 1976) -- but not as having *cities* (even of middle level size). Clearly, this is a serious disjunction between the RV/early post-RV texts and archaeological finds that has not yet been explained. However, as mentioned above, local people are made fun of as not understanding how to deal with cattle (RV 3.53.14 "what is the use of cows with the Kīkaṭa?"), obviously a jab at their non-pastoral habits and/or their lack of proper Vedic ritual. The Indo-Aryans regarded themselves, in opposition to the agricultural 'Panjab/Haryana' people, as excellent pastoralists. The exact nature of their relationship is in need of a reinvestigation that would transgress the simple Ārya-Dasyu ("us : them") paradigm followed so far.

In this context, one has, however, to exclude the possibility, favored by van Driem (1999) and apparently also by Parpola (forthc., and by letter June 1999), that the IA at first met only with *local Gandhara/Panjab* people (= Para-Munda) and that they only came into contact with *Dravidian* speakers of the successor cultures of the Harappan civilization in Panjab proper, much later on. On the contrary, Indus influence is felt in Gandhara/Kurram and even well as beyond the Hindukush, in Shortugai (summary by Shaffer 1999, 246). The influence of the Harappan language therefore should have been noticeable in the languages of these areas as well. However, early Ṛgvedic does exhibit (almost) no traces of it, even though the texts are not restricted to Afghanistan and Gandhara but clearly indicate a pattern of (transhumance) movement from (W. and E.) Afghanistan to the Ganges (see above). Somewhere along their trails, the speakers of early Ṛgvedic should have picked up a few Drav. words.

This leaves no choice but to assume that speakers of Dravidian entered into the orbit of the Ṛgvedic tribes only during the middle RV period, -- most likely from their stronghold in more southern territories, in Greater Sindh, where their presence is still revlealed by certain place names (Sindh, Gujarat and Maharastra). This scenario explains why the c. 300 foreign words of the RV (in the Panjab) with their (agricultural) vocabulary are relatively free of Drav. influence.

It seems that the speakers of Indo-Aryan entered the Panjab and acquired local words from the Northern dialect (*śaṇa, lāṅgala, vrīhi, godhūma, kaṅgu, Gandhāra*), and that the Dravidians entered Sindh at or about the same time and acquired the related words from the southern dialect (*gōnu, ñāñcil, variñci, godī, kaṅku/kampu*). Perhaps it was them who made horses statues at Pirak (1700 BCE), not the IA(?) Bhalānas. Obviously, use of horses is not linked to speakers of an IA language.

The Middle RV period, therefore, is of great interest: it witnesses the emergence of a new coalition of tribes, that of the Bharata led by king Sudās, who overcomes the older loose confederation, led by the Pūru, of the older Ṛgvedic tribes, the 'Five Peoples': Anu-Druhyu, Yadu-Turvaśa (Witzel 1995, 1997a). The Bharata and their Kuru successors, a new group of IA immigrants from the other side of the Indus (such as Vasiṣṭha RV 7, cf. JB 3.238-9 §204), apparently brought with them a slightly divergent, sometimes more Iranian looking, and more modern form of Vedic that became the standard in the post-RV texts (Witzel 1989).

In general, this is a period of great unrest and changes on all levels of society and religion (Witzel 1997a). Importantly, it is only in this period that the 'forth class', the Śūdra, first appear; they are added to but not *included into* the traditional list of three Ārya classes (Brahman poets/priests, Kṣatriya noblemen, and Viś/Vaiśya 'the people'). That not all, and perhaps not even the majority, of the local people and of the incoming Dravidians were automatically included into the Śūdra class is clear from the many names of 'kings' (chieftains), noblemen, poets and priests that are not Indo-Aryan (Balbūtha, Bṛbu, Varo Suṣāman, Kāṇva, Kavaṣa, etc., Witzel 1999.) The last mentioned is of great interest as he is the main poet/priest on the side of the Pūru, and his third generation descendant Tura Kāvaṣeya becomes one of the main protagonists and 'reformers' of the new solemn Śrauta ritual (Proferes 1999) under the Kuru.

