

MOTHER TONGUE  
Special Issue, Oct. 1999

## Early Sources for South Asian Substrate Languages

by

**Michael Witzel**

Department of Sanskrit and Indian Studies  
Harvard University (witzel@fas.harvard.edu)

*Ut somniorum interpretatio,  
ita verborum origo;  
pro cuiusque ingenio iudicatur\**  
Augustinus, 354-430 AD

*Es war die etymologische Arbeit,  
was am Anfang der Linguistik  
als Wissenschaft stand\*\**  
V.I. Abaev, 1952 CE

The recent articles in *Mother Tongue* on the isolated South Asian languages Burushaski, Nahali, and Kusunda offer a welcome peep into the complicated linguistic prehistory of the subcontinent. South Asia is, also in its genetics (L. Cavalli-Sforza 1994), a text book case for the continuing coexistence of many subsequent levels of immigrants. In fact, the subcontinent offers a virtual laboratory of linguistic, cultural and social systems. To echo H.C. Fleming, *MT* II 74: "... given India's role as cultural diffusion cul-de-sac of Asia, ... we may have missed the lower strata of prehistory after all!" Such items have kept me occupied, on and off, over the past few years. I offer some additional data here, and I will draw attention to some other remnants of ancient South Asian languages, most of which have come down to us only as substrates. In the following pages I will be brief with regard to cases that have been noticed before (Burushaski, Nahali and Kusunda, *MT* II and III), but I will add data from substrates not yet adequately recognized. Obviously, the more remote

---

\* "Just as the interpretation of dreams, thus the origin of words : it is determined according to one's own inclination (or, 'talent')."

\*\* "It was etymological work that constituted the beginning of linguistics as a science", p.39 in: V. I. Abaev, *Die Prinzipien eines etymologischen Wörterbuchs*. Österreichische Akademie der Wissenschaften, Phil.-Hist. Kl., Sitzungsberichte 368, Heft 11, 1980, 29-45. German translation of the Russian paper in *Voprosy Yazykoznaniiya* 1952/5, 50 sqq. -- I thank J. Bengtson, H. Fleming, and R. Wescott for their very careful reading of a slightly earlier version of this paper, and for many suggestions and improvements in style and substance. Any remaining errors are of my own making. I also thank the discussants; the paper was written, indeed, to facilitate such discussions: in the traditionally multilingual South Asia, such cooperation is sorely needed.

data we can acquire in this fashion, the better our chances will be for the reconstruction of the early settlement of (South) Asia and for the languages spoken by the first modern humans that entered the area several ten thousand years ago (see Cavalli-Sforza 1994).

I will concentrate on those areas of the subcontinent that are best known from early sources (the Vedas), that is the Panjab and parts of the Gangetic plains, and I will pay special attention those items that allow us to place such linguistic data in *place and time*. For there is testimony enough for a number of additional languages that are of importance in this early period; they indicate that we have to rethink the substrate and adstrate relationships of the South Asian languages, even those belonging to the three major language families (Indo-European, Dravidian, Austro-Asiatic) found there.

All are in need of more detailed study by area specialists, who should provide the philological and linguistic background information so that comparativists can make use of them. I propose to do so, incrementally, for a few of the substrate languages mentioned below, especially those found attested in early texts (Vedic, Epic Sanskrit, and Pali).

These texts provide our most ancient sources for non-Indo-European (that is, non-Indo-Aryan) words in the subcontinent. The Vedas were orally composed (c. 1500-500 BC) in northern Pakistan and northern India. They are followed by Dravidian sources represented by the ancient Tamil "Sangam" (*Caṅkam*) texts of South India (from the beginning of our era); these are virtually unexplored as far as non-IA and non-Drav. substrates and adstrates are concerned. From a slightly earlier period come the Middle Indo-Aryan (MIA) Pali canon and the Epic texts (Mahābhārata, Rāmāyaṇa).

Since I am not a Dravidian specialist, I will concentrate on the Vedas, which are earlier than Drav. texts by at least a thousand years, and contain a host of so far comparatively little studied data. This procedure also has the advantage that the *oldest* linguistic data of the region are used, which is important because of the quick changes that some of the languages involved have undergone. Such changes obscure the relationships and make comparisons, based on later attested forms, more difficult (cf. below, §8, on Semitic loans, and cf. P. K. Benedict, *MT* III 93). So far, linguists have concentrated on finding Dravidian and Munda reflexes, especially in the oldest Veda, the Ṛgveda (RV). 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 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.

## Sources

The oldest text at our disposal is the Ṛgveda (RV), in archaic Indo-Aryan (Vedic Sanskrit). It is followed by a number of other Vedic texts, usually listed as Saṃhitās, Brāhmaṇas, Āraṇyakas and Upaniṣads. Linguistically, however, we have to distinguish five distinct levels: Ṛgveda, other Saṃhitās (Mantra language), Yajurveda Saṃhitā prose, Brāhmaṇas (incl. Āraṇyakas and Upaniṣads) and the late Vedic Sūtras (Witzel 1987, 1997; for abbreviations of names of texts, their dates and their geographical location see attached list).

At the outset, it must be underlined that the Vedic texts excel among other early texts of other cultures in that they are "tape recordings"<sup>1</sup> of this archaic period. They were not allowed to be changed: not one word, not a syllable, not even a tonal accent. If this sounds unbelievable, it may be pointed out that they even preserve *special* cases of main clause and secondary clause intonation, items that have even escaped the sharp ears of early Indian grammarians. These texts are therefore better than any manuscript, and as good -if not better- than any contemporary inscription.

Consequently, these texts are invaluable as early sources for non-IA loan words in Vedic Sanskrit. Recently, F.B.J. Kuiper (1991) has prepared a very valuable collection of some 380 'foreign' words found in the RV. However his intention, in this particular book, was not to present etymologies but to demonstrate their non-IA type by phonetic and structural analyses. Some words indeed stand out immediately because of their non-IA phonetical shape (Burrow 1976), for example *busa* 'chaff, fog?' RV(cf. Pinnow 1959: 39), *ṛbīsa* 'oven/pit with coals, volcanic cleft' RV, *Bṛsaya* 'name of a sorcerer or demon' RV, *musala* 'pestle' AV, *kusīda* 'lending money' KS, TS, *Kusurubinda* 'name of a clan' TS, *Kosala* 'name of the Oudh territory' ŚB, etc. In IA, *s* is not allowed after (long or short) *i, u, e, ai, o, au, ṛ* and *k*. Many of the other words investigated by Kuiper (1991) are clearly of non-IA origin, but often neither of Munda or of Drav. origin. Kuiper occasionally gives Drav. and Munda etymologies but he also cautiously states that the word in question must belong to some unknown language. I think we can proceed further on this basis by adding a growing number of words from the later Vedic texts, especially from the more popular Atharvaveda, which contains several hundred sorcery spells abounding in non-IA words. The ensuing periods of Yajurveda Saṃhitās, Brāhmaṇas, Upaniṣads, and Sūtras, (see Witzel 1987, 1989 for geographical spread and chronology) have a large number of so far little studied loan words.

In the sequel, I will proceed geographically, region by region, indicating, in each case, the source of our knowledge. (Vedic accent marks are omitted, as they play no role in foreign words, see Kuiper 1991.)

### §1. The Northwest

This is the area of the first Indo-Aryan influx into the subcontinent reflected by the hymns of the RV. It includes the mountainous regions of Afghanistan and Northern Pakistan as well as the plains of the Panjab. In the Veda we find few place names; river names, as ancient tribal boundaries, are much better attested. However, the Ṛgvedic area is characterized by an almost total substitution of local river names by those of IA type, such as *Gomatī* 'the one having cows' (mod. *Gomal*), *Mehatnu* 'the one full of fluid', *Asiknī* 'the black one' (now *Chenāb*).

Tribal names, much more difficult to locate, complement this account. Next to typical IA ones (*Druhyu* 'the cheaters', *Bharata* 'the ones who carry (sacred fire?).' There are many that have no plausible IA etymologies, including names such as: the *Gandhāri* tribe of *Gandhāra*, the area between Kabul and Islamabad in Pakistan; *Śambara*, a mountain

---

<sup>1</sup> A restricting factor is the middle/late Vedic redaction of the texts in question. However, this influenced only a very small, well known number of cases, such as the development : *Cuv* > *Cv*.

chieftain; *Vaiyiu* and *Prayiyu* (chieftains on the *Suvāstu*, modern Swat); *Mauja-vant*, a Himalayan peak. This is the typical picture of an intrusive element, the IA, overlaying a previous population. Unlike Northern America for example, only a few pre-IA river names have survived, such as: *Kubhā* (mod. Kabul river), *Krumu* (mod. Kurram), and maybe even the *Sindhu* (Indus); these have no clear or only doubtful IA/IE etymologies (see below).

North of this area, at the northern bend of the Indus (Baltistan/Hunza), Burushaski is spoken. However, the language and the tribal name are indirectly attested in this general area ever since the RV: *\*m/bružā* (mod. *burušo*) > Ved. *Mūja-vant*, Avestan *Muža* (see below).

However, already the RV contains a few words which are still preserved in Bur., such as Bur. *kilāy*, Ved. *kilāla*- 'biestings, a sweet drink' RV 10.91.14, (note AV 4.11.10 next to the loan word *kināśa*, see below); *kilāla* cannot have a IA etymology (EWA I 358 'unclear'); continuants are found in Dardic (Khowar *kilāl*), Nuristani (*kilā* etc.), in later Skt. *kilāṭa* 'cheese', cf. DEDR 1580 Tam. *kilāan* 'curd'). For details see Kuiper 1955: 150f., Turner, CDIAL 3181, Tikkanen 1988. Further the following words (mostly treated in some detail further below),

- *mēṣ* 'skinbag', CDIAL 10343 < Ved. *\*maiṣiya* 'ovine', *meṣa* 'ram' RV,
- *gur* 'wheat' pl. *guriṅ/gureṅ* < *\*yorum*, *gurgán* 'winter wheat', cf. Ved. *godhūma*,
- *bras* 'rice', different from *briú* 'rice' (< Shina *briú*), cf. Ved. *vrihi*,
- *bus* 'sheaf', CDIAL 8298, cf. Ved. *busa*, *br̥sī* 'chaff'
- *ku(h)á* (Berger *γúá*) 'new moon', cf. Ved *kuhū* 'deity of new moon'.
- *γupas* (Berger *gupás*) 'cotton', cf. Ved. *karpāsa*, Kashm. *kapas*,
- *baluqa* 'stone' (in a children's game), cf. *báltaṣ* 'stone thrown at someone', cf. Ved. *paraśu* '(stone) ax', Greek *pélekus*, see EWA II, 214; J. Bengtson, by letter of 4/19/99, draws my attention to PEC *\*bølv̥gwi* 'hammer' > Chechen *berg* 'pickax', Archi *burk* 'hammer'; as for *baluqa*, *báltaṣ* he also draws attention to PEC *\*bəHə́ V* 'hill, mountain' > Rutul *bāl* 'rock', etc.
- *bañ* 'resin of trees' ~ Iir *bhaṅga* 'hemp, cannabis', cf. Khowar *boñ*, or rather, with J. Bengtson (by letter) to be compared with PEC *\*bhinkwV* 'pine tree' > Ingush *baga* 'resinous root of pine tree'.

In Proto-Burushaski (or in its early loans from the lowlands) and in the pre-Vedic Indus language there is interchange of *k/ś*, and retention of *-an-* (not > *-o-*, see below): Bur. *kilāy* : Ved. *kilāla*, but *šon* 'blind one-eyed' : Ved. *kāna*;

- *γoro* (Berger *γuró*) 'stone, pebbles', cf. Ved. *śar-kara*, cf. also (Witzel 1999) *yoqares*, Berger *yókurac* 'raven', Ved. *kāka*; Ved. *γású* 'onion', cf. Ved. *laśuna*, Shina *kaśu*; J. Bengtson informs me, by letter of 4/19/ 99, of the following Caucasian connection: PNC *\*lem̥zi* 'garlic' > Andi *raži*, Lak *lã :i*, or alternatively also Bur. *γású* and Basque *hausin* ~ *asun* 'nettle'; -- cf. also (?) Bur. *yon*, Berger *γúun* 'quail' with Ved. *laba*?

Most of the words from IA languages in Turner's CDIAL that have Bur. correspondences are, however, late loan words from the neighboring Dardic languages, especially from Shina and Khowar (cf. Lorimer 1937, Berger 1959, 1998). I merely mention those which are restricted to the Northwest and may have local substrate origins:

- *bəru* CDIAL 11313 < Ved. *varaṭa*, *barāṭa* 'seed of safflower' GS,

- *chomar, chumer, chumər* 'copper' 14496 : Skt. *cīmara-kāra*, Nur. (Ashkun) *ciməkāra*, Khowar *cūmur*, Shina *cimer* etc.
- *ḍiru* 14547 < Shina *ḍiru, diḍu* < \**dhiddha* 'belly'
- *gindāwər* 4199 < Shina *gunēr* 'small tree with red berries', Skt. *gundra* 'Saccharum sara'
- *gupās* 2877 < Kashm. *kapas*, etc., Late Ved., Skt. *karpāsa* 'cotton plant'
- *kuyōc* 'subjects of a ruler' 14404 < Shina *kuōc, kuiōch, \*kūpatya*
- *mēṣ* 'skinbag' 10343 < \**maiṣiya* 'ovine', Ved. *meṣa* 'ram'
- *sinda* 13415 < \**sind* (> Shina *sin* 'river', Dumaki *sina* 'river'), Ved. *sindhu*
- *tayay* 5626 < Shina *tagá* 'mud' \**tagga* 'mud'

Unfortunately the new dictionary by H. Berger does not contain etymological annotations going beyond CDIAL. For some initial ideas, see Witzel 1999. Further early evidence comes from the names in the Gilgit inscriptions and the Gilgit manuscripts of the later first mill. CE. (see v. Hinüber 1980, 1989, cf. Tikkanen 1988).

It has occasionally been maintained that Burushaski extended into the Panjab in earlier times (L. Schmid 1981, Tikkanen 1988), but the Vedic evidence does not support this. We cannot be sure exactly how far Ṛgvedic geographical knowledge extended northwards, and how much practical interaction existed between RV and Proto-Burusho people. Yet, the RV knows of some small right side contributory rivers of the Indus that are located north of the confluence with the Kabul River; they have IA names: RV 10.75.6. *Trṣṭāmā* '< *trṣ* 'the rough, (or) the dried up (river)', *Susartu* 'the one running well', *Rasā* 'the one full of sap', *Śvetī* 'the white one'.

While it is questionable how far south Burushaski territory extended at this early time, some of the loan words mentioned above indicate that there was early contact. That extends perhaps also to medicinal and other herbs (cf. below on *Kirāta*), for it may be that the name of the Burušo is reflected by the RV mountain name *Mauja-vant* "having *Mūja* (people)", cf. the east Iranian equivalent, Avestan *Muža*. This is the mountain where the best Soma, a hallucinogenic plant, comes from. The RV and E. Iranian (Avestan) forms look like adaptations of the local self-designation, \**Mruža*, Vedic *mūja-*, Avest. *muža*, and are attested since the middle of the first millennium in early Tib. *bru-ža*, Sanskritized *puruṣa* (von Hinüber 1989, 1980), local 10th cent. inscriptions *prūśava* (Jettmar 1989: xxxvii), mod. Bur. *Burušo*.

Phonetic reflexes of Bur. have been seen (Tikkanen 1988) in the Vedic (and Dravidian) retroflex consonants that have otherwise found a number of explanations, from a Dravidian substrate to an internal East Iranian and Vedic development. The occurrence of these sounds clearly reflects an areal feature that is strongest in the Northwest, but extends all the way to Tamil in the South, and has also influenced Munda to some extent. Below, it will be shown that it is an ancient feature of the Indus language as well, and that it must not be traced back to Bur. influence, which seems to have been limited, even in Ṛgvedic times, to the upper Indus valley.

Some early syntactic influence by Burushaski on Vedic in the formation of the Absolutive has been assumed by Tikkanen (1988); it is found already in earliest RV but only as past verbal adverb/conjunctive participle. This clearly S. Asian feature, unknown in the sister language of Vedic, Old Iranian, is also found in various degrees in Drav. and Munda,

and may have been an early regional feature whose ultimate origin remains unclear (cf. Witzel 1999)

## §2. Khowar

Another modern language in the same area is Khowar which belongs, along with Kashmiri, Swati, etc. to the Dardic branch of IA. In its phonetics and vocabulary, however, it shows a strong local substrate, similar to Burushaski. Unique for Khowar, however, is a particular substrate whose origin remains unclear so far. It seems that the Khowars are a late immigrant group who have taken over a Dardic language. Substrate(?) words in Khowar which are neither IA nor Burushaski include (Kuiper 1962: 11, cf. Morgenstierne 1947: 6, Lorimer 1935 : xxi): *ȳec* 'eye', *ap'ak* 'mouth', *krem* 'back', *camoṭh* 'finger', *iskī* 'heel', *askār* 'lungs'.

Kuiper (1962: 14) compares *ȳec* 'eye' with Bur. *ȳai(c)-*, *ȳ'i-*, *ȳe-ic-* 'to appear, seem, be visible', and with *g'e-* 'to look, seem, appear', *da-g'e-* 'to peer' of the Munda language Sora and with Parengi *gi-* 'to see'. (Differently, Morgenstierne, FS Belvalkar, 2nd section p. 91.)

For Bur. loans in Dardic and in Nuristani see Tikkanen 1988: 305 (*cumar* 'iron', *ju* 'apricot', etc.), cf. Fussman 1972 II, 37 sqq.; Lorimer 1938: 95, Morgenstierne 1935: xxi sqq., 1947: 92 sqq.; Schmidt 1981, Berger 1998.

Finally, one must be open to assume the influence of other substrate languages in the Hindukush/Pamir areas. There are local personal names such as RV *Śambara Kaulitara* and his father \**Kulitara* who are 'in the mountains', *Prayiyu* and *Vayiyu* in Swat; names of demons (as always, intentionally confused with those of real, human enemies) such as *Cumuri*, *Namuci*, *Uraṇa*, *Arbuda*, *Pipru*, *Śambara*; tribal names such as *Gandhāri*, *Dṛbhika(?)*, *Varc-in(?)*; river names such as *Gandhāra*, *Krumu*, *Sindhu(?)*. Note also that the Avesta (V.1) speaks about some of these areas, notably *Varəna* (Varṇu) as *an-airiia* 'non-Aryan'.

## §3. The Kashmir substrate.

The prehistory of Kashmir is little known. In the Neolithic, there were relations with Central Asia and China, but the influence of the Indus civilization (2600-1900 BCE) is strong and long-lasting; of course, this does not tell us anything about the language(s) spoken then. Unfortunately, the Vedic texts, which know of the neighboring Indus valley do not mention Kashmir by name. It is first mentioned by name only by the grammarian Patañjali (150 BCE). The native Kashmiri texts (*Rājatarāṅgiṇī*, *Nilamata Purāṇa*, cf. Witzel 1994, Tikkanen 1988, L. Schmid 1981), however, know of the previous populations, the *Pisāca* 'ghouls' and the *Nāga* 'snakes' (that can change into human shape at will). These are common Indian names for 'aboriginals'; cf. the Tib.-Burm. Naga tribe on the Burmese border. Yet, these designations may retain some historical memory. The chief of the *Pisāca* is called *Nikumbha* (*Nikumba* in *Milindapañho*), and the *Nāgas* have such 'foreign' names such as *Karkoṭa*, *Aṭa*, *Baḍi*, *Bahabaka*, *Caṭara*, *Cikura*, *Cukkaka*, etc. The list of some 600 Kashmir *Nāga* names in the local *Nilamatapurāṇa* contains many such non-Sanskritic names; they have not been studied (see Witzel, in press).

An interesting case is that of a tribe in or near Kashmir that is attested only in 550/600 CE: the *Kīra* (*Bṛhatsaṃhitā* 14.29, c. 550 CE). Its name is close to that of the *Kirāta* who are attested in the early inscriptions of Nepal (464 CE sqq.) but who already appear in

the Atharvaveda (c. 1200 BCE). Hsuan Ts'ang, Hsiyuki (c. 600 CE, cf. T. Funayama 1994: 369) knows of them as *Kilito* (Karlgren 1923, no. 329-527-1006), a people in Kashmir who had their own king shortly before his time. The *-ta/ -ṭa* suffix is common in many North Indian tribal names (Witzel 1999, cf. below).

The rich medieval Kashmiri literature in Skt. has preserved other substrate words, such as the river and place names: *Ledarī*, a river in the SE of the Valley (also in the place name *Levāra* < *Ledarī-agrahāra*); *-muṣa*, a 'suffix' in the names of several villages: *Khonamuṣa* (mod. Khun<sup>a</sup>moh), *Katīmuṣa*, (mod. Kaimoh, next to *Kati-kā*), *Rāmuṣa* (mod. Ramuh); also, the *Pañcāla-dhāra* mountain, (mod. (Pir) *Pantsāl* range, south of the Valley), may reflect an old name, cf. the Ved. tribal name *Pañcāla*, and Grierson, Dict. of Kashmiri III : 744; cf. Nepali *himāl* 'Himalaya range', CDIAL 14104. Such names have not been studied in detail (cf., however, L. Schmidt 1981, Witzel 1993).

Just as in Northern India and Nepal, most river and place names in Kashmir have been Sanskritized, or they have been transmitted in their Middle Indian forms (e.g., the *Mahurī* river in N. Kashmir < Skt. *madhurī* 'the sweet one'. Frequently, like many Indian place names, they have been "telescoped" beyond recognition (e.g. Ved. *Kāpiṣṭhala* > Kaithal, *Rohitakakūla* > Rohtak, Class. Skt. *Pāṭaliputra* > Patna, *Nāgapura* > Nagor, *Indrapaṭṭana* > Indarpat, or the river (Pali) *Sundarikā* > Sai); thus we have, in Kashmir: *Kuru-agrahāra* > Skt. *Kuruhāra*, *Levāra* (above). Such shortening is not unheard of elsewhere (e.g., New Orleans [nɔ'ɔrlɪnz] or [aʃəberg] for Aschaffenburg near Frankfurt; Worcester [wüstə], or as John Bengtson tells me, Engl. Featherstonehaugh [fænʃɔ], Cholmondeley [ɔmlɪ], cf. further below, on Nepal); however, this feature seriously affects the interpretation of river and place names in S. Asia when we do not have early sources.

The Kashmiri language itself has not been thoroughly scrutinized for more substrate materials, cf., however, the report by L. Schmidt (1981), who assumes that 25% of the vocabulary and toponymy belong to a pre-IA substrate. A. Parpola (Tikkanen 1988: 305) thinks of a Proto-Tib. or Sinitic substrate. However, the peculiar phonology of Kashmiri (and Dardic in general) sustains the assumption of a strong *northwestern* substrate influence.

We now turn to a region for which we have larger amount of early sources, the Greater Panjab.

#### §4.1. The Greater Panjab

The RV reflects the Panjab and its immediate surroundings of c. 1500-1200 BCE., most clearly visible in its river names, extending from the Kabul River to the Yamunā (mod. Jamna) and even the Ganges (*Gaṅgā*, mentioned only twice).

In order to use the linguistic evidence contained in this text properly, it is important to realize that it has been composed not just in two layers ('main' and 'late', as found in the handbooks), but in *three* clearly distinguishable, and very roughly datable layers (Witzel 1999, J. R. Gardner, Thesis Iowa U., 1998, Th. Proferes, Ph.D. Thesis, Harvard U., 1999, Witzel 1995):

- I. the *early Ṛgvedic period*: c. 1700-1500 BC, especially the hymns in books 4, 5, 6 (and maybe book 2);
- II. the important *middle Ṛgvedic period*, c. 1500-1350 BC: RV 3, 7, parts of 8.1-66 and 1.51-191;
- III. the *late Ṛgvedic period*, c. 1350-1200 BC: RV 8.67-103; 1.1-50; 10, 8.49-59.

It is important to note that level I has no Dravidian loan words at all (details, below); they begin to appear only in level II and III.

Instead, we find some three hundred words from one or more *unknown* languages, especially one working with prefixes. Prefixes are typical neither for Drav. nor for Burushaski (cf. Kuiper 1991: 39 sqq., 53). Note that the "prefixes" of Tibeto-Burm. (Benedict 1972) do not agree with those of the RV substrate either. Their presence apparently excludes also another unknown language which occasionally appears in the RV and more frequently later on with typical gemination of certain consonant groups (perhaps identical with Masica's "Language X" (1979), see below; cf. Zide and Zide 1973:15). The prefixes of the RV substrate are, however, close to, or even identical with those of Proto-Munda; taking my clue from Kuiper (1962: 51,102; but see now Zide MT II, 1996, 96), I will therefore call this substrate language *Para-Munda* for the time being.

#### §4.2. Para-Munda loan words in the Ṛgveda

We can start with the convenient list of Kuiper (1991), who does not, however, discuss each of the 383 entries (some 4% of the hieratic RV vocabulary!) This list has been criticized by Oberlies (1994) who retains "only" 344-358 words, and minus those that are personal names, 211-250 'foreign' words.<sup>2</sup> One can, of course, discuss each entry in detail (something that cannot be done here), but even Oberlies' lowest number would be

---

<sup>2</sup> Oberlies' criticism is written from an IE-centered point of view similar to that of Mayrhofer (EWA). This is fine from the point of view of someone who has to write an etymological dictionary of OIA; however, due to the clear attestation of cultural, ethnical and religious amalgamation of IIR/IA and local elements visible already in the oldest IA text, the RV, the existence of such a large number of 'foreign' words must not be minimized in its importance. Nor does Oberlies offer an explanation or analysis of the remaining 250 words; they are simply 'non-IA'. In a similar vein, R.P. Das has written a much more 'engaged', nit-picking review of Kuiper's book, tellingly entitled 'The hunt for foreign words in the Ṛgveda' (*IJ* 38, 1995, 207-238), which induced Kuiper to write a well-deserved, rather scathing reply in the same volume ("On a Hunt for 'Possible' Objections". *IJ* 38, 1995, 239-247). It is difficult to understand, in view of the well-known evidence (added to in this paper), how one can regard the language (and religion, culture) of the Ṛgvedic Arya as 'relatively free from foreign influences' (Oberlies 1994: 347). "Pristine" languages and cultures do not exist, nor did they at c. 1500 BCE.



significant enough, in a hieratic text composed in the traditional poetic speech of the Indo-Iranian tradition, to stand out, if not to surprise. It is much more difficult to discern Munda/Austro-Asiatic words, and to distinguish them from those of an unknown local substrate (remnants of the Gangetic "Language X"), or the still unknown language of the Indus inscriptions than to establish IA or Dravidian etymologies, as an etymological dictionary of Munda is still outstanding (in preparation by David Stampe et al.). One can also sympathize with Kuiper (1991: 53): "Burrow and Emeneau understandably and rightly ignore the Pan-Indic aspects, but ... their dictionary [DEDR], by omitting all references to Munda, sometimes inevitably creates a false perspective from a Pan-Indic point of view." Nevertheless, one can, for the time being, make use of Pinnow's reconstructions of Proto-Munda in his investigation of Kharia (1959), Bhattacharya's short list (1966: 28-40), Zide & Zide's discussion of agricultural plants (1973, 1976), and Kuiper's relevant studies (especially 1955, 1991; his 1948 book is still useful, in spite of his own disavowal of it, as a collection of relevant materials). By way of caution, it must be stressed that neither the commonly found Drav. nor Munda etymologies are up to the present standard of analysis, where both the root and all affixes are explained. This is why most of the subsequent etymologies have to be regarded as preliminary.

Among the c. 380 'foreign' words of the RV, those with certain prefixes are especially apt to be explained from Munda (viz. directly from Austro-Asiatic). Instead of finding Munda prefixes just everywhere in Skt., as was done earlier in this century, we have to be more cautious now: "Owing to the typological change that has taken place in these languages, only some petrified relicts remain" (Kuiper 1991: 39). Typical prefixes in modern Munda are such as *p-*, *k-*, *m-*, *ro-*, *ra-*, *ma-*, *a*, *ə-*, *u-*, *ka-* (Pinnow 1959:10 sqq.; cf. also the plural suffix *-ki* in Kharia, p. 265 §341a, 211 §145c); some of them are indeed attested in the c. 300 'foreign words' of the RV.

Of interest for 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) and which can be compared, in the rest of Austro-Asiatic, to the 'article' of Khasi (masc. *u-*, fem. *ka-*, pl. *ki-*, cf. Pinnow 1959: 14). The following words in the RV are important, even if we cannot yet find etymologies. (In the sequel, Sanskrit suffixes and prefixes are separated from the substrate word in question).

- *ka-*:
- *kakardu* 'wooden stick', 10.102.6 EWA I 286 'unclear';
- *kapard-in* 'with hair knot', Kuiper 1955: 241 sqq.; EWA I 299 'non-IE origin probable'
- *kabandh-in*, *kavandha* 'barrel' Kuiper 1948: 100. EWA I 327 'unclear'
- *kavaṣa* 'straddle-legged', probably Drav., EWA I 327 'unclear'; cf. Kuiper 1948: 130.
- *kākambīra* 'a certain tree', EWA I 334 'unclear'
- *ki-*:
- *kimīd-in* 'a demon', 10.87.24; 7.104.2, 23 (late); EWA I 351 'unclear'; cf. *śimida*, *śimidā* 'a demoness', Kuiper 1955: 182
- *su-kimśu-ka* 'a tree, 'Butea frondosa', CDIAL 3149 and Add., EWA I 348 'not clear'
- *kiyāmbu* 'a water plant' 10.116.13, AV 18.3.6, PS 18.69.4 *k[i]yāmbū*, EWA I 352 'not clear'; with Kuiper 1955: 143 connected with Up. *ambu* 'water', Nur. *abu*, cf. CDIAL 576, *V[i]-yāmbura* 'a demon'; Drav. according to DEDR 187, Kur. *amm* 'water', Malto *amu*, Tam. *am*,

*ām*; Austro-As.: Sant. *um* 'to bathe', Khasi *ūm* 'water', etc. (Berger 1959: 57), more likely because of prefix *ki-*

- *kilāsa* 'spotted, leprous', 5. 53.1, EWA I 354 'unclear'; Kuiper 1955: 170 'derivation unknown'
- *kilbiṣa* 'evil action', 5.34.4, 10.71.10; EWA I 354 'not sufficiently clear', Kuiper 1955: 175 compares TS, VS *kalmāsa* 'spotted' and Epic *kalmāsa*, Pkt. *kamaḍha* (cf. Pinnow 1959: 379 sqq., Kuiper 1991:36 sqq.), Kuiper 1948: 38, 138 on prefixes *kal-*, *kil-*, *kar-*; Sant. *boḍor*, *bode*, *murgu'c* 'dirty', with adaptation *-ṣ-/ḍ-* into Ved. similar to *Vipās-/Vibāl-/\*Vipāž*, cf. Kuiper 1948: 6, 38
- *kīkaṭa* 'a tribe' 3.53.14; EWA 'foreign name of unknown origin'; prefix *kī-* points to Austro-As.; 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'? cf. Shafer 1954: 107, 125
- *kīkasā* (dual) 'vertebra, rib bone' 10.163.2, EWA I 355 'unclear'; "formation like *pi-ppala*, etc. and connected with lex. *kaṣeruka...*" Kuiper 1955: 147
- *kija* 'implement, spur?', 8.66.3; EWA I 355 'loan word possible'; KEWA I 214 and Kuiper 1955: 161, 165: 'doubtful Drav. etym.' (Burrow, BSOAS 12: 373)
- *kīnārā* dual, 'two ploughmen' 10.106.10; EWA I 356 'probably artificial for *kīnāśa*', rather *ś/d/r*, Kuiper 1948: 6, 38, 1991: 30-33, and 1955: 155f., 1991: 26 on suffixes *-āśa/-āra*, (cf. also *-na/-ra* in *rāspina/rāspira*); on *ś* as hyper-Sanskritization for *ṣ/r* cf. *Vipās*; Kuiper 1991: 46 on suffix *-śa*; if *kīnāra-* contains a suffix, then probably no prefix *kī-*.
- *kīnāśa* 'plough man' 4.57.8 (late), AV; Kuiper 1955: 155, 1991: 14, 26, 46 see *kīnāra*; EWA I 356 'unclear'.
- *kīlāla* 'biestings, a sweet drink' 10.91.14; in AV 4.11.10 next to *kīnāśa*; EWA I 358 'unclear'; discussion, above: Khowar *kīlāl*, Nuristani *kīlā* etc., Bur. *kīlay*, Kuiper 1955: 150f., CDIAL 3181.
- *kīsta* 'praiser, poet' 1.127.7, 6.67.10, to be read as [*kisētāsaḥ*] Kuiper 1991:23, 1955:155; the unusual sequence *-is-* (see introd.) points to a loan word (Kuiper 1991:25); EWA I 358 'not clarified'; cf. Kuiper 1991: 20, 23, 25; to be compared with RV *śiṣṭa* 8.53.4 with var. lect. *śiṣṭeṣu*, *śīrṣṭeṣu*, *śīrṣṭrēsa*, Kuiper 1991: 7, 71; this is Sanskritization of *\*k'īsṭeṣu*, Witzel 1999; cf. EWA II 644
- *ku-*:
- *kuṇāru* 'lame in the arm?' 3.30.8; EWA I 362 'unclear'; Kuiper 1948: 53f., 1955: 175, 176 on a Drav. and Munda explanation
- *kupaya* 'shimmering?' 1.140.3, in a 'intentionally ambiguous hymn' (Geldner), EWA I 366; Kuiper 1991: 56 compares *kupaya* with other formations in *-ya*.
- *kumāra* 'boy, young man', 4.15.7 etc. EWA I 368 'not convincingly explained'; cf. CDIAL 3523, 13488; Kuiper 1955: 146f. compares Tel. *koma* 'young', Tam. *kommai*, etc.; note, however, *śi(m)śu-māra* (see below), cf. Munda *māndra*, *mār* 'man'.
- *kurīra* 'women's hair dress', 10.92.8, EWA I 371 'unclear', Kuiper 1955: 152, 1991: 14, 29-31 compares Tam. *koṭu* 'horn, coil of hair', DEDR 2200
- *kuruṅga* 8.4.19, name of a chieftain of the Turvaśa (cf. Kuiper 1991: 6, 17); EWA I 371 'unclear'; however, cf. *kuluṅga* 'antelope', and the frequent totemistic names of the Munda

- *kulāya* 'nest' 6.15.16; EWA I 373 'unclear'; 'foreign', Kuiper 1991:14
- *kuliśa* 'ax' 3.2.1, 1.32.5, EWA I 374 'not securely explained'; Kuiper 1955: 161, 163 compares Tam. *kulir* 'battle ax'; Skt. *kuthara*, *kuddāla* 'hoe', and Sant., Mundari *kutam* 'to beat, hammer', Mundari, Ho *kutasi* 'hammer', Kan. *kuṭṭu* 'to beat, strike, pound'; cf. Kuiper 1991:14; Berger 1963: 419 \**kuḍiśa*, from \**kodeś* in Kharia *khonḍe'j* 'ax', Mundari *konḍe'j* 'smaller kind of wood ax', with prefix *kon-* and Kharia *te'j* 'to break'
- *kuśika* name of a poets' clan, RV 2 etc.; EWA I 379 'not clear'; cf. Kuiper 1991: 7
- *kuṣumbhaka* 'poison gland of an insect' 1.191.15-16; EWA I 381 'unclear'; if not one of the common IA animal names in *-bha* (*śara-bha* etc.), then: \**ku-ṣumb(h)*.

- Double prefixes in *Cər-*.

More important, perhaps, are the so-called 'double prefixes' in Austro-Asiatic, composed of a prefix (e.g. *k-*) followed by a second prefix (mostly *-n-*, see Pinnow 1959: 11). The use of *k-n-* is clear in names of domesticated animals, in Sora *kin-sod* 'dog' : Kharia *solog* 'dog'; Sora *kim-med* 'goat' : Remo *-me'*; *kəm-bon* 'pig' : Juang *bu-tae* (see Pinnow 1959: 168, cf. Jpn. *buta*, Austr. > Sino-Tib. \**mba*(*ɣ*)); Sora *ken-sim* 'chicken' : Mundari *sim*; Remo *gi-rem* 'cat' : Sora *ram-en*. Such double prefixes seem to be rarer in Munda now than in Eastern Austro-Asiatic; cf., nevertheless, Kuiper 1991: 94 on *śar-varī* 'night': *śa-bala* 'variegated'; Kuiper 1948: 38 on the prefixes *kal-*, *kil-*, p. 138 on the 'Proto-Munda prefix *k-*, 1948: 49f. 'prefix *kər-*, *kar-*, and *gala-*'; further cf. above, on *kuliśa*, Kharia *khon-ḍe'j*. Note also the prefixes of Sora *kār-dol* 'being hungry' (D. Stampe, oral communication, June '99) and Skt. *sṛ-kaṇḍu* 'itch', Khasi *śyr-toñ* 'comb', Stieng *sər-luot* 'sweet' (F.B.J. Kuiper, letter 8/24/98; *tur-/tər* also in Ved., Khasi, Senoi, and Austronesian (Kuiper, 1/29/99).

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 s. below). Double prefixes, however, are typical for the R̥gvedic loans, especially formations with consonant-vowel-*r* = *Cər-*, and due to the common Vedic interchange of *r/l*, also *Cəl-*, that were adapted in Vedic with various vowels (*ṛ*, *ur*, etc., see Kuiper 1991: 42 sqq.; cf. below §9, on Nepalese substrate words for similar substitutions). The cases with *Cər* (note also *Cən-*, *Cəm-*), include:

- *karañja* name of a demon, 1.53.8, *karañja-ha* 10.48.4; EWA I 310 'unclear', cf. the tree name *karañja*, DEDR 1507 Kan., Tel. *kānagu*, Konda *karañ maran* etc.; CDIAL 2785.
- *karambha* 'gruel', for a discussion, see below; Kuiper 1991: 51 sqq., 63 compares loan words with *-b-* > *-bh-* (Pkt. *karamba* 'gruel'); -- rather prefix *kar-* and popular etymology with *ambhas-* 'water' RV, or *ambu* 'water' Up., Mbh. Kuiper 1991: 63; cf. also Kurukh, Malto *amm* 'water', but also Tamil *am*, *ām* DEDR 187
- *karkandhu* later, a tree name 'Zizyphus Jujuba', but personal name in RV 1.112.6; EWA I 313 'not clear'; the Drav. word the meaning of *karkandhu*, DEDR 475, 2070, 3293.
- *karkari* 'lute' 2.43.3 (late), probably onomatopoeic, but from which language? Cf. the echo words of IA, Drav., Munda (Sant. *karkur*, *gargaṛ*, *gargor*, etc.); cf. also Kuiper 1948: 55f. on Class. Skt. *karkarī* 'water jar', Epic *gargara/-ī* (based on the body of string instruments),

therefore with CDIAL 4043, CDIAL 2817 *karkata* 'name of various plants, curved roof of a plant', NIA 'cucumber'

- *kārotara* 'sieve, filter' 1.116.7, EWA I 341 'not clear'
- *khargalā* 'owl' 7.104.17 (late), EWA I 448
- *a-kharva* 'mutilated' 7.32.13, EWA I 448; cf. Avest. *kauruua*, then not a loan word; see, however, Kuiper 1955: 176.
- *kalmalik-in* 'shining' 2.33.8; EWA I 325 'unclear'; however, cf. *kalmāṣa* 'spotted', Kuiper 1948: 38; see above on *kilbiṣa*

Further: *kṛ-* [*kər-*] see Kuiper 1991: 40 sqq., 23;

- *kṛkadāśū* 1.29.7, unclear meaning, personal name? cf. *kṛkalāsa* YV?; EWA I 388 'unclear'
- *kṛpīta* 'bush, brush' 10.28.8 EWA I 394 'unclear', cf. also *kṛmuka* 'faggot, wood' KS, CDIAL 3340a; 'unexplained' Kuiper 1955: 160
- *kṛśana* 'pearl' 1.35.1, 10.68.11, 10.144.2 *ūrdhva-* 10.144.2, *kṛśanā-vat* 1.126.4, EWA I 396 'not securely explained'; Kuiper 1955: 152 compares *kṛ-śana* with other words for 'thick, round', such as Skt. lex. *śāni* 'colocynth?'
- *khrgala* meaning unclear: 'staff, crutch, amulet, armor, brush?' 2.39.4; EWA I 494; cf. *khargala* 'owl', above, *Khārgali* PB? -- Kuiper 1948: 49f. 'well-known prefix *kər-*, *kar-*, and *gala-*'

Due to the frequent interchange *k*[*k'*]/*ś*, (see below) the prefix *śar-/śal-* belongs here as well (cf. *kar-kota-ka* RVKh ~ *śar-kota* AV):

- *śaryāta* name of a person, 1.112.17, *śaryāta* 1.51.12, 3.51.7; EWA II 615 compares *śara* 'arrow'
- *śarvarī* 'night' 5.52.3, *api-śarvara* 3.9.7, 8.1.29; EWA II 621 compares \**śarvar*, *śarman* 'protection'; Kuiper 1955:144 u. 1955: 170 compares *śambara*, *karbura*, Kuiper 1991: 30 *śabala* 'variegated' with simple prefix, as compared with prefix + infix ("double prefix") in *śambara* (cf. Kuiper 1948: 136)
- *śalmali* name of a tree, 'Salmalia malabarica', EWA II 622 'probably not to be separated from RV 3.53.22 *śimbala*', CDIAL 12351 (not related Tib.-Burm. \**siñ* 'tree'); Kuiper 1991: 65 on cases with *-lm-* for *-mm-*: 'different dissimilations of \**śamma/śimmal*'.
- *śṛñjaya* a name of a person 6.27.7 (next to Turvaśa), 4.15.4 (next to Daivavant), *sārñjaya* 'descendent of S.' 6.47.25; EWA II 743 supposes connection with *śṛjaya* 'a certain bird' KS, which would agree with the totemistic names in Munda; cf. Kuiper 1991: 7, on non-IA tribal names in RV
- *śṛbinda* name of a demon 8.32.2; EWA II 744 with Kuiper 1991: 40,43 (and earlier) on names such as *Ku-surubinda* TS, PB, ŚB, *Kusur-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 §203; (cf. Kuiper 1939 = 1997: 3 sqq., 1955: 182, Witzel 1999).

In the same way, the prefixes *jar*, *tar*, *nar*, *par*, *bar*, *śar*, *śṛ* = [*jər*, *tər*] etc.: *jarāyu*, *jarūtha* (cf. also Ved. *jar-tila* : *tila*); *taranta*, *tarukṣa*, *trkṣi*, *trtsu*, *nār-miñi*, epithet of a fort; *nār-mara*, probably the area of or the chief of *Ūrjayantī*; *parṇaya*, *parpharī-ka*, *parśāna*; *prakaṅkata* (next to: *kaṅkata*), *prakala*, *parpharvī*, *pramaganda*, *pra-skaṅva*, *pharva-ra*, *phāriva*; *pr̥thi*, *pr̥thī*, *pr-dāku* [*pər-dak-u*] < Munda *da*'k 'water?', *barjaha*; (cf. also *Nār-śada*

RV, *Nār-vidāla*, *Nār-kavinda* PS and \**ku-bind* in: Ved. *ku-sur(u)-binda*, *bainda*, *vi-bhindu*, *vi-bhindu-kī-ya*).

Furthermore, the formations with other vowels that are adaptations of [-ər] as above in [kər]: *tirindi-ra*, *turīpa*, *turphari*, *turva/turvaśa?*, *turvīti*, *tūrṇāśa*, *sūrmī*.

Instead of Cər, the much more common double prefix of Munda, Cən-, Cəm-, is found as well: *kaṅkata*; *śamba*, *śambara* (cf. *śabala!*), *śāmbara*, *śiṃśapā*, *śiṃśumāra*, *śiṅjāra*, *śimbala*, *śimbāta*, *śimyu*. Compare also the prefixes in Cəs-: *puškara*, *puśya*, *rāspina*, *rāspira*.

Kuiper (1991: 39 sqq.) also discusses other prefixes, such as ā-, i-, u-, o-, ni-, bhṛ-, ma-, sa-, śa-, hi-. Among them, the old prefix u- (o-) would be of special interest; however, is found in the RV only in some 5 or 6 cases.

A very clear case is *śa-kunti(-kā)* 'bird' RV, *śa-kunta* 'bird' AV, Ved. *śa-kunta-ka* 'bird', *Śa-kuntalā* 'name of a nymph', Ved. *Kunti* 'a tribal name', next to the Matsya (IA, 'the Fishes'). The Ved. words belong to Kharia *kon-the'd*, Sora *on-tidən*, etc.; Korku *ti-tid* 'a certain bird', Ved. *tit-tir-a* 'partridge', Pinnow 1959: 160 §336; cf. however RV *śa-kuna* 'a (larger) bird', *śa-kuni* 'bird (of omen)' (Kuiper 1991: 44).

Munda-like prefixes are thus very common in the RV. One has to agree with Kuiper 1991: 39f: "According to some scholars Munda was never spoken west of Orissa, Bihar, Madhya Pradesh and eastern Maharashtra... The obvious occurrence of Old Munda names in the Rigveda points to the conclusion that this statement should be revised." If (some of) these words should not go back directly to Proto-Munda, one may think, especially in the case of the untypical formation Cər, of an *unknown* western Austro-Asiatic language, "Para-Munda" (cf. Kuiper 1962: 51, 102).

If this initial interpretation is correct, several far-reaching conclusions can be drawn. The very frequency itself of non-Drav. loan words in the early (as well as in the later) RV is remarkable: it indicates a much stronger *non-Drav. substrate in the Panjab than usually admitted*. Because of the great similarity with Austro-Asiatic formations and because of some already established (Para-)Munda etymologies (such as *śa-kunta* ~ Kharia *kon-the'd*, etc., Pinnow 1959 160 : 336), this substrate is likely to be an early form of western Austro-Asiatic (cf. below, at the end of §4.3.)

Is the Indus language therefore a kind of Proto-Munda? Against this may speak first of all, as Kuiper states (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', *pramaganda* 'name of a chieftain of the Kikāṭa non-Aryans', *śa-kunti* 'bird' < PMunda \**śa-kontid*, *sṛ-binda*, and in post-RV, e.g., *ku-sur(u)-binda*, *bainda*, *vi-bhindu*, *vi-bhindu-kī-ya* 'name of a tribe'. Yet, the substrate may be a very early form of Munda (or another variety of Austro-Asiatic) which still used prefixes actively, just like the eastern Austro-As. languages, e.g. Mon, Khmer, do even today (cf. also below, on Sumerian). 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-*, *śə-r-*, *pə-r-* etc. (Pinnow 1959: 11). If this is correct, then R̥gvedic Proto-Munda represents a *very* old stage of Austro-Asiatic indeed.