In sum, it is during the middle RV period (c. 1450-1300 BCE) that acculturation of elements of IA, Para-Munda, Drav., took place, along with a restructuring of society into Ārya and Śūdra, centered on Kurukṣetra in the Haryana region. The late Ṛgvedic Kuru confederation, supplanting the 50-odd Ṛgvedic clans and tribes, became the cultural center, with a new society and a new elite, its 'status kit' being that of Vedic orthopraxy. This culture is characterzied by pastoralism (cattle, horse, sheep, goat), IA ritual and acculturated customs, IA religion and ritual, but also post-Indus type agriculture (barley, wheat, rice, millet) and local artisans (potters, etc.). The new culture then spread eastwards into the Gangetic plains, and ultimately to Bihar, and along the Chambal, towards the Vindhyas.

It remains to be seen what Archaeology can add to or contradict of this sketch. One could, for example, expect some well-preserved "Indo-Aryan" camp sites in the drier parts of the Ghaggar-Hakra system), such as the camps signaled by Rafique Mughal (1995), not to

speak of some undisturbed parts of the Gandhara/Kurram areas. If indeed found, their relationship to the Cemetery H sites and the post-Harappan smaller towns in Haryana (J. Shaffer 1999) should be investigated. Was there any exchange, or did the Indo-Aryan pastoralists remain aloof, somewhat in the nature of the modern Godar pastoralists in Sistan (Falk 1997), and just relied on local (Para-Munda) agriculturists?

At the present moment, we can only state that linguistic and textual studies confirm the presence of an *outside*, Indo-Aryan speaking element, whose language *and* spiritual culture has definitely been introduced, along with the horse and the spoked wheel chariot, via the BMAC area into northwestern South Asia.

However, much of present day Archaeology denies that. To put it in the words of J. Shaffer (1999: 245) "A diffusion or migration of a culturally complex "Indo-Aryan" people into South Asia is *not* described by the archaeological record." According to what has been described above, it is also not to be expected that (a group of) Indo-Aryan tribes ("a people") decided, on a nice spring day, to "invade" the Indian subcontinent (cf. Witzel 1995, 1997a, n. 54).

Instead, the *influx* of *some* speakers of early Indo-Aryan from beyond the Hindukush --whatever their enthnic and genetic composition--, and the importation of their spiritual and material culture *must be explained*. So far, clear archaeological evidence has just not been found. Yet, any archaeologist should know from experience that the unexpected occurs and that one has to look at the right place: people on the move (such as the Huns) leave few traces. Had it not been for ample classical testimony and the very recent finding of their graves in Hungary -- they would simply 'not have entered Europe', to apply J.Shaffer's *dictum*: "their diffusion or migration ... [was] *not* described by the archaeological record."[22]

Finally, much or even most of the IA cultural and spiritual data can simply not be 'seen' by Archaeology: it would look just like the remains of any other group of second millennium pastoralists -- unless we will be lucky enough to come across a complex that presents us with *exactly* those material remains described above (chariots, handmade pottery used in rituals, fire altars, Soma residue, etc.)

J. Shaffer's 'anti-linguistic' diatribes[23] (1984, 1999) therefore are misplaced as they run in the open doors of present day linguists and Vedic specialists, and aim into a void.

---

[22] Actually, the case of the White Huns in India is parallel. All we get of them is testimony in Sanskrit and Chinese texts about India, plus a few Rajput clans -- and their Brahmins! -- who are called *Hūṇ* and who do not belong to the sun or the moon lineage but have the Fire god as their clan ancestor; finally there is a *Hūṇ-deś* on the Sutlej, high up in the Himalayas, just across the Tibetan border.

[23] While one may regard his 1984 paper as a justified reaction against the dominant 'invasionist' fashion of (post-)Wheeler S. Asian archaeology, his most recent characterization of the history of linguistic research and the history of 19th and 20th c. thought (Shaffer 1999) still is one-sided and flawed. For example, the "Romantic period" was one of the early, post-Napoleonic 19th century, not one of the late 18th when comparative IE linguistics first were suggested by Lord Monboddo and W. Jones at a time when European intellectuals were strongly interested in their ancient (non-Classical, local, folk) literatures. The first 19th cent. linguists (1809 and later: Bopp, Rask) did not strive to reconstruct the "language of paradise", but were interested in the earlier stages of humankind in general; note also that IE linguistics and others were developed at the same time as Dravidian,

Somehow, he does not seem willing to recognize that we deal with two entirely different, *largely incompatible sets of data*: The South Asian discontinuity of the second millennium is not one of the local food (or pottery) producing cultures but one of language, poetry, spiritual culture, though it also includes some material culture such as the -- not yet discovered -- Vedic chariots. Vedic pottery, produced as is was by local artisans, is of little help (Rau 1983). And, as has been pointed out above, the general remains of any pastoral society are not enough either. Specific evidence is *to be sought for in places that are likely to contain it*, from the eastern hills of Afghanistan to Cholistan and even in parts of Haryana.