#### §4.3. Munda and Para-Munda names

However, direct contact of the non-Indo-Aryan words in the RV with predecessors of present day Munda languages is more problematic. Some of the substrate words may, at least in part, have entered the RV through the *intervention* of the Indus language (*lāṅgala* etc., see below). Yet, there also are a few *direct* correspondences with reconstructed Proto-Munda (*śa-kunta* < \**kon-ti'd*) which indicate the archaic character of the para-Mundic Indus language. For example, the name of *Pramaganda*, the chieftain of the *Kikaṭa* (RV 3.53.14) who lived south of Kurukṣetra (cf. Witzel 1995). Both words are non-Indo-Aryan and they show clear indications of Mundic character: *maganda* can be explained as *ma-gand* with the old, now unproductive Munda prefix *ma-* that indicates possession. The word *gand* may belong to Munda \**gad/gaḍ*, *ga-n-d/gaṇḍ* (Pinnow 1959: 351 §498) that is also seen in *Gaṇḍa-kī*, *Gaṅgā* (Witzel 1999, if not modeled after the tribal names *Aṅga*, *Vaṅga*, see below), W. Nepali *gāḍ* (as 'suffix' of river names, Witzel 1993) and apparently also in *Ma-gadha* (with Sanskritization > *dh*). Kuiper 1991: 43f. (8, 21, 96, also 1955) has explained the prefix *pra-* [*pər*] (cf. prefixes such as *kər-/śər-*) from Munda, which looks perfectly Indo-Aryan but in this case certainly is 'foreign'. The tribe of chief *Pra-maganda* (*pər* 'son of'? Kuiper 1991: 43), the *Kikaṭa*, has either the typical 'tribal' suffix *-ṭa* (see below) or the old Austro-As. plural prefix *ki-*, or maybe both. Cf. further the prefix *kī-/ki-* in: *kīnāśa/kīnāra* 'plough man', *Kimīdin* 'a class of demons', *kīkasa* 'vertebra, breast bone', *kīlāla* 'biestings', *kiyāmbu* 'a water plant', all of which may be compared with the Munda prefix *k-* for designation of persons (and the plural prefix *ki-* of Khasi; note that in RV, *k-* also applies to items merely *connected* with humans and animals).

Further RV substrate names of persons, tribes and rivers include some exactly from the areas where Indus people are to be expected: in their late/post- Indus new settlement area (J. Shaffer 1995: 139) in the eastern Panjab, in Haryana (Kurukṣetra), and especially east of there, well into the Gangetic plains. Even during the middle/late Vedic period, the local rivers of E. Panjab are still designated by non-Indo-Aryan names: the famous Bharata chieftain Sudās crosses (RV 3.33) the Śutudrī and Vipās and settles on the Sarasvatī. They are not explainable from IA:

Śutudrī (Satej) < \**śə-tu-da*? from Munda \**tu* 'float, drift', Kharia *thu'da* < \**tu-da* (*da* 'water'), Khasi *pər-tīu* 'outflow', (note the later popular etymology *Śatadru* 'running with a hundred streams'); for the Ved. substitution of 'k' by *r* cf. \**kul-do* 'tiger' > *kulitar-a*? and \**ganda* > *gandhar-i*? -- Vipās < \**vipāž*/\**vibāl* (cf. *Vibālī* RV 4.30.11-12), and note that the Sarasvatī still has a similar name, *Vaiśambhalyā* (with many variants, always a sign of foreign origin: TB 2.5.8.6, *-bhalyā*, *-pālyā*, *-balyā* ĀpŚS 4.14.4, *-bhalyā* Bhāradvāja Śikṣā; cf. also RV *viśpalā*?) < \**viśambaž*, \**viśambāl*, probably with the prefix *śam/k'am-* (as in *Śambara*, *Kam-boja*) from \*(*vi*)-*śam-bāž* (note the popular etymology from *vi-śambala* 'having widespread blankets').

The land of *Tūrghna* (TĀ), north of this region, has no Indo-Aryan etymology either (see EWA), and *Khāṇḍava* (TĀ) with its suspicious cluster *-ṇḍ-* (K. Hoffmann 1941), south of Kurukṣetra, is inhabited by the *Kikaṭa* under their chieftain *Pra-maganda*. Note also, in the same area (Kurukṣetra), the appearance of Pinnow's *u*-suffixes in 'foreign words', e.g. *Khāṇḍava*, *Kārapacava*, *Naitandhava* (Pinnow 1953-4).

The Greater Panjab names of *Gandhāra*, *Kubhā*, *Krumu*, *Kamboja* may be added. -- *Gandhāri* RV, *Gāndhāra* Br., OP *Ga<sup>n</sup>dāra*, Herodotos *Gandārioi*, EWA I 462, cf. Munda \**ga(n)d* 'river', the river names of the Gangetic plains, *Gaṇḍakī* and *Gaṅgā*, the *Gandhina* people on its upper course, and Nep. -*gād* in river names. *Gandhāra* is formed with the common suffix -*āra*, -*āla* (Witzel 1993, 1999); -- *Kubhā*, cf. Skt. *kubja* 'bent', Kuiper 1948: 42f., Sant. *kubja* which belongs to Munda *ḍui'j*, *kəb-ḍuj* etc. (Pinnow 1959: 21, 91: §108, 249 §286 Kharia *ḍui'j* 'bend', Santali *kəbḍuj* 'ugly', *kəbḍuju'd* 'crooked', p. 435e Santali *kəbnūj* 'bent', etc.) -- *Krumu* from Munda \**kə-rum* 'luke warm'?? cf. Kharia *rum* 'to burn', Sant. *ur-gum* 'luke warm', Mon *uj-ruñ* 'humid, warm'; --The *Kamboja* (AV, PS < *ka-mboj*??) settled in S.E. Afghanistan (Kandahar); cf. OP *Ka<sup>m</sup>bujīya* (or *Kambaujīya*?) 'Cambyses'; however, their name is transmitted as *Ambautai* by Ptolemy (Geography 6.18.3), without the typical prefix; cf. also *Bulitai*). This change in the first syllable is typical for Munda names (see below *Aṅga* : *Vaṅga*, *Kaliṅga* : *Teliṅga*; *Kulūṭa* : *Ulūṭa*, etc.) - Mundas that far west cannot be excluded (Kuiper 1991: 39).

It may be asked, how far Austro-Asiatic speakers extended westwards during and before the RV period. Until now, the *present* distribution of the Munda languages has led to rather far-going conclusions, for example by Burrow (1958, cf. Southworth 1979: 200). Starting from the modern settlement areas of the Mundas in Eastern India (Bihar, Orissa, W. Bengal) and on the River Tapti (in northwestern Maharashtra and Madhya Pradesh) he regarded it as impossible that the Munda could ever have settled in the Panjab. Kuiper, however, has been of a different opinion (1955: 140, 1991: 39, see also 1948: 8, cf. Witzel 1980, 1993 on the substrate in Nepal, and 1999 for the Panjab area). The cases discussed above indicate a strong Austro-Asiatic substrate in the Panjab, and there are some hints which point to Munda influence in the Himalayas (Konow 1905, Witzel 1993, see below) and even in E. Afghanistan (*Šambara*, *Kamboja*).

An important result therefore is, *that the language of the Indus people*, at least those in the Panjab, must have been *Para-Munda or a western form of Austro-Asiatic*.

In view of the recent comparison by the late I. M. Diakonoff of Munda and Sumerian (*MT* III, 54-62, but note the criticism by P. Bengtson *MT* III 72 sq., and cf. still differently, A. R. Bomhard, *MT* III 75 sqq.) this characterization of the pre-IA Panjab acquires special importance (cf. already Przyludski 1929: 145-149). If Munda were indeed related to Sumerian, names such as *Ki-kaṭa*, *Ki-nāša*, *Ki-rāta* may no longer surprise, cf. Sum. *ki* 'country'. To follow up, the role of compound nouns in Sumerian versus old 'prefixes' in Munda would need further investigation. Consider, as a very vague possibility, Para-Munda *pər-* (*pra-magandha*, *pra-skaṅva*) and Sum. *bala* 'term of office, reign, dynasty'. In this regard, it should be noted that Sumerian has 'implosive' (unreleased) consonants, just as Munda, Khasi, Khmer, the Himalayan language Kanauri, the Kathmandu Valley substrate, and Sindhi, all of which may point to a S./S.E. Asian areal feature (For 'implosive', unreleased stops including labials, in Munda and Eastern Austro-Asiatic, see Pinnow 1959: 313 sqq, 316, cf. Zide 1969, 416 sq. The final consonants *j*, *ḍ*, *d*, *g*, which are also called 'checked consonants', are preglottalized *and* unreleased in Munda.)

If Diakonoff's proposal was borne out, the Rgvedic Para-Munda substrate in the Panjab of c. 1500 BCE would represent an early link to Sumerian. Notably, Sumerologists,

though without any firm reasons going beyond some vague mythological allusion to more eastern territories (Dilmun, etc.), think that the Sumerians immigrated from the east, from the Indus area.

If a relationship with Munda could not be confirmed by obvious etymologies, a minimal position would be to define the c. 300 non-Dravidian loan words 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-Vipās or simply, the Harappan language.

Finally, in reviewing the evidence of the Ṛgvedic Para-Munda, it should 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), though much less known, and that both 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 or even may have been due to a north Indian substrate such as Masica's "Language X".

#### §4.4. Other substrates

If the Indus language is a kind of Para-Munda, a 'western' Munda, it cannot, however, be excluded that one or more *unknown* languages are involved (cf. Zide and Zide 1973:15) in the Ṛgvedic substrate. From the older RV onwards, we find a number of words that cannot be determined as Para-Munda. Examples include the words with geminates (see below) e.g. *pippala* RV 5.54.12 and an undetermined number of the c. 300 'foreign words.' Some of them can be traced as being loan words from more distant eastern (Austro-As.) or western (Near Eastern) languages; the path the loans have taken is clear (see below) in the case of RV *lāṅgala* <-- Indus \**langal* (<-- Sumer. *nig-gala<sub>x</sub>+l* or *nig-gál* 'sickle?', see §5.3), <-- > PMunda \**ñan-keḷ*, Austric (Makassar) *naṅkala* (see §5.3); Ved. *vrihi* < Indus \**vrijhi* <-- PMunda (c. 1500 BCE) \**ərig*/ Tib./Malay (')*bras* <-- S.E. As. \*\**əβərij* (?); Ved. *mayūra* 'peacock' <-- N. Indus \**mayur* <-- PMunda *mara'k* 'crier' <-- Austr. (Malay) *merak* --> Sino-Tib. \**raka* 'cock'. Note also the various substrates in Burushaski, Nahali and "Dhimal" (Kiranti languages in E. Nepal) discussed in *MT* II, III and by Kuiper 1962: 14 sqq., 40, 42, 46f, 50f., Berger 1959: 79; and cf. those of the Kathmandu Valley and Tharu (s. below).

In short, the Panjab is an area of a Pre-Ṛgvedic, largely Para-Munda substrate that apparently overlays a still older local level which may be identical with Masica's "language X" found in the Gangetic plains (Hindi). In general, the vocabulary of Para-Munda and "language X" words is limited to local flora and fauna, agriculture and artisans, to terms of toilette, clothing and household; dancing and music are particularly prominent, and there are some items of religion and beliefs as well (Kuiper 1955, 1991). Since no traces of the supposedly Dravidian "Trader's Language" of the Indus civilization (Parpola 1994) are visible in the RV, the people who spoke this language must either have disappeared without a trace (cf. below on Meluhḥa) or, more likely, the language of the Panjab was Para-Munda already during the Indus period (2600-1900 BCE).

The large number of agricultural words alone (Kuiper 1955) that have no Dravidian explanation indicates 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 Kīkātā?" (*kīm te kṛṇvanti Kīkāteṣu gāvah*, RV 3.53.14).

As we can no longer reckon with Dravidian influence on the early RV (see immediately below), this means that the language of the pre-Ṛgvedic Indus civilization, at least in the Panjab, was of (Para-)Austro-Asiatic nature.

This means that *all proposals for a decipherment of the Indus script must start with the c. 300 (Para-)Austro-Asiatic loan words in the RV* and by comparing other Munda and Austro-Asiatic words. (For the Indus script see Fairservis 1992: 14, Parpola 1994: 137 sqq., Possehl 1996b). The decipherment has been tried for the past 35 years or so mainly on the basis of Dravidian. Yet, few Indus inscriptions have been "read" even after all these years of concerted, computer-aided attempts, and not yet in a fashion that can be verified independently (cf. a summary of criticism by Zvelebil 1990). Perhaps that is not even attainable, due to the brief nature of the inscriptions (7 signs on average and hardly more than 20). Yet, Kuiper's '300 words' could become the Rosetta stone of the Indus script.

Further, investigations of the South Asiatic linguistic area (*Sprachbund*) must be reformulated accordingly, for example the question of the retroflex sounds, see Tikkanen 1988, and cf. Zvelebil 1990: 71 on the distinction between true retroflex sounds (domals, 'cerebrals') and cacuminals. In the RV they cannot go back either to Proto-Drav. influence, as usually assumed, because they are already found in the older part of RV (books 4,5,6) where no Drav. loans are present; they also cannot go back to Proto-Munda influences because Munda originally had no retroflexes (Pinnow 1959, except for *ḍ*, an isolate in the reconstructed consonant system, see Zide 1969: 414). The clear increase of the retroflexes in RV books 1, and especially in 10 is remarkable. In the older RV one can only detect very few cases of not internally conditioned, original and clearly non-IA retroflexes: RV 6: *kevaṭa* 'hole'; *reṇu-kakāṭa*; *rāṇḍya*, *śāṇḍa*, (*hiranya-*)*piṇḍa* (late hymn), RV 4, 5: *kṛiḷ-*; RV 2: *śaṇḍika*, *mārtāṇḍa*, *pipiḷe* ( from *piḍ*, < IIr \**piḷd*); cf. also *jaṭhāra* in RV 1,2,3,5,6,9,10. None of these old words is Dravidian (see below). In short, the people of the (northern) Indus civilization must have spoken with retroflexes.

Almost the same situation exists with regard to another item of suspected substrate influence, the innovation in Vedic of the grammatical category of absolutes (not found in Old Iranian!, see below). They occur in RV 4 with 1, RV 6 with 1, RV 2 with 4 cases (a relatively high number in this short book!); equally, in RV 3 with only 1, RV 7 with 4, RV 8 (Kāṇva section) with 0, RV 8 (Āṅgīrasa section) with 2, RV 9 with 4; even RV 1 (Kāṇva section) only with 5. - Really innovating are only the late books RV 1 (Āṅg.) with 34, and RV 10 with 60 forms.

#### §4.5. Dravidian in the Middle and Late Ṛgveda

As has been repeatedly mentioned, there are no traces of Dravidian language in the Panjab until c. 1500 BCE, not even of the supposedly Dravidian speaking traders and rulers of the Indus civilization; however, 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). These include personal and tribal names, as well as cultural terms.

For comparisons, 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.

To begin with, many words that have been regarded as Drav., are now explained as coming from Munda or another substrate language, for example, *mayūra* 'peacock' whose correspondence in Munda \**ma-ra*' still has an appellative meaning, 'crier'; (PMunda \**ra*'k 'to cry,' Pinnow 1959: 76 §57). However, this is not so for the Drav. designation, where 'peacock feather' is reconstructed at a level earlier than 'peacock' itself. Indeed, 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 as early Dravidian loans in Vedic.

Even if one would regard *all* of them, for argument's sake, as Dravidian, only *kulāya* 'nest' 6.15.16, *karambha* 'gruel' 6.56.1, 6.57.2, *ukha-cchid* 'lame in the hip' 4.19.9 occur in early R̥gvedic. These words can, however, no longer be explained as Dravidian:

- *karambha* 'gruel' CDIAL 14358, no longer in DEDR; Kuiper 1955: 151 Drav. etym. as 'doubtful', EWA I 310 'unclear'; Kuiper 1991: 51 sqq. compares loan words with *-b-* > *-bh-* (Pkt. *karamba* 'gruel').
- *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 word formation? and Drav. \**-ḍ-* > Ved. *-l-?*; EWA I 373 'not clear', comparing N.Pers *kunām*, East Baluchi *kuḍām* < *kudāman*, with the same problems; 'foreign word', Kuiper 1991: 14.
- *ukha* 'pan, hip' in *ukha-chid* 'breaking the hip, lame' 4.19.9, cf. MS 4, p. 4.9 *ukhā* (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; EWA I 210 compares Latin *auxilla* 'small pot', Lat. *aulla* 'pot' (Pokorny 88), but declares 'not sufficiently explained'. As RV 4.19 is not seen as a late hymn, this might be the oldest Drav. loan in Vedic (RV I).

Only cases in the middle and late RV remain: In the early RV (2,4,5,6) possible Drav. words are found only in some additional, late hymns (insertion after the initial collection of the RV, c. 1200 BCE, cf. Witzel 1995):

- *-phala* 4.57.6 'fruit' DEDR 4004, Tam. *paḷu* 'to ripen', *paḷam* 'ripe fruit', etc., see Zvelebil 1990: 78 with literature, Parpola 1994: 168; CDIAL 9051, 9057; EWA II 201 doubts Drav. origin, and derives it from IA *phal/r* 'to coagulate, condense', but finds 'origin of IA \**phal/p<sup>h</sup>ar* not explained'; that means, a Middle RV loan from Drav. remains possible, or from Munda: Sant. *piṭiri* 'swelling of glands as in mumps', Sora *pēl* 'to swell, grow in bulk (seeds)'; cf. Kuiper 1955: 144, 158, 183 (cf. also, 1948: 163, Kharia *poṭki* 'to sprout', *potri* 'pregnant' ); Pinnow 1959:173, §378.
- *phāla* 'plough share' 4.57.8, Turner, CDIAL 9072, connects *phalati*, Iran. \**spāra*, and thinks that it has been influenced later on by Drav./Munda; not in DED(R); EWA compares N.Pers. *supār*, Pashto *spāra*, Iškašmi *uspir* < \**spa/ārya*?
- *-piṇḍa* 6.47.23 'ball, dumpling'; the many divergent NIA forms speak for a loan word, see CDIAL 8168 and add.; Drav., Burrow 1946: 23; Munda, Kuiper 1948: 142, 162, cf. 1991: 14; DEDR 4162 Tam. *piṇṭi*, Konda *piṇḍi* etc. 'flour'? - EWA II 128 'unexplained'; cf. also K. Hoffmann 1941: 380 sqq. and perhaps Armenian *pind* 'compact, firm' < Iran. (< Ved.?)

In middle RV (3,7,8):

- *kuṅāru* 3.30.8 'lame in the arm?', or name of a person, see EWA I 362 'unclear'; perhaps connected with Epic *kuṅi* 'lame', *kuṅṭa* 'defective'; however, compare Drav.: Kan. *kuṅṭa* 'cripple', Mal. *kuṅṭan* 'cripple', etc., CDIAL 3259-60, DEDR 1688
- *mayūra* 3.45.1 DEDR 4642, 'peacock' PS, *mayūri* 'pea hen' RV 1.191.14, *mayūra-roman* 'having hair like peacock ('feathers)' RV 3.45.1, *mayūra-śepya* 'a peacock-like tail' RV 8.1.25; generally regarded as Drav.: DEDR 4642 Tam. *mañṅai*, *mayil*; northern Kasaba dialect of Irula *muyiru*, Tulu *mairu*, Konda *mriḷu*, *miril*, (\**mayil/mayir*, see Zvelebil 1990: 77, with discussion and lit.). However, originally from Munda: PMunda \**mara*' 'crier', Kharia *mara*', Santali, Mundari, Ho *mara*', Kurku *mara*, Sora *mārān* 'peacock, Pavo cristatus', see Pinnow 1959: 205 §90; cf. also Skt. *marūka* (lex.) 'peacock, deer, frog, Curcuma Zerumbet', and Khotanese Saka *murāsa* 'peacock' (EWA II 317, KEWA II 587, CDIAL 9865, add. 9865, DEDR 4642, Bagchi 1929: 131, Southworth 1979: 191 sqq., 200, cf. Zvelebil 1990: 77, Hock 1975: 86). The rare tribal name *Mara-ṭa* PS 5.2.1, 12.2.1 (Witzel 1999) belongs here; the *Marāṭa* probably lived south of the Ganges and north of the Vindhya. The above may indicate that the Dravida entered into contact with some groups of Munda speakers fairly early (before the Middle RV); however, just as in the Vedic case, one or two intervening language(s) (\**mayil* / \**mayur*) must have delivered the word to Drav. and Vedic, for example the "Language X" or rather a Northern and Southern Indus language; in the south, this must have occurred before Sindh was practically deserted in the post-Indus phase (Allchin 1995: 31 sqq.). The Ved. form *mayūra* may have been influenced by *māyu* 'bleating'.
  - *phala* 3.45.4 see above
  - *kāṇa* 7.50.1 'one-eyed' EWA I 336 'unclear'; cf. Avest. *karəna* 'deaf' : *karəna* 'ear' and cf. DEDR 1159 Tam. *kaṇ* 'eye' and 1443 *kāṇ* 'to see', both now without reference to Skt.; Zvelebil 1990: 79 compares DEDR 1159 and finds, 'rather speculative', the Drav. negative suffix *-a/-ā*; cf. Kuiper 1991: 79. --However, cf. Burushaski *śon*, *śōn* 'blind' (see above, with northwestern interchange of Ved. *ś/k*, Witzel 1999); note also that *kāṇa* is found as hapax RV 10.155.1 next to 'mountain', a 'foreign' name and an onomatopoeic: *girim gaccha* 'go to the mountain!', *Śirimbiṭha*, *budbud-* 'making bubbling sounds' (cf. Sant. *buḍu'c buḍu'c* 'to bubble up').
  - *kulpha* 7.50.2 'ankle', CDIAL 4216, from Drav.; cf. DEDR 1829 *kuḷampu* 'hoof?'; EWA I 376 'completely unclear', Kuiper 1955: 148 loan word because of AV *gulpha* and points (1991: 35) to variant forms in Ved. (*gulpha*) and MIA (*gopphaka*, *guppha*, *gompha*).
  - *daṇḍa* 7.33.6 (late) 'stick', DEDR 3048 Mal. *taṅṭa* 'forearm, arm', Tel. *daṇḍa*, etc., cf. DEDR 3051, CDIAL 6128; Munda, Kuiper 1948: 76: Sant. *ḍaṅṭa* 'thick stick, club', *ḍa(ṅ)tiṭit* 'stem (of mushrooms)', *ḍaṅḍi* 'stick, staff, stalk', cf. Mundari *ḍāṅḍi* 'small stick'; EWA I 691 'not explained'
  - *kuṅḍa-* 'vessel' 8.17.13 can be compared with Avest. *kunda/-ī*, *kundižā*, the name of demons; Dravid., DEDR 1669 Tam. *kuṅṭam* 'deepness, pond', Tel. *kuṅṭa*, *kuṅḍu*, Kur. *xoṅḍxā* etc., DEDR 2082; Kuiper 1948: 76 Drav., 1991:14 'foreign'; CDIAL 3265; EWA I 363 points to the difference in meaning between Drav. and Ved. and concludes 'unclear, perhaps loan word'
  - *mayūra* 8.1.25, see above

- *nala* 8.1.33 'reed', *naḍa/nala/nada*, EWA II 7 from Iir. \**nada* (Nuristani *nō* < \**nada*, Parthian *nad* 'flute', N.Pers. *nāy* 'flute') < IE \**nedo* (Hitt. *nata* 'reed', Armenian *net*), however without actual explanation of the variation \**d* > *ḍ* (cf. Mayrhofer 1968); DEDR 3610 compares, strangely, Tam. *nal* 'good' with the Skt. name *Nala*, idem Zvelebil 1990: 82; however, *Nala* is found in Vedic, ŚB 2.3.2.1-2 *Naḍa Naiṣidha*, and in Mbh. *Nala Naiṣadha* as king of the (probable) Munda tribe of the *Niṣidha/Niṣadha* = Ved. *Niṣāda* (MS, VS, see below); cf. Kuiper 1991: 33 on *ḍ/d*, and p. 19 *nālī* 10.135.7 'flute, pipe' (cf. 1948: 82).
- *kāṇuka* 8.77.4; (poet: Kurusuti Kāṇva) next to *saras* 'pond'; unclear in meaning and etym., EWA I 336; Kuiper 1991 as foreign.

In late RV (1, 10):

- *ulūkhala* 1.28 'mortar' DEDR 672 Tam. *ulukai*, Kan. *olake*, Koḍagu *olake*, and Kota. *olka*, *olkal kal* '(stone) mortar', Malto *lorā* 'stone to grind spices' (S. Palaninappan, by letter); EWA I 231 'problematic'; cf. Zvelebil 1990: 79 with lit., Kuiper 1991: 14, 41 'still unexplained', compares loan words with prefix *u-*; note Sumerian *ur5* 'millstone', Proto-South Drav. \**ur-al* 'mortar' (Blažek and Boisson 1992: 24); is there a connection with *khala* 'threshing floor' RV 10.48.7?
- *vriś* 1.144.5 'finger', DEDR 5409 Tam. *viral*, Go. *wirinj*, now without reference to Skt. *vriś*; EWA II 597 from IA \**vres* 'to bend', Avest. *uruuvaēs* 'to bend, curve'
- *bila* 1.11.5, 1.32.11 'hole, cave' CDIAL 9245 'Dravid.'; DED 4459 = DEDR 5432 now without reference to Skt., cf. also DEDR 4194; Kuiper 1991:14 'foreign', EWA II 225 'not clear'
- *a-phalā* 10.71.5 'without fruit', see above;
- *phal-inī* 10.97.15 'having fruits', see above;
- *mayūra* 1.191.14, see above;
- *piṇḍa* 1.162.19, see above
- *kūṭa* 10.102.4 'hammer' DEDR 1651, 1655, 1883, app. 29; previously explained by Burrow as Drav., later explained by him as IE (German *hau-en*), but see EWA I 384 'unclear'
- *phāla* 10.117.7 'plough share', see above
- *phala* 10.146.5 'fruit', see above
- *kāṇa* 10.155.1, see above
- *kaṭu(ka)* 10.85.34 'pungent'; CDIAL compares *khaṭṭa* 'pungent'; EWA I 290 Lithuanian *kartūs* 'bitter'? or DEDR 1135 Tam. *kaṭu* 'to pain; pungent; cruel, harsh, bitterness', Kurukh *xarxa* 'bitter', Malto *qarqe* 'bitter', Brahui *xarēn* 'bitter' etc.
- *bala(?)* RV 1,3,5,6,7,9,10 'strength, force'; EWA compares Latin *de-bilis* etc., IE \**belo-*, otherwise not found in Iir. (perhaps in Osset./Sarmatian); see, however, Kuiper 1990: 90, on the rare IE (initial) *b-*, and on the impossibility of an IE etymology; cf. CDIAL 9161; now, against Drav. origin Burrow, see EWA II 215; cf., nevertheless, DEDR 5276 Tam. *val* 'strong', Kurukh *balē* 'with the help of', Brahui *balun* 'big'.

The same is the case with some words that have later on been added and discussed (Sanskrit Index of the DEDR, p. 759-763) and elsewhere. Most of them are too late to be of interest here. In DEDR we find:

Early RV: *phalgu* 'minute, weak' 4.5.14, *kalaśa* 'vessel' 4.27.5, 6.69.2, 3.32.15, 7.69.6; and later: *taḍit* 'flash' 2.23.9 (late), 1.94.7 *phāla* 'plough share' 4.57.8 (late); -- middle RV: *ukhā* 3.53 'pan, hip' (late), *kavaṣa* 'straddle legged', a personal name 7.18.12, *kāla* 'slope, bank' 8.47.11. -- late RV: *ukhā* 'pan, hip' 1.162.13,15; *khala* 'threshing floor' 10.48.7.

Of these, only *phalgu* 'minute weak' (RV 4) remains as a possible early loan into IA, if it indeed belongs to DEDR 4562, Tam. *pollu* 'empty husk of grain'; EWA II 203 has an IE etymology. Again, all other words regarded as Dravidian appear only in the *middle* and especially in the in *later* RV.

Southworth (1990, 1995) adds the following examples of early contact between Drav. and Indo-Ar., however, without ordering the texts historically.

- *car-*, *carati* RV : Tamil *cel* 'to go, flow, pass, be suitable' (already Perunkunrūr Kilār, c. 160-200 CE); DEDR 2781 "probably from IA", CDIAL 4715; IA, without problems from IE \**kwel(h)*; perhaps accidental agreement with Drav. *cel*.
- *māyā* 'confusion, wonderment, awe' RV (found in all of RV, just as *māy-in*, *mayā-vat*, *mayā-vin*), = Avest. *māiiā* 'awful power' :: Tam. *maya-* 'mistake, misunderstand'; *mayakku-* 'bewilder, confuse, intoxicate, alcohol' etc.; DEDR 4706, without comparison with Skt.; the Skt. and Drav. meanings do not agree; also, as attested that early in the RV and Iran., Drav. origin (only Middle-RV Drav. influence!) is unlikely, -- unless it would have taken place in Iran (Southworth 1979: 196f.: "high degree of contact ... at the earliest period for which we have records and possibly before"); however, see below, on *tanū*.
- Southworth 1979: 203, 228 f., 1990: 222-3, 1995 reconstructs as further indication of early contact between Drav. and Indo-Ar. in Iran, a word \**tanu* 'self', Tamil *tān/tan* 'oneself', *tanū* RV 'body, self/oneself', for this meaning see now J. R. Gardner, U. of Iowa Ph.D. thesis, 1998. The variation in vowel length in the Drav. pronoun (Tam. *tān/tan* 'oneself') is old (Krishnamurti 1968). However, next to the RV instances, there is Avest. *tanū* 'body, self', OP *tanū* 'body', however, they all have no clear IE etymology. Pokorny 1959: 1065, 1069 derives them from IE \**ten* 'to stretch', in other IE languages the meaning mostly is 'thin'; EWA II 622 connects *tan-ū* '\*Ausdehnung, ausgespannte Hülle' with *tan*. The comparison of the Iir. and Drav. words would presuppose a very close relationship between Drav. and (pre-)Indo-Ar. tribes, as pronouns are not taken over easily. Such early Drav.-IA relationships are not found otherwise: there are no early loans in designations of material culture, e.g. pastoralist terms in Vedic/Drav.: horse: *aśva* : *ivulī*, *kutira*, cow: *gau-* : *ā(ū)*, sheep: *avi* : *(y)āṭu*, *korī*, goat : *aja* : *(y)āṭu*, *korī*, dog: *śvan* : *nāy*, *nāi*. This would rather point *against* a neighborly relationship of both languages in any pre-South Asian context.
- *garda-bha* 'donkey' RV late, only 1.23.5, appendix hymn 3.53.23 next to *rāsa-bha* 'donkey!', RV Vālahilya 8.56.3 :: Tam. *kaḷutai*, Gondi *gārḍi*, etc., to which DEDR 1364 compares Skt. *gardabha*; CDIAL 4054; EWA I 473 cf. *gard* 'to cry shout', not from Drav.
- *piśāca*, *piśāci* AV, *piśāci-* 'demon' RV late: 1.133.5 :: Tam. *pēy-* 'devil, goblin, madness' DEDR 4468, without comparison with Skt., and without suffixing *-śāci-*, only: *pēytti*, *pēycci*, *pēcci* 'demoness'. -- Ved. *piś-* may derive from Tam. *pēy* etc. if, with Zvelebil 1970: 111, Drav. *-c-* > *s* > *y*.

- *śava* (not in RV, diff. Southworth 1979: 197), only AVP : Tam. *cā*- 'to die' (Kural), Ko. *ca-v*- 'corpse' DEDR 2426 compares Skt. *śava*; EWA II derives *śava* from *śav* 'to swell' AVP; CDIAL 12356 not from Drav. As the word is early in Drav., perhaps accidental look-alike.
- *paṭhati* 'to recite' RVKh., TĀ, Up. : Tam. *pāṭu* 'sing, chant', *pāṭṭu* 'song', attested already in Perunkunrūr Kilār, DEDR 4065 without reference to IA; EWA II 69; CDIAL 7712 < \**prṭhati*; Drav. <-- Indo-Ar., Burrow-Emeneau 1962: 46, no. 242. Rather to be derived from MIA pupil's slang Ved. *prath* 'to spread out (a text, in recitation)'; compare the frequent loan words in the context of Vedic teaching and learning: *maṇḍala*, *kaṇḍa*, *kāṇḍa*, *prapāṭhaka*, *paṭala*, *daṇḍa*, MIA: *orimikā* 'a section of KS' etc.
- *nagara* 'town' TĀ, but cf. already *nagar-in* JB :: Tam. *nakar* 'house abode, town, city'; cf. EWA II 5, CDIAL 6924; DEDR 3568 IA --> Tam. *nakar* 'house, town, etc.' But why *nakar* from Skt.? There is no IA etymon, nor is there one in Drav. and Munda. Drav. for settlements: DEDR 3568 *nakar* 'house, town', 1655 *kuṭi* 'home', 3868 *paṭṭi* 'cow stall, village', 5393 *viṭu(ti)* 'temporal residence', 2007 *cēri* 'street, village', 752 *ūr* 'village', 4362 *pūṅṭi* 'town, village', 4047 *pākkam* 'seaside village', 4646 *maṭappam* 'agricultural town', 807 *eyil* 'fortress'; 4064 *pāṭi* 'town', 4112 *pāli* 'temple, town', 4555 Kan. *poḷal* 'town', 5549 *vai*, 3911 *pati*, 2814 *cēr*; 3638 *nāṭu* 'open country' (opp. *nakaram*); -- cf. also Skt. *haṭṭa* 'market'~ Santali, Mundari, Ho *hatu*, Korwa *watu* < PMunda \**watu* Pinnow 1959: 79 §69.-- In short, the word may be a loan from the southern Indus language or one from the Malwa area.

Thus, the words added by Southworth are post-Ṛgvedic (*śava*, *paṭhati*, *nagara*), or they are attested in relatively late RV sections (*gardabha*, *piśāci*), or they are of dubious nature (*car*, *māyā*, *tanū*). Therefore, it is not possible to suppose, with Southworth, an *early* close contact, *even in Iran*, and on all levels of society, of Dravidas and Indo-Aryans. Rather, one has to agree with Kuiper, who stresses the very hesitant acceptance of non-Indo-Aryan words and forms in the high level, poetic language of the RV. The words collected by Southworth in his second list can have been taken over into Drav. at *any* time after the RV, e.g. *accu* 'axle' < *akṣa* RV.

Furthermore, most of the c. 800 words in the list provided by DEDR, p. 759-764 are attested only in the Epics or in class. Skt. Of the c. 61 words listed in the appendix of DEDR which are supposed to come from Indo-Aryan, only a few can be regarded as (possible) early loans; they all should be checked in early Tamil before something that even approaches a final decision can be made.

Finally, among the words in Zvelebil's recent list (1990: 77-82) of 22 "early" Drav. loans into Skt., most have already been discussed above; yet, none of them nor the ones newly mentioned are Ṛgvedic: 8. *bilva* 'Aegle marmelos, Bel tree' AV, 10. *kunapa* 'corpse' AV, 11. *kurkura* 'dog' AV, 12. *arka* 'Calatropis gigantea', ŚB, 12a. *candana* 'sandal wood, paste' Nirukta, 13. *kavaca* 'armor' PS, ŚB, *kavacin* AV, 13a. *jaṭā* 'matted hair' GS, 13b. *mālā* 'flower necklace', GS, *mālya* RVKh, 13c. *eḍa* 'sheep' KŚS, *eḍaka* JB, *aiḍaka* ŚB. The rest of the words are only post-Vedic.

Zvelebil's summary is: "as Emeneau (1971) 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). According to what has been said above, this has to be modified drastically: Ṛgvedic loans from Drav. are

visible, but they also are now datable *only* to middle and late Ṛgvedic (in the Greater Panjab), and they can both be localized *and* dated for the Post-Ṛgvedic texts (Witzel 1987, 1989).

Of all the words mentioned so far that have been regarded as Drav., only the following few are possible, though not uncontroversial, for the early RV :

*ukha[-chid]* 'hip[-breaking]' 4.19.9; *phalgu* 'minute' 4.5.14, *āṇi* 'lynch pin' 5.43.8 (whose ultimate source is unclear, and, very tentatively, *bala* 'force' 5.57.6, 5.30.9, probably from IE, cf. Latin *de-bilis*).

Whether this is enough to ensure the presence of (even a small number of) speakers of Dravidian in the Panjab during early RV times may remain in the balance. These few village type words would constitute a strange legacy of the c. 700 years of the great Indus civilization, had it been speaking Dravidian. From the middle RV, however, come: *kavaṣa* 'straddle legged', (a personal name) 7.18.12, *kāla* 'slope, bank' 8.47.11 and perhaps also *kunḍa* 'vessel' 8.17.13. Burrow (1955, 1958) regards the Drav. element in Vedic as having come from Northern Drav., but cf. Zvelebil 1990: 46.

If the middle and late RV words mentioned above are accepted as Drav. and even if some of the words *excluded* above for the early RV should be accepted, this would not change the general picture: There is *very* little Dravidian, but there are about 300 words of the Indus substrate.

For it *cannot* be said, conversely, that there were, during the older and middle RV, clear indications (or: "a precious handful", Zvelebil) of a strong Drav. substrate in the Panjab. At best, one can speak of a few very isolated cases which have been taken over into the RV; clearly this indicates an adstrate rather than a substrate.

This result is important for the time of the immigration of speakers of Dravidian into the Panjab and it specifically underlines that the Indo-Aryans *did not at once* get into contact with speakers of Drav. but only much later, when the tribes speaking IA were already living in the Panjab and on the Sarasvatī and Yamunā. Apparently, Dravidian speakers began influencing the Panjab *only at this moment in time* (cf. Allchin 1995: 31 sqq., see above). *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.

It cannot be argued that the immigration of the Dravidians into the Panjab should have taken place *earlier* than discussed above, for the simple reason that Drav. words do not exist in that early period; the same is the case if only the 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, Fairervis in: Southworth, 1979: 208, 228; contra, Hock 1975: 87f., cf. Southworth 1992: 663), and that in consequence, the Indus inscriptions should be read as Dravidian. In this case, one would expect, after some 400-700 years of the flourishing of the Indus civilization, cases of bilingualism. Consequently, much more Drav. influence should have been retained than visible in the few (late) words found in the c. 380 'foreign' words. One would expect at least a few important loan words from the fields of trade, handicraft or state organization (at least, from the post-Indus, village level type cultures). This, again, is *not* the case. *Paṇi* '(rich) foreigner, demon' cannot be connected with 'trader' inside the RV,

and *paṇ* 'to barter' appears first only in (post-Rgvedic) KS, *pra-paṇa* 'trade' AV, *prati-paṇa* 'exchange' (see EWA II 69, DEDR 3884 does not help: *paṇ* 'work, service', *paṇikkān* 'carpenter'; cf. Kuiper 1955: 168, on *vāṇa*, *vaṇij*.) In addition, there are not many designations of RV artisans, except for IA *takṣan* 'carpenter', etc. (see below). Even if Drav. had been the traders' language, one would be at loss to answer the question why Drav. influence is only seen in the middle and late RV as well as later one (AV+).

Summing up, *early* Dravidian influence in the Panjab can be excluded, but must be *explained* for the following middle and later RV periods (cf. also Kuiper 1997: 7 sq). This is best done by the scenario mentioned above: middle and later RV immigration of Drav. speakers from Sindh. Incidentally, it must be noted that in *all* of the RV, there are no typical Drav. words for agriculture which should be expected if the Indus people of the Panjab had been speakers of Dravidian. This agrees with the reconstruction of Fairservis (1995), Southworth (1979, 1988, 1990: 663 'an "Indus" or "Harappan" language or group of languages'), and McAlpin (1979) of early Dravidian: an originally pastoral society that acquired agriculture only in South Asia. All of this indicates that we have to take a closer look at the regions bordering the Panjab in the South, especially Sindh.

### §5.1. Greater Sindh

In contrast to the clear picture of the Panjab in Rgvedic times, the situation in Greater Sindh is much more vague and the following results must remain tentative. The RV does not mention this area as such, yet there are some indications that Sindh and neighboring Baluchistan were known. First of all, the *Bhalānas* tribe took part in the Ten Kings' Battle (RV 7.18) that settled the suzerainty of the Bharata chieftain over the Panjab tribes. The *Bhalānas* are identified with the *Bolān* pass and river near Quetta in Baluchistan. Unfortunately, southern local rivers are not mentioned anywhere in the RV south of the Gomatī (Gomal River).

However, data from RV book 8 may supplement our scanty information. Book 8 has long been connected with Eastern Iran: K. Hoffmann (1940 = 1975: 1 sqq.) has pointed to Iranian looking names such as *Kaśu* ~ Avest. *Kasu-* (EWA I 330), *Kaśu Caidya* 8.5.37, *Kanīta* ~ Scythian *Kanītēs*, cf. further *Tirindira* 8.6.46 ~ *Tiridatēs* ~ Avest. *Tīrō.nakaθβa*, *Kṛśa* 8.59.3 ~ *Kērāsāspa*, *Parśu* 8.6.46 ~ OP *Pārsa* 'Persian', *Paktha* 8.22.10 (mod. Pashto, Paktho), *Varo Suśāman* 8.60.18 (with unusual Sandhi), *Arśāsāna* 8.12.9, 2.20.6, etc., *Anarśani* 8.32.2 ~ Iran. *ərśan-*? All such names, if Iranian, belong to pre-Iranian tribes that spoke a dialect close to the one that later developed to E. Iranian (cf. the similar case of the Mitanni-Aryans, below). Book 8 also knows of camels (*uṣṭra* 8.4.21-24, 31, 46-48, O. Iran. *uṣṭra*, as in *Zaraθ-uṣṭra*), that are first attested archaeologically in S. Asia in the Bolān area, at Pirak, c.1700 BCE.

However, data from RV book 8 may supplement our scanty information. Book 8 has long been connected with Eastern Iran: K. Hoffmann (1940 = 1975: 1 sqq.) has pointed to Iranian looking names such as *Kaśu* ~ Avest. *Kasu-* (EWA I 330), *Kaśu Caidya* 8.5.37, *Kanīta* ~ Scythian *Kanītēs*, cf. further *Tirindira* 8.6.46 ~ *Tiridatēs* ~ Avest. *Tīrō.nakaθβa*, *Kṛśa* 8.59.3 ~ *Kērāsāspa*, *Parśu* 8.6.46 ~ OP *Pārsa* 'Persian', *Paktha* 8.22.10 (mod. Pashto, Paktho), *Varo Suśāman* 8.60.18 (with unusual Sandhi), *Arśāsāna* 8.12.9, 2.20.6, etc., *Anarśani* 8.32.2 ~ Iran. *ərśan-*? All such names, if Iranian, belong to pre-Iranian tribes that spoke a dialect close to the one that later developed to E. Iranian (cf. the similar case of the Mitanni-Aryans,



below). Book 8 also knows of camels (*uṣtra* 8.4.21-24, 31, 46-48, O. Iran. *uštra*, as in Zaratuštra), that are first attested archaeologically in S. Asia in the Bolān area, at Pirak, c.1700 BCE.

The area west of Sindh, Makran or Gedrosia, is known in Old Persian as *Maka* and its people as *Maciya*; this continues the old Mesopotamian designation *Makan* (Sumer. *Má-gan*, Elam. *Ma-ak-qa*, Akkad. *Ma-ak*, Greek *Mákai*) which included the other coast of the Gulf, in Oman. It may be that indigenous populations held on in this area for a long time as it is altogether missing in the list of "Aryan" countries in the Avesta (V. 1). Along this coast and the few rivers flowing into the Gulf, there were many Indus settlements. Further inland, the oasis along the Bampūr river was known to the Mesopotamians as *Marhaši*, an area that no longer belonged to the Elamite speaking lands which extended from Susa and Anšan to Simaški (Tepe Yahya/Shahdad).

Now, apart from RV 3 and 7, Drav. words occur first in the Middle RV book 8, more specifically in its Kāṇva section (RV 8.1-48, and 8.49-59, 60-66); they include *kuṇḍa-* 8.17.13, *mayūra* 8.1.25, *naḍa/nala* 8.1.33 (see below); note also the many words in RV 8 with retroflexes (Kuiper 1991: 17, Hoffmann 1941, 1975:16, Kuiper 1967: 84 n. 18, 86 n. 26).

If one takes all of this seriously and locates at least the Kāṇva sections of book 8 in East Iranian lands, that is in (S.W.) Afghanistan and Baluchistan, one can also adduce the very name of this clan of poets. K. Hoffmann (and I) have connected the name with *kṛ* 'to act magically, to do sorcery' (Hoffmann 1975: 1 sqq., Witzel 1983-5). Kuiper (1991: 80) has correctly objected there also is *Pra-skaṇva*, with the common Indus prefix *pra-* \*[p̄r-]. This may mean that the Indus language extended to Eastern Iran, especially to the area west of Sindh, 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?).

Dravidian influence in Middle R̥gvedic (the time of king Sudās) can be traced back, with some probability, to the areas from Arachosia to Sindh as well. It is here that Drav. place names are assumed to appear first (cf. L.V. Ramaswamy Iyer 1929-30). These names (showing MIA development *p > v*) extend from Sindh via Gujarat and Maharashtra to the South: Sindhi *-vali*, Gujarati *-wāri/warī* (Sankalia 1949), Mar. *-oli*, all from a Drav. word for 'village' (Tam. *palli* 'hamlet', Kan. *palli*, *halli*, Tel. *palli* 'village', Kur. *palli* DEDR 4018, CDIAL 7972, see Parpola 1984, 1994: 170 sqq., 1997; Southworth 1995: 271, see further, below).

A similar view has been proposed, on the basis of linguistic and archaeological observations, by Zvelebil (1972, 1990: 48, 123), Southworth and McAlpin,<sup>3</sup> and Fairservis (1992: 17, 21). It has to be underlined, however, that McAlpin's reconstruction of an Elamo-Dravidian language family has not been accepted by Dravidologists. Fairservis and Zvelebil think of an immigration by Drav. speaking tribes at c. 4000/3500 BCE, from the

---

<sup>3</sup> McAlpin 1981 is based on the lexico-statistic calculation of P. Gardner 1980; he 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.

mountainous lands of East Iran into the Indus valley. Both underline data that characterize the Dravida as originally pastoral hill tribes.