Instead of fighting the ghosts of the 19th and early 20th century, we should, in the 21st, look for the similarities or differences, for overlap or linkages, and for the obvious cultural changes that are visible both within the archaeological and the linguistic/textual record.

####END ####

Finno-Ugrian (Uralic), (Uralo-)Altaic, and (Hamito-)Semitic (now 'Afro-Asiatic') linguistics. Certainly, all these language families were not all regarded as 'languages of paradise.'

Fig 1. Overview of the languages of N. India during the Indus and the Vedic period

----

2600-1900 BCE	INDUS CIV. :	'Harappan' (Para-)Munda and a southern variety in Sindh (Meluhhan)

various substrate languages on the rim ("language X", Proto-Nahali etc.); Proto-Munda in
Central and E. India; Arattan and Marhašian in Afghanistan/Baluchistan,
BMAC language(s) in Central Asia

1900/1700?	immigration of speakers of (coming with a C. Asian substrate)	Indo-Aryan

1700-1450?   RV I : some hymns of books 2, 4, 5, 6:
                   Yadu-Turvaśa, Anu-Druhyu tribes

                                                                         (Para-)Munda substrate

1450-1300   RV II: books 2,4,5,6 + 3, 7, 8.1-66, 1. 51-191
                   Pūru and Bharata immigration
                   Ten King's Battle & Bharata victory
                   (immigration of Vasiṣṭha from E. Iran)

                                                                         first Drav. words

1300-1200   RV III: books 8.8.67-103, 1. 1-50, 10.1-854; 10.85-191
                   emergence of Kuru realm   more Drav., Munda words
----

Abbr.

DEDR
Drav.
EWA
Gk.
IA
IE

IIJ
KEWA
MT
NIA
PB
PS
RV

WZKS
ZDMG

ABBREVIATIONS

Note: for ready reference, the five historical levels of Vedic are indicated by numbers (1-5), followed by their geographical location, W: western North India = Panjab, Haryana, C: central North India = Uttar Pradesh, E: eastern North India = N. Bihar; S: southern N. India = between the Jamna/Ganges and the Vindhya mountains).

AA	Austro-Asiatic
AB	Aitareya Brāhmaṇa (4, W & E)
Akkad.	Akkadian
ĀpDhS	Āpastamba Dharmasūtra (5 C)
ĀpŚS	Āpastamba Śrautasūtra (5 C)
Armen.	Armenian
Austro-As.	Austro-Asiatic
AV	Atharvaveda Saṃhitā (2 C)
Avest.	Avestan
AVP	Atharvaveda Saṃhitā, Paippalāda version (2 W)
Beng.	Bengali
Brah.	Brahui
BSL	Bulletin de la société de linguistique de Paris
BŚS	Baudhāyana Śrautasūtra (4-5 C)
Bur.	Burushaski
CDIAL	Turner 1966-69
DED	Burrow, T. and M.B. Emeneau 1960
DEDR	Burrow, T. and M.B. Emeneau 1984
Drav.	Dravidian
ep.	Epic Sanskrit
EWA	Mayrhofer 1956-76
Gr.	Greek
GS	Gṛhyasūtra(s) (5)
Guj.	Gujarati
HŚS	Hiraṇyakeśi Śrautasūtra (5 C)
Hitt.	Hittite
IA	Indo-Aryan
IE	Indo-European
IIJ	Indo-Iranian Journal
IIr	Indo-Iranian
Indo-Ar.	Indo-Aryan
Iran.	Iranian
JB	Jaiminīya Brāhmaṇa (4 S)
Jpn.	Japanese
Kan.	Kannada, Canarese
Kaśm.	Kashmiri
KaṭhĀ	Kaṭha Āraṇyaka (4 W)
KauśS.	Kauśika Sūtra (5 C)
KB	Kauṣītaki Brāhmaṇa (4 C)