In sum, we may reckon with early Drav. pastoralists (Fairservis 1992, 1997) in Baluchistan and later on, after a period of acculturation with the Indus people, we may encounter Drav. farmers (Southworth 1979, 1990, 1995) who practiced intensive rice (Kenoyer 1998: 178, Jarrige 1985) and millet cultivation in Sindh.

### §5.2. The languages of Sindh

In addition to these western (Dravidian, pre-Iranian) elements there also are local 'Sindh' ones. First of all, it is precisely in this area that rice was first introduced into the Indus civilization. It occurs first as *odana* 'rice gruel' in the (partly E. Iranian) Kaṇva book (RV 8) in the Emuṣa myth, which clearly smacks of 'foreign' origin: RV 8.69.14, 8.77.6-11, 8.77.10, (cf. also 8.96.2, 1.61.7, and in *vy-odana* 8.63.9; summary and discussion by Kuiper 1991: 16 sqq.) He had explained it earlier on (1950) as Austro-Asiatic, but is more cautious now (Kuiper 1991: 18f., cf. below). On closer observation, we can notice a mixture of an IA, Austro-Asiatic and possibly Drav. myth.

Kuiper (1991) now shows that the Kaṇvas, non-IA local sorcerers, introduced this myth into the RV. At any rate, the motif is unusual for the RV. Its hero is a divine bow shooter (probably seen on an Indus copper plate, only at Mohenjo Daro, in Sindh, Parpola 1997: 39; cf. also Avesta, Yt. 8.6.37 *ərəxša*, Kṛṣānu RV 4.27.3, Rudra, and Murukan in S. India; for 'bow' see KS *dālbhūṣī*, MS *drumbhūṣī*; with PDrav *-r-* > [l] / [ž], Kuiper 1991: 26). This bow shooter splits a mountain, finds the *odana* rice gruel and kills the boar *Emuṣa*. The myth is an imitation of the well known Ṛgvedic Vala myth (splitting the mountain cave containing the cows/dawns), but is otherwise completely alien to the RV.

Now, the suffix *-uṣa* (Kuiper 1991) of *Emuṣa* clearly indicates a name taken from the (Para-Munda) Indus language. This points to a *late* myth (because a latecomer, rice, is important), adopted from the local southern or southwestern Indus region and from beyond.<sup>4</sup> Second, the word for 'rice' occurs in a Sindh and a Panjab variety (see below). The Sindh version, closer to Dravidian, has been transmitted further west, along the southern trading route to Fars and has entered western languages from there (Greek *oryza*).

Whether rice was otherwise known to the Ṛgveda is doubtful. Rice was introduced towards the end of the Indus civilization in its southern areas, in Sindh (Kenoyer 1998: 178, in Pirak, along with newly introduced sorghum and millet, and also horse, donkey, camel). In this case, we have again to reckon with a (West-)Munda word: *odana* is connected with *oḍi(kā)* 'wild rice' (lex., CDIAL 2546) and Santali *horō*, *huru* 'rice plant' (EWA I 280) and explained as Munda loan (Berger 1963: 420, Kuiper 1950: 179; but cf. Zide and Zide 1973: 8-9 on Mundari *kode*, Kharia *kuda* 'millet, *ragi*'). Together with the introduction of rice its *charter myth* (Malinowski) may have been taken over as well. As has been mentioned, the Dravidians originally had neither a word for 'rice' nor for the staple food of the Indus civilization, wheat.

<sup>4</sup> It has to be observed that the boar does not play a role in the Indus civilization: "apparently not domesticated, not used in Indus economy" Kenoyer 1998: 165; this rather seems to be an eastern phenomenon (thus Munda?); cf. below Munda and Sino-Tib. 'pig' and cf. the ancient boar cult on the Nicobar Islands.

In sum, it can be said that we may have to reckon with a combination of several factors in the southern Indus area: with the (Para-Munda) Indus language, with some more eastern Munda influences, with immigration from E. Iran in the person of Vasiṣṭha (RV 7) and of (pre-)Old Iranian tribes into Baluchistan and the neighboring Kachi plain of the Indus valley (e.g. at Pirak, 1700 BCE), and with Dravidian immigration.

As mentioned above, Zvelebil (1970, 1990) is of the opinion that the Dravida entered South Asia from the Iranian highlands. Their oldest vocabulary (Southworth & McAlpin) is that of a semi-nomadic, pastoral group, not of an agricultural community. They are thus not expected to have their own word for 'wheat'. Wheat, however, was the staple of the Indus civilization, and was 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' (details below). If the Dravidians acquired agriculture only in the hills bordering S. Asia, they may very well have been inhabitants of Baluchistan at the time. At any rate, neighboring Sindh, just as Gujarat and Maharashtra, show place names that are explainable from Dravidian \**palli* (see above). Then, according to archaeology, a large section of the population of Sindh left this area towards the end of the Indus period. They moved 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.

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 Kassites came to Mesopotamia before the Mitanni-IA. In that case they knew the horse already in Central Asia, but would not have taken it over directly from the Indo-Iranians (as may be indicated by Brahui (*h*)*ullī*, O.Tam. *ivulī* 'horse', etc., different from IIr. *aεva*). In other respects as well, they have not been influenced by the Indo-Iranians.

One can even assume that the early testimony of the introduction of horse and camel from the Iranian plateau into Sindh (Pirak and Kachi plain in western Sindh) is due to the Dravida (c. 1700 BCE, Kenoyer 1998: 178; Allchin 1995: 31). In that case, it must be investigated why they apparently did not preserve a word for 'camel'. In this fashion, that is through the mediation of the Dravida in Sindh, Drav. \**variñci* 'rice' must have reached Iran (> M.Pers. *brinj*), that is not, as otherwise common, via the northwestern Khaiber Pass, as in this region another form of the word is found, with \**vrijhi* > Pashto *wrizē*, etc. (see below).

This may mean, on the one hand, that the Dravida themselves were immigrating at the time of the older RV, or that they only influenced the Panjab in the later, Middle R̥gvedic period, coming from Sindh. This is perhaps supported by archaeological facts, for Sindh was practically deserted by its population in the post-Indus phase (Allchin 1995: 31 sqq.) It is from this Southern basis that they suddenly appear in mid-level RV, with names such as *Kavaṣa* 'straddle legged' (*K. Ailūṣa* RV), cf. *Śailūṣa* "dancer, singer" VS (EWA II 655, Kuiper 1991:20, 25, 42) which Kuiper 1991: 24 explains with reference to Dravidian: initial *c-* is often dropped in South (!) Dravidian; further examples in RV are : *Śirimbiṭṭha* : *Irimbiṭṭhi* EWA II 639, cf. also *śiriṇā* 'hiding lace, night?' : *iriṇa* 'salt pan, hiding place (for gambling)' (Witzel 1999).

Ailūṣa is important, as it was this poet who was an important priest, on the side of the opponents of the Bharata. (These opponents included the Bhalānas). His great-grandson Tura Kāvāṣeya, however, is an important priest of the Kuru realm that succeeded the Bharata 'kingdom'; he developed the Agnicayana ritual (Th. Proferes, Harvard Ph.D. thesis 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 (that of the post-RV, classical Śrauta ritual, see Proferes). Further, he was not classified as Śūdra but obviously as a Brahmin who had learned to compose RV hymns in the traditional poetic IA language! All of this is indicative of a high degree of amalgamation and language acquisition at this time, during the middle and late Ṛgveda period (see below).

### §5.3. The Southern Indus language: Meluhhan

However, there are indications that another 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. It is perhaps best to call this language "Meluhhan" after the name the Sumerians gave to the country, Meluḥḥa. Its language 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). This indicates that 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, which allows for a hypothetical identification of the Marhaši language with that of Meluḥḥa and makes a thorough investigation of the data of RV 8 (see §5.1.) even more important. 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 Meluḥḥa include *giš-ab-ba-melu-ḥḥa* (*abba* wood, a thorn tree), *mêsu* wood ('of the plains'), ships of Meluḥḥan style (*magilum* boat) (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 Meluḥḥa', *zaza* cattle (zebu?), elephants, etc. (cf. Landsberger, *Die Welt des Orients* 3. 261). As coming from Dilmun (Bahrain), we may add the Meluhhan(?) trees *giš-ḥa-lu-ub* or *ḥaluppu* wood, *giš-mes-makan* or *mêsu* wood of Magan, and the *gišgišimmar* wood (cf. above \**šimma* in *šimbala*, *šalmali* 'Salmalia malabarica!') A slightly later(?) loan-word relationship is seen in Sumer. *ili* 'sesame', Akkad. *ellu/ūlu* 'sesame oil', which is only found in South Drav. with *eḷ*, *eḷḷu* 'Sesamum indicum' (D. Bedigian 1985); the word can be compared, however, with Ved. *tila* and *jar-tila* 'sesame' which shows the typical Para-Munda prefix *Cər-* (cf. Kuiper 1955: 157 for a Munda origin). The ultimate source, \*(*t*)*il*, however, is unclear (cf. Blažek and Boisson 1992 on Sumer. loans in Dravidian, see below §6).

The word *meluḥḥa* 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 *Meluh̥ḥa* 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 MIA *\*mliccha* > Sindhi *milis*, Panjabi *milech*, *malech*, Kashmiri *brichun* 'weep, lament' (< *\*mrech-*, with the common *r/l* interchange of IA), W. Pahari *melēch* 'dirty'. It seems that, just as in other cases mentioned above, the original local form *\*m(e)luḥ* (i.e. *m(e)lukk* in IA pronunciation, cf. E. Iranian *bāxδī* 'Bactria' > AV *\*bahli-ka*, *balhi-ka*) was preserved only in the South (Gujarat? > Pali), while the North (Panjab, Kashmir, even ŚB and Bengal) has *\*mlecch*. The sound shift from *-ḥḥ-/-kh-* > *-cch-* is unexplained; it may have been modeled on similar correspondences in MIA (Skt. *akṣi* 'eye' ~ MIA *akkhi*, *acchi*; *kṣetra* 'field' ~ MIA *khetta*, *chetta*, etc.)

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 *Meluh̥ḥa*, in Baluchistan-Sindh, and have been transmitted for a long time in a non-literary level of IA as a nickname, before surfacing in E. North India in Middle/Late Vedic as *Mleccha*.<sup>5</sup>

Further examples of the Southern Indus (Sindh) language include the designations of plough, rice, wheat, and millet.

### Plough

The old agricultural word *lāṅgala* 'plow' (RV, 4.57.4, a late hymn) is found, in a divergent form, in Tam. *ñāñcil*, *nāñcil*, Kan. *nēgal*, Gadba *nāngal* (DEDR 2907). Southworth (1988; 1979: 200, 205; 1995: 268, cf. Kuiper 1948: 127, 1955: 156, Przuludski BSL 24, 118 sqq., cf. Parpola 1994: 168) assumes a popular etymology PDrav. *\*ñān-kal*, *\*ñān-kel* 'earth stone' and traces the term back an Austro-Asiatic source, Munda *\*ñā-kel*, *ñān-kel* (Zide & Zide 1973: 5), Santali *nahel*, Khasi *lynkor* [*lənkor*] < *\*lēnkol*, Khmer *aṅkal*; cf. also the Austronesian forms, Malay *tengala*, Makassar *naṅkala* (Bagchi 1929, 9). V. Blažek and C. Boisson (1992: 17-19) add cognates from Austroasiatic (Vietnamese *cay* < *\*kal*, etc.), Austronesian (Cham *langal*, *langar*, Batak *tingala*, Bugi *rakala*), Sino-Tibetan (Kanauri *hālon*) etc.; they think of a Sumerian, and ultimately perhaps even an Afro-Asiatic origin of this widespread word of culture: Sumer. *níg-gala<sub>x+1</sub>* or *níg-gál* 'sickle' ('the semantic shift ... may seem far-fetched', 1992: 19), and Afro-As. *\*nigal* 'to reap; reaping sickle'.

<sup>5</sup> Pali *milāca* is influenced by a 'tribal' name, Piśāca, as is Sindhi *milindu*, *milidu* by Pulinda; the word has been further 'abbreviated' by avoiding the difficult cluster *ml-*: Prākṛt *mecha*, *miccha*, Kashmiri *mī.c.(h)*, Bengali *mech* (a Tib.-Burm tribe) and perhaps Pashai *mecə* if not < *\*mēcca* 'defective' (Turner, CDIAL 10389. -- Parpola 1994: 174 has attempted a Dravidian explanation. He understands *Meluh̥ḥa* (var. *Melah̥ḥa*) as Drav. *\*Mēlakam* [*mēlaxam*] 'high country' (= Baluchistan) (= *Ta-milakam*) and points to Neo-Assyrian *baluḥḥu* 'galbanum', *sinda* 'wood from Sindh'. He traces *mlech*, *milakkha* back to *\*mlekš*, which is seen as agreeing, with central Drav. metathesis with *\*mlēxa* = *mēlaxa-m*. Kuiper 1991:24 indicates not infrequent elision of (Dravid.) *-a-* when taken over into Skt. -- Shafer 1954 has a Tib-Burm. etymology *\*mltše*; Southworth 1990: 223 reconstructs PDrav. 2 *\*muzi/mizi* 'say, speak, utter', DEFR 4989, *tamil* 'Tamil' < 'own speech'.

However, the Munda words do *not* agree with Ved. *lāṅgala*, though one can easily assume dissimilation of *n-l*. The word underlying RV *lāṅgala* must have come from an intermediate language, in short, the Panjabi form of the Indus language (Para-Munda), with *\*laṅgal*. This form cannot have been that of the Southern Indus language (Meluhhan) as this has resulted in Drav. *\*ñāṅkal, ñāṅkel*. While the difference is small here (*g/k, n/l*), it is more substantial in other agricultural words.

### Rice

The word for 'rice' shows a difference between a Northern form, approximately *\*\*\*(ə)βərij*, versus a southern one, *\*vari, (v)ariki, variñci*. Note that this indicates the same difference in tenuis/media as met with in the word for 'plough':

N. *\*laṅgal, \*vəriji* :: S. *\*naṅkal, \*variñci/variki*.

Still another form exists in Proto-Munda *\*ə-rig*; it has provided Dravidian *\*(v)ari, variki* > Tam. *arici, ari*, Kan. *akki* (DEDR 215), and also Tam., Tel. *vari* (DEDR 6565).

Though rice is indigenous to S. Asia, the domesticated version can be traced back to S.E. Asia and S. China.<sup>6</sup> It has been found in India since the 3rd millennium BCE (Glover & Higham 1996, Kajale 1991), and appeared late in the southern Indus civilization, at Pirak c. 1700 BCE. However, it appears first (as *vrihi*) only in post-RV texts (AV, c. 1200 BCE), though it probably was an ingredient in the RV offerings *puroḍāśa* 'rice cake' and *odana* 'rice gruel'. The older IA grain is only *yava* 'barley', but later on we have 7 or 10 agricultural products: in the Yajurveda Saṃhitās, the 'seven agricultural plants' (*sapta grāmyā ośadhayah*); ŚB 14.9.3.22 has even ten: *vrihi* *Oryza sativa* L.; *yáva* *Hordeum vulgare* L. subsp. *hexastichum* (L.) Schinz et Kell.; *tíla* *Sesamum indicum* L.; *māśa* *Phaseolus mungo* L. var. *radiatus* = *Phaseolus Roxburghii*; *áṇu* *Panicum miliaceum* L.; *priyāṅgu* *Setaria italica* (L.) Pal. Beauv. = *Panicum italicum* L.; *godhúma* *Triticum aestivum* = *Triticum sativum* Lam.; *masúra* *Lens culinaris* Medic. = *Ervum lens* L.; *khálva* *Phaseolus radiatus* L. a variety of *Phaseolus mungo* L. = *māśa*(?); *khalá-kula* *Dolichos biflorus* L. (W. Rau 1997: 203-206).

Southworth (1979, 1988: 659-660) supposes an Elamo-Dravid. origin: *\*var* 'seed, grain', Elam. *bar* 'seed', PDrav (stage 1, c. 2000 BCE) *\*vari* 'rice grain'. (McAlpin 1981, Tyler 1968, Southworth 1988). Achaemenid Elam. *umi* 'grind (grain)', *\*um* 'to process grain', PDrav1 *\*um* 'husk, chaff' DEDR 637; (this should be compared with *\*gant-um-a, gandh-um-a*!). However, the Elamo-Drav. family has not been proven to the satisfaction of Dravidianists (McAlpin (*et al.*) 1975, Krishnamurti 1985, Zvelebil 1985), and the N. Drav. language Brahui, seen as a link by McAlpin, is a late-comer to Baluchistan (Elfenbein 1987). Southworth (1988: 664) stresses the difference between northern (Gangetic) and southern rice, which might have been dry land rice.

On the other hand, Southworth later on mentions that PDrav *\*(v)ariki* DEDR 215, has been taken over from PMunda at c. 1500 BCE: *\*ərig* 'millet, *Panicum militare*' (Zide & Zide 1973: 8) --> *\*ariki(i)* 'staple grain' (Southworth 1988: 660), because the South Drav. sound change *k > c* took place only between the second and third stage of Drav.

<sup>6</sup> The earliest archaeologically found rice is said to come from Koldihwa near Allahabad (c. 5440/5430 BCE or even earlier); this has been doubted. A more probable date is c. 4000 BCE, at Chirand in Bihar. -- Some trace the terms for rice back to Sino-Tibetan (see Blažek and Boisson 1992: 27 n. 40).

(Krishnamurti 1969); thus: Munda \**ərig* --> Drav. \*(*v*)*ariki* > Tamil *ari*, *arici*. This South Dravidian form *arici* has been transmitted westwards, probably by sea trade, Greek *óryza*, *óryzon* and Arab. *ruz*, Engl. *rice* etc. (Southworth 1979: 202, cf. EWA II 598).

Southworth also reconstructs PDrav. \**vari*, \**variñci* DEDR 5265. This, too, was transmitted westwards, but via the Baluchistan-Bampūr trail, to Old Iranian as \**brinj*, M.Iran. *brinj*, N.Pers. *birinj*). It must have been this form that was the basis of the word in the late Southern Indus civilization.

The northern track westwards is attested by Ved. *vr̥hi* < pre-IA \**vrijhi*- and reflected in the E. Iran. (and N. Iran.?) languages: Pashto *wrižē*, (but Khotan. *rr̥iysua* [*r̥izua*]!), Nuristani *wrič*, *r̥ič*. (cf. Fussman 1972).

The Northern Indus dialect had \**vrij* > Ved. \**vrijhi* > *vr̥hi*, Nuristani *wrič*, Pashto *wrižē*. The Southern dialect is indicated by M.Pers. *brinj*, N.Pers. *birinj*, going back to \**vəriñji*, Dravidian \**variñci*, a form with "infix" *-n-*, found in central Dravidian: Gondi *wanj̥i* (Pengo *verci(l)*, Gadba *vasil*, DEDR 5265). The form with *-n-* points to Munda origin and to a relatively far-reaching influence or expansion of the Munda in this early period (cf. Kuiper 1955: 140, 1962: 14, 51, 1991: 39f.) Again, this distribution also suggests a difference between, on the one hand, northern or north-western form, including the northern Indus language, and on the other, the southern Indus language and the rest of the subcontinent.

However, these forms have to be reconciled with Tibetan '*bras* [*əbras*] > mod. Tib. [*je*], Purik *bras*, with the neighboring, linguistically isolated Burushaski *bras* (Kuiper 1962: 40, 1955: 143 n. 17, Tikkanen, 1988: 303-325), Dumaki *bras*, and even with some Austronesian forms such as Malay *bəras*--> Somali *baris?*; cf., however, Dayak *bari*, Malegasy *vare*, *vari* --> Bantu *wari*, *wali* (Nurse 1983, Southworth 1988: 664, Witzel 1995) and O.Jpn. *uru-shine*, (cf. mod. Jpn. *uru-chi* < \**uru-ti*). Both *bras* and pre-Vedic \**vrijhi* must go back to a source such as \*\**əβərij* (Witzel 1997b).

In the study of the Asian words for 'rice' we have to take into account words from S., S.E. and E.Asia:

- S. Asia: Ved. *vr̥hi* < \**vrijhi*,  
           Burushaski *bras*<sup>7</sup>, Tib. '*bras*,<sup>8</sup>  
           Drav. \**arici*, \**variñci*;<sup>9</sup>  
           Munda \**ə-rig*,

<sup>7</sup> Southworth 1990: 229, n.10: PIA \**camala/cāvala* < TB *ca-*? (*dza*); cf. Southworth 1974, with an early Drav. substrate in the northwest and in the Gangetic plains: < Tib.-Burm. \**cā* + *vāl/vār* < Drav. *vari*? -- Other IA words for 'rice' (*oryza sativa*): OIA *taṇḍula* < Drav. (Southworth 1988: 660); OIA *śāli* < Tib.-Burm. *cau-* / Austr. *Cśamaq* (Benedict 1990); P.Drav.1 \**manji(k)* DEDR 3790, 'rice plant', but also 'seed' in Kurukh.

<sup>8</sup> Benedict 1972: 123 [*əbras*, '*əbras*]; cf. also TB \**mruw* 'grain, seed' Benedict 43: no.150 Tib. '*bru* 'grain' (and Nepal. inscriptions, with *-brū*, *-bū*, see below), and (?) Lushai *buh* 'boiled rice'

<sup>9</sup> Southworth 1990: 229 n. 9. - In Drav. the word for 'rice' cannot be reconstructed for the early stages (PDrav. 1), where only the meaning 'seed' is found: Kurukh *manji* 'seed in general' and Tamil *arici* 'seed' in: *ēlav-arici* 'cardamom seed' DEDR 768. -- Cf. also Guj. *varī* "particular kind of grain", Mar. *varī* 'grain Coix barbata', Pkt. *varaia* 'a kind of rice'; CDIAL 11328 *varī*, -- all on the Drav. trail South from Sindh.

Tib.-Burm. \**dza*-<sup>10</sup> < Austr. \**C<sub>s</sub>amaq*

Kusunda *cusum* 'rice in husks', *kādiyun* 'cleaned rice'

- S.E. Asia: Munda \**ruŋ-ku*'g (Zide & Zide 1973: 17)

Austr. \**C<sub>s</sub>amaq*

Austrones. \**pajay*;

Austrones. \**i-may*

Thai \**xau* > *khaw* (Haudricourt, in Shafer 1966-7: 522)

Austro-Thai \**kru-may* (> Jpn. *kome*)

- E. Asia: Chin. \**miər*, Tib.-Burm. \**may*<sup>11</sup>

The distribution of the various words for 'rice' points to an old (South)East Asian word of culture. Just as in the modern spread of the E. Asian word 'tea', several routes of distribution have to be distinguished:

1. an approximate reconstruction of the S.(E.) Asian word \**əvrij(h)i*/\**əbras*, probably < \*\**əβərij*,<sup>12</sup> which is spread out in a wide arch between

2. E. Asian \**may*, \**xau*, \**krumay* (< \**kru*-\**may*?)<sup>13</sup> and

3. S. Asian \**ə-rig*<sup>14</sup>, \**ruŋ-ku*(?g).

PMunda \**ruŋ-ku*(?g) (Zide & Zide 1973: 17, \*(*r*)-(n)-*ku*, Kuiper 1962) may be an Austro-Asiatic form with prefix *r*-. This might be connected, via metathesis, with Benedict's Austro-Thai-Japanese \**krumay* (> Jpn. *kome*, *kuma-shine*), a word that may be composed, if Sino-

<sup>10</sup> Ved. *vrihi* has been supplanted in NIA almost everywhere by Tib.-Burm. CDIAL 4749 \**cāmala/cāvāla*, Pkt. *caulā* (pl.), *cavala*, and NIA *bhāt* 'cooked rice' (Southworth 1988: 666); for this see Benedict 1972: 28 no. 66 'to eat', Kanauri *za*, Garo *tsha* 'eat', Lushai *fa*, *fān*, Bahing *dz'a*, Newari *jā* 'cooked rice', *jāki* 'uncooked rice' (cf. Lushai *caw* 'cooked rice', *caw ciar*); the Tib.-Burm. word apparently is a loan from Austro-Thai: \**C<sub>s</sub>amaq*, s. Benedict 1990: 175.

<sup>11</sup> Benedict 1972: 149 n. 408, 491-2 Tib.-Burm. \**may* as early loan-word from Austro-Thai, e.g. Indones. \**imay* 'rice' (but O.Jpn. *yōne*, Jpn. *ine*, *-shine* 'rice plant' < \**yinai*, according to Benedict 1990: 234; cf. also *ne* 'root'); Chin. *miei* < \**miər* 'rice (paddy)', Bodo-Garo \**m[a,e]y*; Karen \**may*; cf. Tib.-Burm. \**s-min* 'ripe, cooked') Benedict 1972: 106 § 432 (< Proto-Miao-Yao \**snañ* 'cooked rice'?, see Benedict 1992: 234).

<sup>12</sup> Benedict 1990: 43 reconstructs Proto-W.-Malayo-Polynes. (Hesperonesian) \**pajay* (Malay *padi*, Javanese *pari*, cf. the Engl. loan *paddy*; however he also has (1990: 77) Proto-Austrones. \**pagr[ə]y*, that differs from the S. Asian/Central Asian cluster \**vrijhi/bras* by a transposed(?) *-r-*, (perhaps: Austric \*\**βə-rəji* / \**pa-Cj/grəy* > \**pagrəy*, \**pajay*??).

<sup>13</sup> Benedict 1990 assumes Proto-Austro-Thai \**krumay*, whence Jpn. *kome*, *kuma(-shine)*. In connection with the Tib.-Burm. and Sinitic forms (\**mi*, *may*, Benedict 1972) a compound \*\**kru* + \*\**may* may be construed. The proto-form \*\**kru* seems to be the source for the words for 'rice' in Sino-Tibetan, Austro-Asiatic and Austro-Thai (including Austronesian).

<sup>14</sup> The Austro-Asiat. words still are very close to those in Austro-Thai: PMunda \**ruŋ-ku*(?g'*b*) < Austro-As. \**ərig*, 'millet, *Panicum militare*'. Pinnow 1959: 96 § 139 derives \**ruñ* from Kharia *ḍuruñ* 'to pound rice' etc. (p. 92 § 116), and *-ku*(?*b*) from Sant. *hoṛo*, Mundari *huṛu* etc. (p.122 § 244), cf. also Kharia *khōsṛō pe*' etc. (p. 171 § 370). -- In Munda there is, next to Kharia *romku*'*b*, also Juang *ru(n)kū*, Sora *ruñkū-n*, Bondo/Remo, Parengi *ruñku*, Gutob *rukū* (Pinnow 1959: 96), and in eastern Austro-As.: Khasi *khau*, Mon *unko*, Khmer *oñkor*; - Thai *khāu* may be a loan word from Austro-As.? Further: Palaung *ra-kō*, Kuoi *añkau*, Sue *rañkao*, Palaung *ra-kō*, Palaung-wa *unko*, Sakai: Krau (Ketiar) *uñ-kuok*, Sakai also: *cənroñ* 'husked rice', Krau (Kuala Tembeling) *rə-kuā*' etc. (Pinnow 1959: 96, Kuiper 1962: 51f.). The variation in Austro-As., already observed by Kuiper, points to a proto-form \*(*r*)(n)-*k(h)u*. - Thus, Dhimal (= Tib.-Burm. Kiranti, eastern Himalaya) *uñkhū* 'rice', according to Kuiper < Munda \**ruñku*.



Tib. (Benedict 1972: no. 65, 128, 149, 192, 193) \**may*, Austrones. *i-may* and Thai \**xau* are compared, of \**kru*-\**may*. In the end, one may think of a Proto-form \*\**kru* as the ultimate source for 'rice' in S.E. and E. Asia (Sino-Tib., Austro-As., Austro-Thai; cf. Blažek and Boisson 1992: 27 n. 40).

The origin of O.Jpn. (Wamyōsho) *uru-shine* (cf. *kuma-shine*), Jpn. *uru-chi* < \**uru-ti* remains problematic. It seem to belong to the S./Central Asian group \**vrij(h)i*/\**bras* and could have been introduced from (S.)China to Japan along with the domesticated plant. The proto-form might have been something like \*\*\*(ə)βərij; the difficult initial cluster \*\*βr- has received various treatments: Drav. *va-/ a-*, Proto-Jpn. \**wuru-* > *uru*, \*βərij > \**vrijhi* > Ved. *vrihi* as *b* is relatively rare in IA and is often replaced by *v* in foreign words.

### Wheat

Further dialect differences between the northern (Panjab) and the southern (Sindh) forms of the Indus language can be observed in the designation of 'wheat'. Though some claim that wheat, the staple of the Indus civilization, is a local domesticate (cf. Allchin 1995: 46, cf. Allchin & Hammond 1978, Kenoyer 1998), it is a western import, as it originated west of the Zagros and south of the Caucasus. In S. Asia it is found as early as the 7th millennium BCE. This leaves several thousand years before the attestation of the S. Asian words for 'wheat', Ved. *godhūma*, Kan. *gōdi* etc.

These are clearly related to Near Eastern ones, e.g. (according to Berger 1959, EWA II 499) \**qend* > Hitt. *kant*, Old Egypt. *xnd*, PSEmit. \**ḥant* (Arab. *ḥinṭatum*). The individual track of the loan word differs, however, just as in the case of the word for 'plough'. A form \**gant-um* (note also PKartv. \**ghomu*), that has entered via the northern Iranian trade route (Media-Turkmenistan-Margiana/Bactria-Aratta/Sistan) has resulted in Avest. *gantuma* and the later Iranian forms: M.Pers. *gandum*, Baluchi *gandim*, Pashto *yanəm* < \**gandūma*?, Yigdha *gondum*, Shugni *žindam*; Khotanese *ganama* < \**gandama*, etc. (see Berger 1959: 40f, EWA II 498). The Iranian form has also been taken over by the Drav. newcomer in the region, Brahui: *xōlum* < IA \**γolum* (CDIAL 4287), according to Berger (1959: 42), however, from Bur. However, Bur. *guriñ*, *gureñ* (pl.), *γárum* < \**γor-um* < \*\**γund-* (Berger), rather seem to have been borrowed from the Indus language. (Berger thought of a loan from Bur. into the Panjab area languages; cf. also Bur. *gur* 'barley, wheat colored', *bur* 'buck wheat' Berger 1959: 43. However, J. Bengtson informs me, by letter of 4/19/99, of the following Macro-Caucasian links: Bur. *gur* 'wheat' ~ Basque *gari* 'wheat' < PEC \**Gōl'e* 'wheat' > Tindi, Karta *qeru*, Archi *qoqol*, etc. (Note that Harmatta, EWA II 499, thinks of an Anatolian \**ghond[ū]*, but cf. Klimov's PKartv. \**ghomu*). How these can be linked to general "Near Eastern" \**qend/kant/gand* remains to be seen. The question of the domestication of einkorn, wheat, etc. in the Near East would play a role in determining when the word could have existed (in PEC) and/or spread east - and westwards.

When this word entered the Panjab it must have changed its initial syllable (\**gan-*) to *go-*, thus \**godum*, a change echoed by the Southern Indus language (\**godi*). Vedic has *godhūma* and similar continuants (Turner, CDIAL 4287). This is a clear folk etymology: the unfamiliar \**gantum/gandum* > \**godum* was analyzed as *go-dhūma* 'cow smoke'.

Another form of the Near Eastern word that has come via the Southern route (Elam/Anšan - Simaški/Tepe Yahya - Marhaši/Bampūr) has resulted in Meluhhan \**gōdi*.

This is retained in Drav. \**gōdi* (Kan. *gōdi*, Tam. *kōti*, cf. DEDR 1906). The change from *-an-* > *-o-* is not unfamiliar in Sindh (see below). A pre-Iranian \**gantum* must have become \**go-tum* or \**go-dum* in Sindh.

The Drav. word, too, seems to be a popular etymology of the unfamiliar \**godum*: 'low red plant', reconstructed by Southworth (1988: 658, 660) as PDrav. 3 at c. 1000 BC as \**kō-tumpai*. Maybe he thought of DEDR 3334 Tam. *tumpai* etc. 'nettle, weed' etc. (cf. Tam. *kōtumam*, Mal. *kōtambu?*). The exact development from \**tumpai* > *-di* would then not clear; (at this supposed late date *kōtumpai* could even be based on RV *godhūma*!)

Obviously, in *this* case both the Northern *and* Southern Indus language have changed *-an-* > *-o-*, while the Northern language otherwise retains *-an-* (see below). The northern form, based on Pre-Iranian \**gantum* would have resulted in Vedic \*\**gan-dhūma* or perhaps \*\**gandha-dhūma* 'perfume smell', cf. CDIAL 4020 Skt. (lex.) *gandhālu* 'fragrant rice', Pashai *gandār* 'a kind of grain'. The Southern (Meluhhan) \**godī* must have influenced a northern \**gantum/gandum* that facilitated a later Vedic popular etymology as 'cow smoke'. The mechanism of this influence is unclear. It may be due to Dravidian influence on the Panjab in the Middle/Late Ṛgvedic period; note that *godhūma* appears only in early post-RV texts.

In short, the inhabitants of the northern Indus region (Panjab) thus must have called their wheat something like \**godum* and those in the Southern Indus region (Sindh), \**godī*.

#### §5.4. Further dialect differences

However, the strange sound change \**an* > *o* is not isolated. It also occurs in the migrant word of culture for 'hemp': Ved. *śaṇa* (AV 2.4.5, PS 2.11.5 *śaṇa*), M.Pers., N.Pers. *šan*, Khotanese Saka *kaṃha* (but Gāndhārī > Niya Pkt. *ṣaṃṇa*), Osset. *gæn*, *gænae*, (Greek *kánnabis*, EWA II 605; Engl. *hemp*, etc.). It appears, again, in Dravidian with popular etymology, as Tel. *gōnu*, *gō:gu*, cf. *gōṅgūra*, Kan. *gōgi*, 'hibiscus cannabinus' (DEDR 2183). The original northwestern form is guaranteed by the North-Iranian (Ossete), Greek and Germanic forms of the loan word: *kanna-bis*, *hemp*, etc. The northwestern dialect has preserved \**-an-*, for example in the Ṛgvedic, yet certainly pre-Indo-Aryan tribal name of the *Gandhārī* (and in the later Vedic country *Gandhāra*). The northwestern name *Śambara* (in the Afghan. hills), too, has not been changed to \**Śobara*, but note the name of a poet in the more southern RV 8, *Sobhari Kāṇva*.

We have a clear distinction between N. Indus *-an-* and Southern Indus *-o-*. (Note that original \**-an-* appears in post-RV texts further east and south, in Dravidian, as *-o-*). This is again a point that may turn out to be of importance for the decipherment of the Indus script which indeed has several features (special signs) that are different in Harappa (N) and Mohenjo Daro (S), (see B. Wells 1998).

This is the opportune moment to briefly discuss another northwestern peculiarity, the interchange of *k/ś* in Vedic. This has occasionally been observed, even one hundred years ago in the case of *Karkoṭa/Śarkoṭa*, but it has not been put into proper relief (Kuiper 1991: 41, 42, 44 as Proto-Munda, cf. KEWA III 309, Witzel 1999). The interchange of *k* and *ś* is not related at all to the well-known Indo-Ir. development of IE \**k'* > Ved. *ś*, as the present

variation occurs only in 'foreign' words; (note also the curious development, in post-RV Skt., of *kṣā* > *khyā*, Witzel 1989).

The name of the snake demon *Śarkoṭa* (AV) appears also as *Karkoṭa(-ka)* RVKh 2.14.8, and locally especially in Kashmir and Nepal; cf. Bur. *hergin* (Berger *hargín*) 'dragon' or rather *ṡárqa* (Berger *ṡárqas*: CDIAL 3418?) 'lizard', Skt. *karkaṭa* 'crab', Mundari *kaṛkom* etc. (Pinnow 1959: 341 §483d). The prefix *śar-/kar-* can be connected with [*sər-*] of the '300 foreign words' (Kuiper 1991: 40-1, 1948: 121), for example in *Sṛbinda* (Kuiper 1939 = 1997: 3 sqq.), *Ku-sur(u)-binda*, *binda* (the mod. Bind tribe; probably also the name of the Mountain range, post-Vedic *Vindh-ya*).

Further materials include *kambala/Śambara* 'blanket/name of a demon', *kabara/śabara*, *kīsta/śīṣṭa* 8.53.4 (with var. lect. *śīṣṭ-*, *śīrṣṭ-*, *śīrṣṭr-*, see above), *Kimīdin/śimidā* 'demon/a demoness', *kambu/śambu* 'shell' (Kuiper 1955: 182), cf. *Kū-śamba*, *Kau-śamba* 'name of a person', cf. *ki-śora* 'filly' AV, 'youth' CDIAL 3190 : *śi-śu* 'baby', *śi(m)-śu-māra* 'Gangetic dolphin', *śiśūla* 'dolphin' RV (EWA II 641-2; Lévy, in Bagchi 1929: 121 sqq.), *Kirāta/Cilāda* 'a mountain tribe', *kiknasa* 'ground grain' AB: *cikkasa* 'barley meal' lex., Bur. *ṣon* ~ Ved. *kāṇa* 'blind' RV.

The realization [*k'*] or [*ś*] of an unknown phoneme (probably *k'*) would easily unite such words as *Śam-bara* : *Kam-bala*, *śabala* : *kabara*; it would also offer a better candidate for Pinnow's unexpected reconstruction for the Munda and Mon-Khmer self-designation \**Ṣqawar* > *Śabara* AB, and in the tribal names > *Sōrā*, *Hor*, *Kora*, *Kherwar*, *Koro/Korku*, *Khmer* etc., Pinnow 154 §311); rather from \**k'awar*, \**k'amwar*.

In consequence, Vedic loan words with the interchange of *ś* / *k* may go back to a phoneme *K'* with realization close to [*k'*] or [*ś*] in the *Indus language*.

## Millet

Another dialect difference can be observed in the "new" import at the time of the Indus civilization, millet. This domesticated plant has originated in China and another variety in Africa (Southworth 1988: 665, Randhawa 1980: 504; Nurse 1983, summarized by Cavalli-Sforza 1995, see now Meadow 1998). The Archaic Chinese words have no similarity to the Indian ones (Karlgren 1923, no. 543 \**liang* < ,*liang* 'millet, sorghum', 1095 \*,*tsi* 'common millet', 1051 \*,*tsi* < *tsiək* 'panicked millet, god of agriculture', 903 \*,*ṣu* < *d'z*'*iuet* 'glutinous millet', 135 \**si<sup>w</sup>ok* 'rice, millet', 914 \**si<sup>w</sup>o* 'glutinous millet'), and the source of the Indian words has not been established so far: any language between the Sahel belt and Baluchistan is possible.

Millet is important as it can be grown outside the winter period (wheat, barley), during the monsoon. The onset of its cultivation in S. Asia coincides with the increasing spread of rice (Kenoyer 1998: 163, 173, 178, Glover & Higham 1996) which has markedly influenced the archaeologically attested emigration of the Indus people towards the Gangetic plains, and towards Gujarat. Even a middle Vedic text, *Aitareya-Brahmaṇa* 3. 45, still knows about this (Witzel 1987: 185).

However, the original source of the S. Asian word in Africa or in one of the intermediary languages has not been determined. It has to be noted, that in the case of this comparatively late import, *-an-*, *-am-* has been preserved both in Proto-Munda \**gaṅgay*, Dravidian DEDR 1084 *kaṅgu* (Tam. *kaṅku*), DEDR 1242 *kampu*, Ved. *priyaṅgu*, OIA

dialects *\*kaṅkuna*, *\*kaṅguna*, *\*taṅguna* (which may provide some indication of the time frame for the words discussed above).

Even though comparisons between the various words for 'millet' can be made, they cannot be traced back, as is the case with many widely spread loan words, to a single source. Hindi *kaṅnī* can be compared with OIA *\*kaṅkunī* CDIAL 2606, with Tamil *kampu* DEDR 1242 and with Munda *\*gaṅ(-)gay* (Southworth 1988: 660, Zide & Zide 1973: 8). The source of these words may have had a form such as *\*\*kaṅ-CV*. From this, Ved. *priyaṅgu* (EWA II 190) can be derived as well, as it seems to have been changed by popular etymology, like several other agricultural terms: prefix *\*pər-* (Kuiper 1991: 42f.) > *\*priya+gu* 'dear cow'. Other IA designations of millet are: Ved. *aṅu* and *\*aṅuni* CDIAL 195. All of this points to a contamination or cross of *\*kaṅgu* and *\*(k/g)aṅgu* --> IA *aṅu*; (*\*al* 'to mill' EWA I 55; rather a Munda change, Pinnow 1959: 198f., *k/\*q* > *0* typical for Sora, Kharia *k* : Sora *0*; thus: *kaṅgu* : *\*aṅgu* --> Ved. *aṅu*, cf. Kuiper 1991: 38). In short, all major language families of S. Asia have taken over the word from an unknown, but not exactly the same source.

Nevertheless, a clear difference between Northern and Eastern/Southern forms is visible: PDrav. *\*kampu* is opposed to PMunda *\*gaṅgay* (Zide & Zide 1973), while the IA forms stand in between the two. The usual IA form is Ved. *aṅu* (cf. Old Indo-Aryan *\*aṅuni*, Turner, CDIAL 195). However, based on Ved. *pri-yaṅgu* < *\*pər-gaṅgu*? and the reconstructed OIA forms *\*kaṅkunī*, *\*kaṅgunī*, *\*taṅgunī* (CDIAL 2606), a northwestern Indian *\*kaṅkun*, a central-northern *\*kaṅgun*, a more eastern North Indian *\*taṅgun* can be reconstructed for the pre-Vedic period, while the Southwest must have had, next to Drav. *\*kampu* DEDR 1242 (= Skt. *kambū* Hemādri) also a form *\*kaṅgu* CDIAL 2605, DEDR 1084. The northern Indus language should have had *\*kaṅku(n)*, its southern dialect (Meluhhan), *\*kaṅgu*.

The modern languages also do not agree: In Hindi (Masica 1979: 76 sqq., 135f.) we find various terms for the many varieties of millet: *kaṅnī* (*\*kaṅkunī* CDIAL 2606); *kuṭkī* (Masica from Skt. *kuṭakā*, not found in the dictionaries; cf. *kuṭaka* 'a kind of tree' KauśS.); *kodon* (CDIAL 3515 *kodrava* 'grain eaten by the poor' Mbh., cf. *koradūṣa* 'idem' Suśr., *-ka* KŚS; DEDR 2163 Tam. *kural*, Kan. *koṛale*, *korle*; Konda *koṛen* 'a grain'); *khil* (Masica: from Skt. *khīḍ*), *junhār*, *j(u)wār*) (*\*yonāla* > *yavanāla* > *juār*, < Drav. *\*connel*, DEDR 2359, DEDR 2896, CDIAL 10437); *bājra* (Vedic: HŚS *varjarī*, CDIAL 9201 *\*bājjara*); *ma(ṅ)ṛūa* (CDIAL 9728 < *madaka* 'the small grain Euleusine corocana'); *sāṅwāṅ* (Ved. *śyamaka* VS, CDIAL 12667). Some of them belong to the c. 30% of agricultural vocabulary in Hindi that comes from Masica's "Language X".

Finally, the word for 'peacock' must go back to a northern Indus form *\*mayur* > Ved. *mayūra* RV level II, and to a southern form *\*mayil/r* > Drav.: Tamil *mayil*, Irula *muyiru*, Tulu *mairu*, Konda *mrīlu*, *miril* etc.

In summing up, it can be stated that in the north-west and also in the Panjab, as represented by loan words in most of the RV, original northwestern *\*-an-* is opposed to southern *-o-*. The same relationship is also found in north-western *ś* : subcontinental *k*, north-western *-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.

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 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 Meluhhan variety of the Indus language was the 'original' language of Sindh. Was it also the Indus trading language? In that case, it has disappeared, just like Sumerian and Elamite, and traces may at best be found in Sindhi -- a step that has not been taken. There is no etymological dictionary of Sindhi.

#### §6. Dravidian immigration

The observations about the early linguistic evidence from Sindh, made above, indicate that Dravidians were not a primary factor in the population of the Indus civilization, even of Sindh, and that they were immigrating into the Panjab only in middle R̥gvedic times. But when could they have entered South Asia?

Earlier scholars (Heine-Geldern 1964, Pinnow 1954: 15) thought that they entered S. Asia (sometime as late as the early 1st millennium BCE) and proceeded via Baluchistan, Sindh and Gujarat to S. India (Zvelebil 1970, 1990: 48, 123). Indeed, their tracks are still visible in certain place names in Sindh, Gujarat and Maharashtra (see above). According to Southworth and McAlpin, however, the semi-nomadic speakers of Dravidian who even had contacts in Iran with the pre-immigration Indo-Aryans (Southworth 1979: 203, 228 f., 1990: 222-3, 1995), came to S. Asia relatively late, but early enough to participate in the Indus civilization, from which they acquired agriculture and the accompanying vocabulary. This scenario, if applied just to Sindh, explains why the c. 300 foreign words of the RV (in the Panjab) with their (agricultural) vocabulary are relatively free of Drav. influence.

According to the indications given above, the Dravidians apparently were just as foreign to Sindh and its agriculture as the Indo-Aryans to the Panjab. As the Northern Indus language (Para-Munda/Harappan) differs considerably from the Southern one (Meluhhan), it seems likely that the speakers of Indo-Aryan entered the Panjab and acquired local words from the Northern dialect (*śaṇa*, *lāṅgala*, *vrihi*, *godhūma*, *kaṅgu*, *Gandhāra*), and that the Dravidians entered Sindh at or about the same time and acquired such words from the southern dialect (*gōnu*, *ñāñcil*, *variñci*, *godī*, *kaṅku/kampu*). It may even be the case that the first who made horses statues at Pirak (1700 BCE) were Dravidians, not the IA Bhalānas. For the first use of horses must not necessarily be linked to speakers of an IA language.

The Drav. words for 'horse' underline this: DEDR 500 Tam. *ivuli*, Brah. (*h*)*ullī*, 1711 Tam. *kutirai*, Kan. *kudire*, Tel. *kudira*, etc., 3963 Tam. *pari* 'runner', 4780 Tam. *mā* 'animal' (horse, elephant), Tel. *māvu* 'horse, (cognates mean 'deer' etc. in other Drav. languages), cf. Nahali *māv* 'horse'. These words are quite different and independent of IA *aśva* 'horse' and various words for 'runner' (*arvant*, *vājin*, etc.), etc.