KEWA   Mayrhofer 1986-96
Khar. Kharia
Khot. Khotanese Saka
KS   Kaṭha Saṃhitā
KŚS   Kātyāyana Śrautasūtra (5 E)
Kur.   Kurukh
LŚS   Lāṭyāyana Śrautasūtra (5 S)
Lit.   Lithuanian
Mal.   Malayalam
Mar.   Marathi
Mbh. Mahābhārata
MIA   Middle Indo-Aryan
MP.   Middle Persian
MS   Maitrāyaṇi Saṃhitā (2-3 W)
MT   Mother Tongue
Mund.   Mundari
Nep.   Nepali
New.   Newari
NP.   New Persian
NIA   New Indo-Aryan
Nir.   Nirukta (5)
Nur.   Nuristani (Kafiri)
OP   Old Persian
Osset.   Ossetic
Panj.   Panjabi
Pkt.   Prakrit
PS   Paippalāda Saṃhitā (2 W)
PSK   Paippalāda Saṃhitā, Kashmir MS.
RV   Ṛgveda Saṃhitā (1, Greater Panjab)
RVKh   Ṛgveda Khila (2 W)
ṢaḍvB   Ṣaḍviṃśa Brāhmaṇa (4 W)
Saṃh.   Saṃhitā(s)
Sant.   Santali
ŚĀ   Śāṅkhāyana Āraṇyaka (4 C)
ṢB   Ṣaḍviṃśa Brāhmaṇa
ŚB   Śatapatha Brāhmaṇa (4 E)
ŚBK   Śatapatha Brāhmaṇa, Kāṇva recension (4 C)
ŚS   Śrautasūtra (5)
Skt.   Sanskrit
Sum(er).   Sumerian
Sū.   Sūtra(s) (5)
Suśr.   Suśruta
SV   Sāmaveda Saṃhitā (2 W)
Suśr.   Suśruta
StII   Studien zur Indologie und Iranistik
TĀ   Taittirīya Āraṇyaka (4 C)

Tam.	Tamil
Tel.	Telugu
TB	Taittirīya Brāhmaṇa (4C)
Tib.	Tibetan
Tib.-Burm.	Tibeto-Burmese
Toch.	Tocharian
TS	Taittirīya Saṃhitā (2 C)
Up.	Upaniṣad(s) (4)
V.	Vīdēvdād
VādhB	Vādhūla Brāhmaṇa (Anvākhyāna) (4 C)
Ved.	Vedic
Ved. Index	Macdonell - Keith 1912
VS	Vājasaneyi Saṃhitā (2 E)
YV	Yajurveda (-Saṃhitā) (2)
ZDMG	Zeitschrift der deutschen morgenländischen Gesellschaft

=======

Allchin, F. R. *The Archaeology of Early Historic South Asia. The Emergence of Cities and States.* With Contributions from George Erdosy, R. A. E. Coningham, D. K. Chakrabarti and Bridget Allchin. Cambridge: Cambridge University Press 1995

Anthony, D. and Vinogradov, Nikolai B. The birth of the chariot. *Archaeology* 48, 1995, 36-41

Bedigian, D. and J. H. Harlan. Evidence for the cultivation of sesame in the ancient world. *Economic Botany* 1985

Beekes, R. S. P. *Comparative Indo-European linguistics : an introduction.* Amsterdam / Philadelphia: J. Benjamins   1995.

Benedict, P. K. *Sino-Tibetan. A Conspectus*, Cambridge 1972

---, *Japanese/Austro-Thai.* Ann Arbor: Karoma 1990.

Berger, H. Review of: K. H. Pinnow, Versuch einer historischen Lautlehre der Kharia-   Sprache. (Wiesbaden   1959), *ZDMG* 112, 1963, 416-421

---, Die Burushaski-Sprache. Wiesbaden: Harrassowitz 1999

Bhattacharya, S. Field notes on Nahāli, *Indian Linguistics* 17, 1957, 245-258

Blažek, V. and C. Boisson, The Diffusion of Agricultural Terms from Mesopotamia. *Archív Orientalní* 60, 1992,   16-37

 Burrow, Th. Some Dravidian words in Sanskrit. *Transactions of the Philological Society*, 1945,   79-120

---, Loanwords in Sanskrit. *Transactions of the Philological Society*, 1946, 1-30

---, Dravidian Studies VII: Further Dravidian Words in Sanskrit. *Bulletin of the School of Oriental and African   Studies* 12, 1947-48, 365-396

---, *The Sanskrit language.* London: Faber and Faber 1955

---, Sanskrit and the pre-Aryan Tribes and Languages, *Bulletin of the Ramakrishna Mission Institute of Culture*,   Febr. 1958, Transact. 19

---, Sanskrit words having dental -*s*- after *i, u,* and *r.* In: A. M. Davies and W. Meid. *Studies in Greek, Italic, and   Indo-European linguistics : offered to Leonard R. Palmer on the occasion of his seventieth birthday, June 5,   1976.* Innsbruck : Inst. f. Sprachwissenschaft d. Univ. 1976, 33-41.