On the other hand, the technical terminology for chariots is IA and IE. It has been taken over into Drav.: *akṣa* 'axle' RV > Parji-Kolami *accu* 'axle'; *āni* RV (of unknown origin) > *āni* 'lynch pin', *ara* RV > *ār* 'spoke' (cf. Southworth 1979: 230 n. 14). Note that the earliest IIR *\*ratha* 'chariot (with two spoked wheels)' (Gening 1977, Pigott 1992, Anthony u. Vinogradov 1995, cf. Littauer u. Crouwel 1996) is found about 2000 BCE, near the Volga

(North Iran. \**Rahā* > Greek *Rhā* = Avest. *Raṇhā*, Ved. *Rasā*). The IIr word for 'chariot', however, is old enough to have resulted in the archaic compounds Ved. *rathe-ṣṭhā*, Avest. *raṭaē-ṣta-* 'chariot fighter', cf. Old Avestan *raṭi*, RV *rathī* 'chariot driver.' Dravidian has nothing of this, but words for 'wagon' or 'bullock cart'.

An early wave of Dravidian speakers might very well have preceded the IAs into Iran and S. Asia. (Note the strange absence of *Maka* in the list of "Aryan countries" in the Avestan records, such as V. 1, cf. Herodotos 3.94). A few IA loans in Proto-Drav. would settle the case, but culturally decisive words, such as for the newly introduced horse, the chariot, or other pastoral terminology do not exist. The Dravidians hardly had any previous contact with the Indo-Aryans while still in Iran. Contra Southworth (1979: 196f.), there 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 (see above, on Drav. evidence in Vedic). There are only a few questionable loans that might have come from the pre-immigration period, that is from hypothetical contact when still in Iran; these remain speculative; cf. perhaps, Ved. *garda-bha* EWA I 473, Drav. *kaḷu-tai* DEDR 1364 'donkey'. -- On the other hand, several agricultural terms in Dravidian are in a close loan word relationship with Sumerian and sometimes beyond, with Afro-Asiatic (Blažek and Boisson 1992). These include words for plough-tail, -handle, plough share, to plough, mortar, threshing floor, and to grind; this close link may point to a more western path of immigration of Proto-Drav. speakers than that of those of pre-Vedic IA (see below §15).

### §7.1. Eastern Panjab and Upper Gangetic Plains

We return now to the epicenter of post-Indus developments, the area of Eastern Panjab-Haryana-Uttar Pradesh, in other words, the lands from the Pakistani border up to Allahabad. In the early post-RV texts, its hub is the Kurukṣetra area, northwest of Delhi.

This is the realm of the middle Ṛgvedic Bharata and the late Ṛgvedic Kuru (Witzel 1997). The Bharata tribe and its successor, the new tribal union of the Kuru, represent a new wave of IA immigrants from the other side of the Indus (Vasiṣṭha RV 7, JB 3.238-9 §204), which brought new linguistic traits with them (*kuru* for older *kṛṇu*, *sarva* for *viśva*, etc., Witzel 1989). The Kuru dialect is remarkably more modern than the language of the bulk of the RV. However, RV book 10 often reads already like the next level, that of the AV and other Mantra texts of the Kuru period.

The Kuru confederation, supplanting the 50-odd Ṛgvedic clans and tribes, became the center of linguistic (Witzel 1989), religious and social (Witzel 1997b) development. They formed, together with partly IA acculturated Indus people (*ārya*-tribes such as the Anu-Druhyu, Yadu-Turvaśa) and with the new addition of Dravida speakers, a new society with a new *elite kit* (Ehret 1988). This included 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. see below). The new culture, Vedic orthopraxy and social system (with four classes) then spread eastwards into the Gangetic plains, and ultimately to Bihar.

Because of the amalgamation of the three groups (IA, Para-Munda, Drav.) we have to suppose a large degree of bilingualism and even trilingualism, and the forming of pidgins. A Vedic pidgin must have been used at home, and proper Vedic Sanskrit was learnt 'in school', at the time of initiation of boys (cf. Kuiper, A bilingual Ṛṣi, in press). While the lingua franca was a form of late/post-Ṛgvedic IA, 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.

Among the post-Ṛgvedic texts, especially the AV is full of non-IA, 'popular' words of plants, animals, demons, local deities, and the like. Their character still is, by and large, Para-Munda, with some words from the 'local' language ("X"), and with some Drav. words included; all of which is clearly visible in the increase of words with retroflexes.

The linguistic situation is reflected, among other items, in the mixture of IA and other river names in the area. The famous Sarasvatī is also called Vaiśambhāyā / Vaiśampāyā / Vibalī; these names and that of the nearby Vipās < \*vipāl/vipāž all seem to go back to a local word, \*vi-śam-paž, (Witzel 1999). However, and typically, there are no Dravidian river names in the whole Kuru area.

A hint of how Drav. influence on Vedic was exerted is contained in the name of the Śūdra. From the late RV (10.90) onwards, this designates the fourth, non-Ārya class; it was added to the three 'Ārya' classes of Brahmins, Kṣatriya (nobility) and Vaiśya ('the people') only at this time. However, Greek sources of Alexander's time still place the *Sudroi* people at the confluence of the Panjab rivers with the Indus; this may still indicate their origin in Sindh/ Baluchistan.

Drav. words first appear in Middle and Late Ṛgvedic, in RV 3, 7, and 8, especially in the Kāṇva section. Interestingly, it is Tura Kāvaṣeya, the great-grandson of the Drav.-named *Kavaṣa* 'straddle legged', a priest on the 'wrong side' in the great Bharata battle (RV 7.18)

who becomes an influential priest in the Kuru realm and who developed the new, post-Ṛgvedic (*śrauta*) rituals (Proferes 1999).

It has been stressed by Burrow (1973 : 386) that the post-Vedic texts have more Dravidian words; indeed, the evidence of Para-Munda words, too, is not diminishing but increasing during the Vedic period. This is the case right from the Mantra texts, and includes the Yajurveda Saṃhitās whose territory can be easily established (Witzel 1987, 1989, 1997) as that of the area between E. Panjab (Lahore), Allahabad and the Chambal River area (Ujjain).

A complete discussion of the c. 200 longer or shorter Vedic texts must be postponed to a separate paper (for some lists, see below). In the mean time, one can compare the word index to the AV (Whitney 1881), or Vishva Bandhu's Vedic Word Concordance (in Devanagari script), in conjunction with EWA, KEWA (and DEDR).

### §7.2. The Post-Ṛgvedic period

The new tribal union of the Kuru (and their more eastern allies, the Pañcāla), with their new social set-up and solemn rituals expanded, incorporating the surrounding tribes, eastwards into the Gangetic plains, in a partly military, partly peaceful fashion until it reached northern Bihar (Witzel 1995, 1997). The eastern tribes were at first regarded as half-barbarian (JB 1.337 §115) or '*asurya*' (demonic).

The same is seen in archaeology: late Harappan people emigrated towards the Upper Gangetic plain (the only movement of people the archaeologists allow for the whole period under discussion here, Shaffer 1995: 139, cf. Allchin 1995: 33-35), a fact reflected in the Vedic texts as well. The emigration was possible due to a new type of agriculture, permitting cultivation of rice during the monsoon (Kenoyer 1998: 163) as well as wheat and barley in winter, resulting in a food surplus. The settlement at first occurred along the river banks, (Witzel 1987, 1995), in half-nomadic treks (*grāma*, Rau 1997). This is reflected by the Painted Gray Ware culture, with their clear elite pottery whose regional motifs indicate the split into western Kuru and more eastern Pañcāla, something that is also seen in the Vedic dialects they use (Witzel 1989).

Not everybody is included: The non-IA *Kīkaṭa* (3.53) or the *Paṇi* are clearly described as foreigners (late hymn 6.45.31), and even later, in the Mantra and YV Saṃhitā period, the *Niṣāda* in the Chambal area (MS 2.9.5 etc.) and other *dasyu* 'enemies' (JB, Witzel 1997b: n.161, 163, 278); in RV 10.61.8 as well the South (i.e. the area south of Kurukṣetra) still is the land to banish someone.

As has already been indicated, the features of the Ṛgvedic substrate language are also found in post-Ṛgvedic texts that were composed further east in the Kurukṣetra and in western Gangetic plains, as well as in the Chambal area. These words are not just the same as found in the RV, but there are many new ones.

In the Mantra period, starting with YV (MS, KS, TS) and AV/PS, we can clearly distinguish all three linguistic elements:

- Indo-Aryan with some already incorporated north-western elements such as Nuristani *kāca* 'shining piece of jewelry' or Burushaski *kilāy* ~ RV *kilāla*, *ṣon* ~ RV *kāṇa*, *bus* ~ RV *busa*, etc.;
- The Indus substrate (Para-Munda), that also is found in the Ganges area (next to some elements of language 'X'), such as RV *kuśika*, *karañja*, *kaṅkata*, *śiṃśapā*, *śiṃśumāra*,



*puṣkara*, *puṣya*, especially the words with prefix *Cər* (*pər/kər/sər-*), *kar-koṭa-ka* RVKh ~ *śar-koṭa* AV, *tila* AV: *jar-tila* KS, *kalmaśa* MS, KS, *kal-māśa* PS, *kul-māśa* Up. : *māśa* AV, with the *-ta*, *-śā/śa* suffixes, and with *-nḍ-*: *ka-maṇḍalu* : *maṇḍa-la*, *kaṇṭha?* PS, etc.

• The Middle and Late Ṛgvedic Drav. element also is found in the Ganges area: *godhūma* AV (Hindi *gehū* etc., Kusunda *gabun*), *kuṇapa* AV, *kurkura* AV, *cūda* ŚB, *coḍa* TS, *eḍaka* JB, *arka* ŚB, *bilva* AV 20 (Kuiper 1991:66), *-nīra-* ŚB, etc.

In short, the upper class IA language (of the Vedic priests) used in the upper Gangetic plains contains the same substrate elements as seen in the late Ṛgvedic period of the Panjab. However, due to the increasing stratification of society and increasing specialization among occupations, many words from the sphere of the artisans and from technology were added; furthermore many names of persons, localities and rivers.

Their affiliation can still be ascertained to some extent. With regards to agriculture, Kuiper's RV list (Kuiper 1991: 8, 21, 96, see already Kuiper 1955) contains quite a number of such terms (*kīnāśa*, *lāngala*, *bija*, etc.) Especially among the artisans there is an increasing number of non-IA designations; many of them first appear in the Horse sacrifice (Aśvamedha ritual) (MS *kevarṭa*, *kaivarta* TB).<sup>15</sup> Some of them are, in line with the increasing specialization, new Indo-Aryan formations (*anucara* 'servant', *grāma-ṇī* 'leader of a trek, wagon train' etc.), but especially those of fishermen (*kevarṭa/kaivarta*, *dāśa*, *dhīvan*, *daivara*, *puñjiṣṭha*, *pauñjiṣṭha*, *bainda*, *maināla*) are non-IA (often until today). Furthermore, non-IA specialists are: musicians (*talava* 'musician', *ādambara-āghāta* 'drum beater', *dundubhy-āghāta* 'drum beater' (cf. *dundubhi* RV), *vīṇā-gāthin* 'lute player', *vīṇā-vāda* 'lute player', cf. *vīṇā* 'lute' KS (EWA II 568), artisans (*kaṇṭakī-kārī* worker in thorns', *bidala-kārī* 'female splitter of bamboo', also *kulāla* 'potter', and the *pālāgala* 'messenger' (cf. *pālāgali* 'fourth wife of a chieftain'), *gaṇaka* 'astrologer' (cf. *gaṇa* 'troop, number' RV) and 'money lender' (*kusīdin*, *kusīda* KS).

Such words come up not only in the eastern parts of North India (Bihar, area of VS/ŚB) but also everywhere from the Panjab (RV) and the Delhi area (MS, KS) eastwards,

<sup>15</sup> Details: *kīnāśa* 'plough man' EWA: 'non-IE'; *kīnāra* only RV 10.106.10; -- the following words all mean 'fisher' *kevarṭa/kaivarta* VS/TB; Pali, Pkt. *kevaṭṭa*, \**kevēṭa*, CDIAL 3469 and add., 3479; Drav. according to Burrow, KEWA I 566, DEDR 1252 Tam. *kayal* 'carp', Mal. *kayal* 'a fish', etc.; *kai-* in *kevarṭa*; -- *dāśa* VS, *daśera* lex. CDIAL 6314 a *Jāt* tribe: *ḍahā*; -- *daivara* VS, see *dhī*, CDIAL add. 6819 NIA, Kuiper, KEWA II 105 ~ *tivara* (lex.) = tribal name? -- *puñjiṣṭha* also 'bird catcher?', MS, VS, *pauñjiṣṭha* AV; no NIA etym.; -- *bainda* ~ *Srbinda*, Kuiper 1991, EWA; -- *maināla* < Drav. *mīna* 'fish'; --- *śauśkala* ~ *śuśka* 'dried up'? -- Further: *talava* 'musician' VS ~ *taḍ* Epic 'to play a musical instrument'? Kuiper ZII 8, 1931, 251; -- *ādambara-ghāta* 'drummer' VS, ā- ŚB; Kuiper 1948: 85f. from Proto-Munda, *dundubhy-āghāta* 'drummer' (RV), ŚB EWA: onomatopoeitic, Kuiper 1948: 84 Munda; *vīṇā-gāthin* 'lute player', also in Iran?, see EWA, Mayrhofer 1968, CDIAL 12048; *vīṇā-vāda* 'ditto'; -- *pālāgala* 'messenger' ŚB, *-kālī* ŚS. no NIA continuants; -- *kaṇṭakī-kārī* 'worker in thorns' VS; *kaṇṭaka* 'thorn' ŚB, Iran?, Greek *akantha?* -- *bidala-kārī* 'basket maker' VS, EWA "not clear", but cf. DEDR 5432 *viḷ* 'to split'; -- *sirīn* 'weaver?' only RV 10.71.9 (Ved. Ind. 585-6); -- *gaṇaka* 'astrologer' VS: RV, *gaṇa*, \**gṛna*, CDIAL 3993 and add.; Greek *ageirō* 'collect'; Kuiper 1948: 54 Munda; -- *kusīdin* 'money lender' ŚB, *kusīda* KS, TS; Pali *kusīta* 'lazy', etym.? *ku+sad* > Pali *ko-sajja?* -- *parṇaka?* a tribal name? VS "Bhilla" in later commentary, EWA ~ *paṇi?* -- *paulkasa?* VS a mixed tribe, Kuiper 1948: 54ff. -- Indo-Iran.: *malāga* 'washer man' < AV, *mala*: IE \**mel*; -- *upala-prakṣinī* from IA *upalā* 'mill stone' TS: *kulāla* 'potter' MS, KS, VS; EWA ~ RV *kula* 'hole, hollow', in *mahākula*, Pashai *kōlāla* 'potter' CDIAL 3341; -- *kṛṣī-vala* 'agriculturist' RV, *a-*, AV *kārṣīvaṇa* : suffix variation!; -- *vañij* RV, *vāñija* KS 'trader' < *van-ij* 'winning goods' according to EWA, Mayrhofer 1968.

e.g. *kīnāśa* 'plough man' RV, *gaṇa* 'troop' RV, *duṇḍubhi* 'drum' RV, *vīṇā* 'lute' KS, *kuśīda* 'money lending' KS. The newly attested words have the same 'foreign' grammatical formations as seen in the RV: prefixes (*ke-/kai-*, *duṇ-dubhī?*), retroflexes (*āḍambara*, *kaṇṭakī-*), initial *b-* (*bidala*), suffix *-āla* (*pal-āla*, *main-āla*, cf. Oberlies 1994: 341).

Similar data could be supplied for the spheres of material culture and the surrounding nature: agriculture and domesticated plants, local animals and plants, many items of food, illnesses and poisons, implements and utensils, and ornaments; this would lead to far afield in present context (see the lists in MacDonell-Keith, *Vedic Index*, Delhi 1967 [1912] 517-92). For more examples, one can consult Mayrhofer, *EWA* and for non-IA details especially *KEWA*; these may serve, in connection with CDIAL, DEDR, Kuiper 1948, 1955, 1991 and Pinnow 1959 as a first orientation.

### §7.3. The Para-Munda substrate in Post-Ṛgvedic.

Prefixes with *ka-* are found in the AV, YV and the Brāhmaṇas (here follow only a few proposals for etymologies; it is to be expected that not *all* of the following words can be divided in the way proposed below; ultimately this depends on a fitting etymology):

- *kapaṭu* 'mushroom' AV, PS, cf. Sora *pud-ən*, Sant. *o'd* etc. (Pinnow 1959: 121 §237;
- *kapāla* 'potsherd, skull' AV;
- *kapiñjala* 'partridge' PS;
- *kapola* 'cheek' RVKh, cf. Sant. *puṭi* 'to swell', Kharia *poṭki* 'to sprout' etc. (Pinnow 1959: 173 §378, Kuiper 1948: 148) ~ *puṭa* 'bundle, bag' MS, BŚS;
- *kaphauḍa/kaphoḍa* 'clavicle, elbow?' AV, see Kuiper 1948: 44;
- *kamaṇḍalu* 'water jar' KS cf. *maṇḍala* 'circle' etc.;
- *karīra* 'bamboo shoot' MS, KS;
- *karīṣ-in* 'having dung' AV;
- *karuma* 'epithet of certain spirits' AV;
- *karūkara* 'vertebra of the neck and spine' AV;
- *kalāp-in* 'having a bundle of arrows (or 'peacock feathers')' ŚS;
- *kaliṅga* 'the name of Orissa' AB, cf. Skt. *tri-liṅga* (mod. Telingana), etc., see Kuiper 1948: 45;
- *kavaca* 'armor' PS (but see above, Zvelebil's no. 13);
- *kaśambhūka* 'name of a mythical being' Suparṇākhyāna;
- *kaśipu* 'cushion' AV;
- *kaśiti* 'name of a man' JB;
- *kaśoka* 'name of certain demons' AV;
- *kaśmaśa?* "? ", 'confusion, agitation?' AV, see Kuiper 1948: 39;
- *kaśāya* 'astringent sap, red' ŚB;
- *kaśkaśa?* 'a certain damaging worm' AV;
- *kasarṇīla* 'a certain snake' AV, cf. *sarṇīka* 'water?' TS/*sṛḍīka* 'water?' MS (cf. *sṛḍaku* 'snake, lizard' ?);
- *kasāmbu* 'name of an extract derived from the *devadāru* a tree?' AV, etc.;
- *kastūpa* 'hair tuft', *kastūpa-stopinī* 'woman wearing a hair tuft' PS, cf. *stupa* 'hair tuft, top knot' KS / *stuka* 'hair tuft' RV;

• *kahoḍa* 'name of a teacher, belonging to the Kauṣītaki clan' ŚB, JB.

With 'double prefix' *Cər-/Cəl-* there are the following words in which the many variants of the prefix in *kər-* stand out:

- *karkandhu* 'the tree *Zizyphus jujuba*' MS, KS;
- *karkī?* 'white (cow)' AV;
- *karkoṭa-ka* 'name of a snake demon, *Nāga*' RVKh ~ *śarkoṭa* 'name of a snake demon' AV, PS, cf. Mundari *kar-kom* (Pinnow 1959: 341 §483d), Kuiper 1991: 41, 44, 1948:121, Bur. *yarqas* 'lizard';
- *kardama* 'dirt, mud' KS, cf. Munda *ko-dil*, *ə-dil* 'dirty' (Pinnow 1959: 87 §101);
- *karpāsa* 'cotton shrub' Suśruta, *kārpāsa* 'made of cotton' ŚS;
- *karśapha* 'name of certain demons' AV, PS : *śapha* 'hoof?' RV (note that *śapha* has a clear IE etymology, EWA II 608), cf. *Śaphāla* 'a tribe' BŚS;
- *garmut* 'wild beans' TS, *gārmuta* 'wild beans' MS (Kuiper 1948: 146, CDIAL 4063: Sindhi *gamu* 'a sort of grass');
- *kalkuṣī* 'a bone of the lower arm' PS; ŚB,
- *kalmali* 'shimmering (of stars)? AV;
- *kalmāsa* 'spotted, variegated' MS, KS, *kalmāsa-grīva* 'with spotted neck' ŚS, PS;
- *kārṣmarya* 'the tree *Gmelia arborea*' KS;
- *kharjūra* 'date palm' KS;
- *gulma?* 'shrub, bush' Samh.;
- *jar-tila* 'wild sesame' KS : *tila* 'sesame' AV;
- *jarvara* 'name of a person' PB;
- *jalāsa* 'an ingredient used in medicine, healing?' PS (or *-āsa* suffix, Kuiper 1991: 26);
- *palala* 'crushed sesame' Sū.,
- *palālī* 'straw' AV;
- *palāva* 'chaff' AV;
- *palijaka* 'a certain demon' AV;
- *barkara* 'lam' ŚS;
- *barbara* 'having curly hair' KS;
- *barhiṇa* 'peacock' ĀpDhS;
- *bharūji* 'a certain noxious animal' AV;
- *marīca* 'pepper corn' ĀpDhS;
- *markata* 'monkey' KS
- *markataka* 'a type of grain' ĀpŚS, CDIAL 9884, Shina *makari* 'large millet', Bihari *makrā* 'the grass *Eleusine aegyptica*'; cf. CDIAL 9879 *markaka* lex. > NIA *makāi*, *makai* 'maize'
- *śarkara* 'sand, pebbles' AV, cf. Bur. *ḡoro* 'stones' ?;
- *śarkoṭa* 'name of a snake demon, *Nāga*' AV, PS (see above *karkoṭa*);
- *sardigr̥di* 'part of the female sexual organs' TS.

Double prefix *Cən-/Cəm-*:

- *kaṅkūṣa* 'part of the head' AV, PS ~ *śaṅku*;
- *kaṅṭha?* 'neck' PS, (*saha*)-*kaṅṭh-* AV, cf. Kharia *konko*, Khmer *ko*, Mon *ka* "possibly old compound", Pinnow 1959: 132 §276;

- *kāṇḍa?* 'section, piece, section of bamboo, grass' AV, cf. Kharia *koṇḍen* 'bamboo', (Pinnaw 1959: 132 §275);
- *kāṇḍūy-?* 'to scratch' KS;
- *kandhara* 'neck' Up., cf. *kaṇṭha*;
- *kambala* AV 'woolen blanket, clothes' ~ *śambara?*;
- *kambūka* AV 'chaff' ~ *śambūka*;
- *kamboja* 'name of a people in SE Afghanistan' PS, cf. Greek *Ambautai*;
- *kāmpīla-* 'name of a particular dress, skirt' KS;
- *jāmbīla* 'saliva' KS, TS;
- *taṇḍula* 'rice grain, husked rice' AV;
- *talāśa?* 'a particular tree' AV (if not with *-āśa* suffix);
- *parūśaka* 'a type of plant, *Grewia asiatica*' ŚS;
- *palāṇḍu* 'onion' ĀpDhS;
- *palāśa* 'leaf' TB (if not with *-āśa* suffix);
- *palijaka* 'a certain demon' AV;
- *palpūlana* 'lye, washing water' AV (if not onomatopoeitic);
- *palvala?* 'pool, small pond' Sū;
- *pālāgala* 'messenger, runner' ŚB, *-ī* 'fourth wife of a chieftain' ŚB;
- *barśa?* 'knot' KS,
- *barśva?* 'gums, alveolus' KS;
- *balāśa* 'a certain illness' PS (cf. *kilāśa* 'leprosy');
- *balkasa* 'sediment, residue' ŚB;
- *balbaja* 'a type of grass, Eleusine indica' RV;
- *balbūtha* 'name of person' RV;
- *bhalānas* 'name of a tribe' (of the Bolan Pass area ?) RV.