--- and M. B. Emeneau, *A Dravidian Etymological Dictionary.* Oxford: Clarendon Press 1960   (*DED*)

---, *A Dravidian Etymological Dictionary.* Second Edition. Oxford: Clarendon Press 1984   (*DEDR*)

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

Das, R. P. The hunt for foreign words in the Ṛgveda. *IIJ* 38, 1995, 207-238

Diakonoff, I. M. Hurro-Urartian Borrowings in Old Armenian. *JAOS* 105, 1985, 597- 604

Driem, G. van. On the Austroasiatic Indus Theory. *MT* 4, Oct. 1999, appendix 1-9

Ehret, Christopher. Language change and the material correlates of language and ethnic shift. *Antiquity* 62, 1988, 564-74

Elfenbein, J.H. A periplous of the 'Brahui problem'. *Studia Iranica* 16, 1987, 215-233

Emeneau, M. B. India as a linguistic area. *Language* 32, 1956, 3-16

Erdosy, G. (ed.). *The Indo-Aryans of Ancient South Asia.* (Indian Philology and South Asian Studies, A. Wezler and   M. Witzel, eds., vol. 1). Berlin/New York : de Gruyter 1995

Fairservis, W. A. *The Harappan Civilization and its Writing. A Model for the Decipherment of the Indus Script* New   Delhi: Oxford 1992

---, Central Asia and the Ṛgveda: the archaeological evidence. In: G. Erdosy (ed.) *The Indo-Aryans of Ancient South   Asia.* Berlin/New York : de Gruyter 1995, 206-212

---, The Harappan Civilization and the Ṛgveda. In: M. Witzel (ed.) *Inside the texts, Beyond the Texts. New   Approaches to the Study of the Vedas.* (Harvard Oriental Series, Opera Minora 2). Cambridge 1997, 61-68.

Falk, H. Vedisch *árma. Zeitschrift der deutschen morgenländischen Gesellschaft* 131, 1981, 160-171

---, The Purpose of Ṛgvedic Ritual. In: M. Witzel (ed.) *Inside the texts, Beyond the Texts. New Approaches to the Study   of the Vedas.* (Harvard Oriental Series, Opera Minora 2). Cambridge 1997, 69-88

Gardner, J. R. *The Developing Terminology for the Self in Vedic India.* Ph.D.Thesis, Iowa U., 1998

Gening, V.F. Mogil'nik Sintashta i problema rannikh Indoiranskikh plemen. *Sovetskaya Arkheologiya* 1977, 53-   73

Hillebrandt, A. Vedische Mythologie. Breslau 1927-29.

Hock, H.H. Substratum influence on (Rig-Vedic) Sanskrit? Studies in the Linguistic Sciences 5, 1975, 76-125

Hoffmann, Karl. *Die alt-indoarischen Wörter mit -ṇḍ-, besonders im Ṛgveda.* Ph. Diss. München 1941

---, *Aufsätze zur Indoiranistik.* (ed. J. Narten, vols.1-2) Wiesbaden. 1975-76

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

Iyer, L.V. Ramaswamy, Dravidic place names in the plateau of Persia, *Quarterly Journal of the Mythic Society of   India* 20, 1929-30, 49-53.

Jarrige, J.-F. Continuity and Change in the North Kachi Plain (Baluchistan, Pakistan) at the Beginning of the Second Millennium B.C., in: *South Asian Archaeology 1983*, ed. J. Schotsman and M. Taddei. Naples: Istituto   Universitario Orientale, 1985, pp. 35-68.

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

Koppers, W. *Die Bhil in Zentralindien.* Horn: F. Berger 1948

Krishnamurty, Bh. Patterns of Sound Change in Dravidian. offprint 1998

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

---, *An Austro-Asiatic myth in the RV.* Amsterdam : Noord-Hollandsche Uitg. Mij. 1950.

---, The Genesis of a Linguistic Area. *IIJ* 10, 1967, 81-102

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

---, *Nahali, A comparative Study.* Amsterdam 1962

---, The sources of Nahali vocabulary. In: N.H. Zide (ed.), *Studies in comparative Austroasiatic Linguistics.* The   Hague 1966, 96-192

---, *Aryans in the Rigveda,* Amsterdam-Atlanta: Rodopi 1991

---, On a Hunt for 'Possible' Objections. *IIJ* 38, 1995, 239-247

Littauer, M. and Crouwel, J. H. *Wheeled vehicles and Ridden Animals in the Ancient Near East,* Leiden : Brill 1979

Lubotsky, A. Indo-Iranian Substratum. forthc. in: Proceedings of the Conference "Contacts between Uralic and   Indo-European in the Neolithic, Eneolitic and Bronze Ages in the Light of Linguistics and   Archaeological Evidence" at Tvarminne, 8-10 January 1999

Masica, C. P. Aryan and non-Aryan elements in North Indian agriculture. In: M. Deshpande,   P.E. Hook (eds.).   *Aryan and Non-Aryan in India.* Ann Arbor : Center for South and Southeast Asian Studies, University of   Michigan 1979, 55-151.

Mayrhofer, M. *Kurzgefasstes etymologisches Wörterbuch des Altindischen.* Heidelberg 1956-1976. (KEWA)

---, *Wörterbuch des Altindoarischen.* Heidelberg 1986-96 (EWA)

---, Über den spontanen Zerebralnasal im frühen Indo-Arischen. In: *Mélanges d'Indianisme.* (= *Fs. L. Renou),* Paris   1968, 509-517

McAlpin, D. W. Linguistic prehistory: The Dravidian situation, in Deshpande and Hook 1979, 175-189

---, *Proto-Elamian-Dravidian: the evidence and its implications.* Transactions of the American Philosophical
        Society, 71, Philadelphia 1981

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

Morgenstierne, G. *Irano-Dardica.* Wiesbaden 1973

Mughal, Mohammad Rafique. *Ancient Cholistan. Archaeology and Architecture.* Rawalpindi-Lahore-Karach
Ferozsons   1995

Oberlies, Th. *Review Article: F.B. J. Kuiper: Aryans in the Rigveda.* IIJ 37, 1994, 333-349.

Oldenberg, Hermann. *Die Hymnen des Rigveda, Band I. Metrische und textgeschichtliche Prolegomena,* Berlin
:   Wilhelm Hertz 1888.

Palaniappan, S. personal communication, Sept. 1999

Parpola, A. Interpreting the Indus script. In: Lal, B. B. and S. P. Gupta, *Frontiers of the Indus Civilization: Sir
        Mortimer Wheeler commemoration volume.* New Delhi 1984: 179-191

---, *Deciphering the Indus script.* Cambridge: Cambridge University Press 1994.

---, The Dāsas and the Coming of the Aryans. In: M. Witzel (ed.) *Inside the texts, Beyond the Texts. New
        Approaches to the Study of the Vedas.* (Harvard Oriental Series, Opera Minora 2). Cambridge 1997,
193-   202

---, Deciphering the Indus Script: Basic methodology and some sample interpretations. Forthc. in: *Indo-
Kōko   Kenkyū / Indian Archaeological Studies* 21, 1999-2000

Piggott, S. *The earliest wheeled transport: from the Atlantic coast to the Caspian Sea.* London: Thames &
Hudson   1992

Pinnow, K.H., Zu den altindischen Gewässernamen. *Beiträge zur Namensforschung* 4, 1953, 217-234; 5,
1954, 1-19.

---, *Versuch einer historischen Lautlehre der Kharia-Sprache*, Wiesbaden 1959

Possehl, G. *Indus Age. The writing System.* Philadelphia: University of Pennsylvania Press 1996

---, Meluhha. in: J. Reade (ed.) *The Indian Ocean in Antiquity.* London: Kegan Paul Intl. 1996a, 133-208

---, The Transformation of the Indus Civilization, Journal of World Prehistory 11, 1997, 425-72

Proferes, Th. *The Formation of Vedic liturgies.* Harvard Ph.D. Thesis, 1999

Rau, W. The Meaning of *pur* in Vedic Literature, [Abhandlungen der Marburger Gelehrten Gesellschaft
III/1]   München : W. Finck 1976

---*, Zur vedischen Altertumskunde,* Akademie der Wissenschaften zu Mainz, Abhandlungen der Geistes- u.
        sozialwissenschaftlichen Klasse 1983, No. 1. Wiesbaden : F. Steiner 1983

Reinhard, J. Aperçu sur les Kusuṇḍā. *Objets et Mondes* 9, 1969, 89-106

Sankalia, H. D. *Studies in the historical and cultural geography and ethnography of Gujarat (places and peoples
in   inscriptions of Gujarat: 300 B.C. - 1300 A.D.)* Poona: Deccan College 1949

Shafer, R. Nahālī, A linguistic study in paleoethnography. *Harvard Journal of Asiatic Studies* 5,   1940,
346-371

---, *Ethnogeography of Ancient India,* Wiesbaden 1954

Shaffer, J.G. The Indo-Aryan invasions: Cultural myth and archaeological reality. in: *The people of South
Asia*,   ed. J. R. Lukacz, New York 1984, pp. 77-90

---, and Diane A. Lichtenstein. The concepts of "cultural tradition" and "palaeoethnicity" in South Asian
        archaeology. In: G. Erdosy (ed.) 1995, 126-154.

---, Migration, Philology and S. Asian Archaeology.
        in : J. Bronkhorst & M. Deshpande, *Aryan and Non-Non-Aryan in South Asia.*   *Evidence,
Interpretation   and Ideology.* Harvard Oriental Series. Opera Minora, vol. 3. Cambridge 1999, 337-404

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

---, The reconstruction of Prehistoric South Asian language contact, in E. H. Bendix (ed.), *The Uses of Linguistics.* New York: New York Academy of Sciences 1990, p. 207-234

---, Reconstructing social context from language: Indo-Aryan and Dravidian prehistory. In: *The Indo-Aryans of Ancient South Asia*, ed. G. Erdosy (ed.). Berlin/New York : de Gruyter, 1995, 258-277

Szemerenyi, O. Introduction to Indo-European linguistics. Oxford : Clarendon Press / New York : Oxford University Press 1996.

Toba, T. *A brief introduction to Kusunda.* Kathmandu : Summer Institute of Linguistics 1971

Vallat, F. Eléments de géographie élamite (résumé) PO 11, 1985, 49-54

Wells, B. *An Introduction to Indus Writing.* MA. Thesis, U. of Calgary 1998
[2nd ed.: Early Sites Research Society (West) Monograph Series, 2, Independence MO 1999]

Witzel, M. Die Kaṭha-Śikṣā-Upaniṣad und ihr Verhältnis zur Śīkṣāvallī der Taittirīya-Upaniṣad. *WZKS* 24, 1980, 21-82

---, Tracing the Vedic dialects. In: Colette Caillat (ed.), *Dialectes dans les littératures indo-aryennes.* Paris : Institut de Civilisation Indienne 1989, 97-264

---, Nepalese Hydronomy. Towards a history of settlement in the Himalayas. G. Toffin (ed.), Nepal, Past and Present : *Proceedings of the Franco-German Conference, Arc-et-Senans, June 1990.* New Delhi : Sterling Publishers 1993.

---, Early Indian history: Linguistic and textual parameters. In: *The Indo-Aryans of Ancient South Asia*, ed. G. Erdosy (ed.), = Indian Philology and South Asian Studies, ed. A. Wezler and M. Witzel, vol. 1. Berlin/New York : de Gruyter 1995, 85-125

---, Ṛgvedic history: poets, chieftains and polities. In: *The Indo-Aryans of Ancient South Asia*, ed. G. Erdosy (ed.) 1995, 307-352.

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

---, , 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 1997, pp. ix-xxvi

---, Early Sanskritization. Origins and development of the Kuru State. In: B. Kölver (ed.), *Recht, Staat und Verwaltung im klassischen Indien. The State, the Law, and Administration in Classical India.* München : R. Oldenbourg 1997a, 27-52

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

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

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

Zide, N. H. Munda and non-Munda Austroasiatic Languages, in: Current Trends in Linguistics, 5. The Hague: Mouton 1969, 411-430

Zvelebil, K. The descent of the Dravidians. *International Journal of Dravidian Linguistics* 1-2, 1972, 56-63

---, *Dravidian Linguistics: an Introduction.* Pondicherry: Pondicherry Institute of Linguistics and Culture 1990

=========