From the post-Ṛgvedic materials come words with other prefixes in *Cār-* and with other vowels, etc.:

- *kirika* YV, *girika* MS 'sparkling';
- *kirmira* 'variegated' VS, etc.;
- *kul-māśa* 'an inferior type of grain' Up., cf. (*kal*)*māśa* 'spotted, variegated' AV;
- *ku-ṭaru* 'rooster' YV, etc.;
- *ṣṛḍāku* 'lizard', etc., lex., *ṣṛḍāku/-gu* MS, *ṣṛḍara* 'snake', etc. Mayrh. ZDMG 110, 6189 Munda prefix *ṣṛ-* + *da'k* 'water', see KEWA s.v. *ṣṛḍāku*, etc.;
- *kaśmaśa?* '?' 'confusion' AV, Kuiper 1948: 39;
- *kaśkaśa?* 'a certain type of noxious worm' AV;
- *jāśkamada* 'a certain animal' AV;
- *maśnāra* 'name of an area' AB;
- *masūra?* 'lentil' KS, *masura* TS;
- *prakubrata* '?' ŚB, *prakudrata* '?' ŚBK,
- *pramota* '?' 'deaf, blind?' AV etc.;
- *tilvaka* 'name of a tree, *Symplocos racemosa*' ŚB, *tailvaka* 'belonging, stemming from *tilvaka*' MS, etc.;
- *tumbara* 'a certain tree, *Disopyros embryopteris?*' Kauś etc.

- Further Vedic words which are suspected of a Para-Munda origin are, among others:
- *me-khala* 'girdle' AV: *śṛ-ñ-khala* 'chain, fetters' Skt.;
  - *khaḍga* 'rhinoceros' MS, EWA 443, cf. N.Pers. *karka-dān*, Arab. *karkaddan*, Aelianus *kartázōnos* (\**kargazōnos*) 'Indian rhinoceros', cf. Kuiper 1948: 136 sqq.;
  - *karta/garta* 'hollow'/'seat' to be compared with Kharia *garha* 'river', Mundari *gaḍa*, *garā* 'pit, trench, grave, water course, stream, river'; Sant. *gaḍa* 'hollow, pit, excavation, trench, river'; etc. (Pinnow 1959: 351f. §498);
  - *tittira* 'partridge' KS, MS cf. Korku *titid*, Santali *sengel titi* 'Guinea fowl': Kharia *khonthē'd*, Sora *on-'tid-ən* (Pinnow 344 §488a); probably also:
  - *musala* 'pestle' AV;
  - *jala?* RVKh, PS;
  - *dhūkṣṇa/dhluḥṣṇa/dhḷkṣṇa* 'a bird' PS ~ *dhvāṅkṣa* 'crow' AV, *dhūṅkṣṇā* 'white crow' TS; *jhaṣa* ŚB : *jaṣa* AV, TS : *caṣa* 'a large fish' VādhB;
  - *drumbhūli* MS / *dālbhuṣi* KS / class. *dambholi* 'bow of Indra' see Kuiper 1991: 26 (cf. p. 18, 47, 61, 75).

#### Para-Munda suffixes.

In order to characterize the substrate, certain typical suffixes can be used. Kuiper (1991: 45 sqq.) has isolated the following in the substrate of the RV: *-āla*, *-āṣa*, *-īṣa*, *-ūṣa*/*-āṣa*, *-īṣā*, *-ūṣā*, *-ṭa*, *-nas*, *-ya*, *-ra*, *-śa/ṣa*, *-ha*. Among the suffixes are to be underlined in this context are those often found in personal and tribal names, in *-ṭa* (*Kīkaṭa* 'a tribe', *kr̥pīṭa* 'brush', *birīṭa* 'crowd', *kevaṭa* 'hollow' RV / *avaṭa* 'hollow' SV), and the ones in *-āla*/*-āra* (*kīlāla* 'biestings', *caṣāla* 'snout'; *maināla* 'fisher' VS, cf. IA *karmāra* RV 'smith'; *Gandhāri* 'a tribe' RV, *Gandhāra* 'a country in N. Pakistan', *Abhisāra* 'a region north of Gandhāra' etc., cf. Witzel 1999).

Such suffixes also appear in post-R̥gvedic time in the texts of the Mantra period and in the Yajurveda-Samhitās, e.g. *kalmāṣa* 'spotted' VS, TS; *niṣkāṣa* 'scraping' MS, KS; *yevāṣa* AV, *evaṣa* MS 4.8.1:107:16, *yavāṣa* 'manna plant' KS 30.1, KpS 46.6 (*vṛṣaś ca yavāṣaś ca*); *r̥jīṣa* a name of Indra, RV, 'residue of Soma' AV; *uṣṇīṣa* 'turban' AV; *karīṣa*[-*ja*] PS, 'dung', *karīṣin* AV, *karīṣa* ŚB, (cf. the frequent *purīṣa* 'dung'); cf. also *tūṣa* 'border of garment' KS; later also: *palāṣa* 'leaf' TB, ŚB, *ni-palāṣa* ŚB, *śirīṣa* 'Acacia sirissa' ṢaḍvB, etc.; cf. also *jhaṣa* 'a certain large fish', ŚB *jaṣa* AV, TS, *caṣa* VādhB.

#### Para-Mundas in Kurukṣetra and in the Gangetic plains.

The words mentioned above clearly show that also in post-R̥gvedic, i.e., in the Mantra texts (AV, SV, RVKh, YV), in Yajurveda Prose, and in the Brāhmaṇas, such Para-Munda words can still appear for the first time. Therefore, they had either already existed in Vedic colloquial speech or they entered Brahmanical High Vedic at that particular point in time from the sphere of village life or of the artisans. The area of the early post-R̥gvedic texts (Mantra texts, YV Prose) can be localized fairly well (Witzel 1987, 1989): it contains Kurukṣetra (i.e. more or less, modern Haryana) and the western Ganga-Yamunā-*doāb* (i.e. the Gangetic plains of western Uttar Pradesh).

In these areas, where no modern groups of Munda speakers survive, the same R̥gvedic substrate with its typical prefixes can be found. That means Haryana and Uttar Pradesh once had a Para-Munda population that was acculturated by the Indo-Aryans.

If the late Vedic texts (such as the Jaiminīya Br. and Śatapatha-Br.) are added, the area in question is further enlarged to include the regions south of the Ganges and east of Uttar Pradesh. Here, new Munda words appear as well; however, these regions include those where even today Munda languages are spoken.

In short, a strong Austro-Asiatic substrate is found both in the early Panjab (RV, c. 1500 BC) as well as later on in the Ganges valley (YV Saṃhitās, Brāhmaṇas, c. 1200 v. - 500 BC.), a fact that can also be shown in the names prevailing in these areas (Witzel 1999).

As examples, I mention the river names Gaṅgā (popular etymology of Munda *ga(n)d*), Gaṇḍak-ī (see below), Narma-dā, and tribal names such as Marata, Vibhindu (and Vibhindukīya, cf. Nar-ka-vinda PS 12.2.3, Sṛ-bindu RV Kuiper 1991: 40-43, 1997, Ku-suru-bindu TS, TB, ŚB, Ku-sur-bindu JB, Baidu VS, cf. Munda *bid* 'insert, plant, sow', Pinnow 1959: 143 §285), Śabara (\*Šqawar, cf. Pinnow 1959: 154 §31; rather from \*K'awar/Śawar), Puṇḍra, Aṅga/Vaṅga (cf. also Gaṅgā?; further: Pra-vaṅga), Kaliṅga (cf. Teliṅga/Triliṅga, see S. Lévy in Bagchi 1929: 100, cf. Shafer 1954: 14, 122 as Tib.-Burm.; Kuiper 1948: 45 compares *kuliṅga* 'fork-tailed shrike' Mbh., and \*liṅ in Munda, Khasi, Mon, Khmer, Malay); Ikṣvāku (RV, emigration from the Panjab eastwards, Witzel 1997b: 307 sqq., 321, 1989: 237), Niṣada/\*Niṣadha/Naiṣadha, Mucīpa/Mūtība/Muvīpa, Magadha (cf. Pra-maganda), Śaphāla cf. Śavasa, Vasa etc.

However the truly eastern words (Uttar Pradesh, Bihar) are, next to some remnants of language "X", of Munda nature: there are many personal and place names (Witzel 1999), e.g. that of the river Gaṇḍak(ī), or even that of the Ganges, with popular etymology: Gaṅgā, a sort of intensive formation of *gam* 'to go' (if not modeled after the tribal names Aṅga, Vaṅga). Pinnow (1953-4) has pointed out many river names, from the Gaṇḍakī to the Narma-dā which contain the Munda element -\*da', \*-da'k 'water' (Pinnow 1959: 69), for gaṇḍa(kī) cf. Santali gāḍa, Ho gaḍa 'river' (Pinnow 1954: 3).

The Gaṇḍakī is not attested in Vedic, and is referred to as *Sadānīra* 'always having water'. Apart from the Epic, it appears in local context, the early Licchavi inscription (464 CE), Sanskritized as *Gaṇḍakī* and in other Skt. texts: *Kāla-Gaṇḍikā*, *Gaṇḍārikā*, *Apara-*, *Pūrva-*; the shorter version, *Gaṇḍī*, appears from the Epic onwards, and several times early on in Nepal as *Gaṇḍī-(gulma-viṣaya)* (998, 1092, 1165 CE, see Witzel 1993). The Gaṇḍaka appear as people in Mbh. as well.

Further, tribal names such as *Pulinda*/Pali *Buli*, Pali *Moriya* (from Skt. *mayūra* 'peacock') and also *Mara-ṭa* (PS), from Munda *mara* 'peacock', *Kunti* from Munda *kon-ti'd* 'bird' (note that Munda *kom* is a children's word!), cf. RV *śa-kunti*, Epic *Śa-kuntalā*, etc. (contrast the IA *Matsya* 'fish' RV, a tribe just west of the Kunti), *Mūtība* (var. *Mūcīpa*), *Śabara* (mod. *Saora?*), *Puṇḍra* (Bengal), the *Aṅga*, at the bend of the Ganges, and the neighboring *Vaṅga* (Bengal). The prefix change in *Aṅga* (AV) / *Vaṅga* (AB) is indicative of a Munda formation (Kuiper 1991: 43). Mundas may also have lived in the hills and valleys of the Sub-Himalayas, for example in the Kathmandu Valley (see below, Witzel 1993).

Other typical words of the Gangetic plains are, from west to east: *sardigr̥di* 'part of female sexual organs' TS, *palāśa* 'leaf' TB, *palāṇḍu* 'onion' ĀpDhS, *tumbara* 'a certain tree'

KauśS, *kaṣṭi* 'name of a man' JB, *kirmira* 'variegated' VS, *kaṣāya* 'astringent sap, red' ŚB, *pra-kudrata* '?' ŚBK, *pra-kubrata* '?' ŚB, *ka-hoḍa* 'name of a man' ŚB, JB, *kul-māṣa* 'an inferior type of grain' Up. etc. Especially informative for regional dialect features of the substrate, from W. to E.: *jaṣa* AV, TS : *caṣa* VādhB : *jhaṣa* ŚB 'a certain large fish'.

The Ṛgvedic substrate thus has the same grammatical structure as the words in the Yajurveda-Samhitās and the Brāhmaṇas that newly appear from the substrates of the Kurukṣetra (Haryana) and Ganges regions (*doāb*, Uttar Pradesh). It is of great importance that we can detect the *same Indus substrate as found in the RV*. In other words, *the Ṛgvedic Panjab as well as the post-Ṛgvedic Gangetic Plain were largely settled by speakers of Para-Munda* (including remnants of Masica's 'Language X'). They had been joined, in the early Ṛgvedic period, by speakers of Indo-Aryan and, in the later Ṛgvedic period, by those of early Dravidian (see above).

### Dravidian

In the new IA speaking, culturally Vedic "eastern territories" of the Gangetic plains some Drav. words occur for the first time in literature, e.g. *nīr* 'water' in the name of the eastern river *Sadānīrā*, the modern Gaṇḍak (Witzel 1987), or the verb 'to speak in barbaric fashion', *mleccha-ti*. However Drav. *nīr* is not found in the neighboring N. Drav. languages (Malto, Kurukh), but is only found in Baluchistan (Brahui *dīr*, DEDR 3690). This may be accidental, but it may also indicate that Brahmanical educated speech of the Kuru with their IA-Drav.-Munda symbiosis and acculturation had incorporated some Drav. words which appear only now in the texts. The word *mlecch* has been discussed above. Its appearance in the eastern context is not surprising. From the point of view of the Brahmins, the easterners are 'foreigners', *mleccha*. The word may at first have designated only the southern (Sindh) foreigners, and later on all others. These central and eastern North Indian territories, however, have no Dravidian names; the river names belong to other substrates.

A study of present and medieval north Indian places names has not been undertaken in earnest. We will have to account for such names as that of the town of Goṇḍ(ā) in Uttar Pradesh, some 180 km north of Allahabad. The name Goṇḍ appears nowadays only on the Central Indian Vindhya mountains, and is not known in U.P. from medieval and classical sources. (For some supposedly Drav. river names such as *Sadā-nīrā* from Drav. *nīr* 'water' see above, and for the *Varaṇāvati* at Benares, see Witzel 1999.)

There are, as always, wrong leads, such as the river name *Kankai* in the Eastern Nepal Terai, which looks like the Tamil form of the name *Gaṅgā* (Witzel 1993); there are, however, no traces of an earlier S. Drav. occupation in the area. The Dravidian Kurukh living in the Terai now have recently been imported as laborers from Central India (K.H. Gordon 1976) where they are known as Kurukh or Oraon.

For a different view of early Dravidian settlements in N. India, see R. Shafer 1954, Parpola 1994: 168, and Burrow 1973: 386. Burrow points to the fact that most of the Drav. loan words are found in post-RV texts and concludes: "the influence took place in the central Gangetic plain and the classical Madhyadeśa." Therefore, "the pre-Aryan population of this area contained a considerable element of Dravidian speakers". If that had been the case, we would expect some Drav. river names in the Gangetic plains. However, only Munda (and Tib.-Burm.) names are found (Witzel 1999).

### §8. Substrates of the Lower Gangetic Plains

Next to the Mundas, there must have been speakers of other languages, such as Tibeto-Burmese, who have left us names such as *Kosala*, *Kauśikī* (mod. Kosi), perhaps also *Kāśi* and *Kauśāmbi* (mod. *Kosam*) (from Himalayan *khu*, *ku*, see Witzel 1993). In IA they also have left such words as the designations for cooked rice IA \**cāmala* and probably also PS *śāli* 'rice'.

In Uttar Pradesh and North Bihar (attested in Middle and Late Vedic texts, c. 1200-500 BCE) another apparent substrate appears in which the 'foreign' words do not have the typical Para-Munda structure, with the common prefixes, as described above (§4.2). Masica (1979) called this unknown substrate "language X". He had traced it in agricultural terms in Hindi that could not be identified as IA, Dravidian or Munda (or as late loans from Persian, S.E. Asia, etc.). Surprisingly some 30% of the terms are of unknown, language "X" origin, and only 9.5% of the terms are from Drav., something that does not point to the identity of the Indus people with a Drav. speaking population.

However, only 5.7% of these terms are directly derived from Munda. Obviously, the pre-IA population of the Gangetic plains had an extensive agricultural vocabulary that was taken over into all subsequent languages. F.B.J. Kuiper has pointed out already in 1955: 137-9 (again in 1991: 1) that many agricultural terms in the RV neither stem from Drav. nor from Munda but from "an unknown third language" (cf. Zide & Zide 1973: 15). This stratum should be *below* that of Para-Munda which is the *active* language in the middle and late Vedic texts.

Again, it has been Kuiper who has pointed the way when he noted that certain 'foreign' words in the Vedic substrate appear with geminate consonants and that these are replaced in 'proper' Vedic by two dissimilar consonants (1991: 67). Examples include: *pippala* 'fig' RV (1.164.20,22; 5.54.12, *su-* 7.101.5) : *piṣpala* AV (in Mss.) 9.9.20,21; 6.109.1,2; *su-piṣpala* MS 1.2.2:11.7, *guggulu* 'bdellion' AV, PS : *gulgulu* KS, TS, *kakkaṭa* PS 20.51.6, KSAśv. : *katkaṭa* 'a bird' TS, cf. Pali *kakkaṭa* 'a large deer'. Kuiper adds many other cases of Vedic words that can be explained on the basis of words attested later on.

In RV geminates also occur in 'onomatopoeic' words: *akhhkhalī-kṛ* 'to speak haltingly' or 'in syllables?', apparently not attested again in IA until, now Nahali *akkal-* (*kāyni*) '(to cry) loudly in anguish' MT II 17, L 33 (*kāyni* < Skt. *kathayati* 'to tell' CDIAL 2703, cf. 38) MT II 17; cf. also *jañjan-* RV 8.43.8 etc., *ciccika* 10.146.2 'a bird?', and cf. also *aśvattha* 1.135.8 : *aśvatha* a personal name, a tree, 6.47.24, with unclear etymology, (Kuiper 1991: 61, 68).

Post-RV, new are: *hikkā* PS 4.21.2, *kakkaṭa* PS 20.51.6 (MS *kakuṭha*, TS *katkaṭa* 'a type of bird'), KSAśv in YV: *kikkita* KS, TS, *kiṭkita* *kṛ* 'call to attract birds' JB, *kukkuṭa* 'rooster' VS, *pilippilā* 'slippery' TS 7.4.18.1, MS, VS; cf. also TS *akkhidant*, *prakkhidant* TS 4.5.9.2, *ājya* 5.2.7.3.

Especially interesting is the early gemination \**dr* > *ll*: *kṣullaka* AV 2.32.5, TS 2.3.9.3 *kṣullaka*, < *kṣudra* 'small' (a children's word?); later on, among others, *bhalla-akṣa* ChU4.1.2, *bhalla* Br., MBh (with variants *phala*, *phalla*! EWA s.v.); JB *Malla* 'a tribe' (in the Indian desert, Rajasthan; cf. DEDR 4730), etc.

Though certain geminates, especially in word formation and flexion (-*tt-*, -*dd-*, -*nn-* etc.), are allowed and common, they hardly ever appear in the stem of a word (Sandhi cases



such as *anna*, *sanna* etc. of course excepted). Until the late Brāhmaṇa texts, other geminates, especially *bb*, *dd*, *gg*, *jj*, *mm*, *ll*, but also *kk*, *pp*, etc., are studiously avoided, except in the few loan words mentioned above (*pippala* 'fig', *gulgulu* 'bdellion', *katkaṭa* 'a bird' etc. 1991: 67 sqq.).

It will be readily seen that Kuiper's seminal observation reflects a tendency that can be observed throughout the Vedic texts. Geminates, especially the mediae, apparently were regarded, with the exception of a few inherited forms such as *majj* 'to dive under', as 'foreign' or 'barbaric'. They did not agree with the contemporary Vedic (and even my own) *Sprachgefühl*.

However, starting with Epic Sanskrit, forms such as *galla* 'cheek', *malla* 'wrestler', *palla* 'large granary', *bhallūka* 'bear' (CDIAL 9415, cf. Nahali *bologo*, MT II: 41, III, 48, but note Marathi etc. *bhālūk*; -- Nahali *bologo cikin* 'caterpillar' MT II: 21 would be 'bear insect') are normal and very common (however, *-mm-*, perhaps regarded as Drav.(?) remains rare); such words, in part derive from normal MIA developments, in part from the substrate.

This tendency can be sustained by materials from various other sources. In the language 'X' only a few of Masica's agricultural substrate words that do not have a clear etymology (1969: 135) contain such geminates: Hindi *kaith* < Skt. *kapittha* 'a tree, Feronia elephantum, wood apple' CDIAL 2749 (Mbh), *pipli/piplā* < *pippala* (RV), *roṭī* < \**rottā*, *roṭika* 'bread' 10837 (Bhpr.); *karela* < *karella/karavella* 'a gourd, Momordica charantia' 3061, *khāl* < *khalla* 'leather' 3838-9 (Suśr.); to these one can add the unattested, reconstructed OIA forms (Turner, CDIAL, see Masica 1969: 136): \**alla* 'a tree or plant' (Morinda citrifolia') CDIAL 725, \**uḍidda* 'a pulse' 1693, \**carassa* 'raw hide' 4688, \**chācchi* 'buttermilk' 5012, \**bājara* 'millet' (see, however, OIA \**bājara*, 9201 *bājara* HŚS: *varjari!*), \**balilla* 'ox' 9175, \**maṭṭara* 'pea' 9724, \**suppāra* 'areca nut' 13482, \**sūjji/sōjji* 'coarse white meal' 13552. However, these words have come into NIA via MIA, and that their geminates may go back to a consonant cluster without geminates (see below, on Turner's reconstructs).

All of these tendencies are reconfirmed by what we can discern in the other substrate languages. While there still are but a few cases in the northwest, the substrates located further east and south all have such geminates, (for details on these languages see §8). (Incidentally, the northwest has retained the original, non-geminate consonant groups, such as *-Cr-*, to this day, cf. Ved. *bhrātā* 'brother' > Khovar *bhrar*, Balkan Gypsy *phral*, W. Panj. *bhrā*, E. Panj. *bh(a)rā*: Hindi *bhāi*, etc.).

In the unstudied substrate of the Kathmandu Valley (inscriptions, 467-750 CE, see below), geminates are found in the following place names: *gamme*, *gullatamga*, *gollam*, *jajje-*, *dommāna*, *dankhuṭṭā-*, *bemmā*, cf. also *bhumbhukkikā* (onomat. with double consonant: < \**bhumbhum-ki-kā?*); cf. also village names such as *joñjon-diñ*, *tuñ-catcatu*, *thuṃtuṃ-rī*, *dandañ-(gum)*.

In the substrate of modern Tharu which is spoken in the swampy lowlands of the foothills of Nepal and U.P.: e.g. *gēṭṭī*, *ghaṭṭī*, *ṭippā* (?), *ubbā*; cf. also 'onomatopoetic' words such as *jhemjhemiyā* 'small cymbal or drum', *bhubhui* 'white scurf', *gula-gula* 'mild' (with the usual middle Vedic, OIA, Tamil, etc. form of the "expressive" and onomatopoetic words: type *kara-kara* versus older Vedic *bal-bal*).

In Nahali (spoken on the Upper Tapti River) Kuiper 1962: 58 sqq.) the following substrate words can be found, though apparently various types of consonant groups are allowed: *bekki* 'to reap', *beṭṭo* 'to die', *bokko* 'hand', *coggom* 'pig', *cutṭi* 'to pound', *joppol/jappo* 'a clan name', *kaggo* 'mouth', *kallen* 'egg', *maikko* 'bee', *oṭṭi* 'to pull out, to burn', *poyye* 'bird', *unni* 'to take'. Additions to this list can easily be supplied now from that of A. Mundlay (MT II) which are not obviously from NIA include 8 *aḍḍo* 'tree', 91 *attú* 'to stretch', 221 *bijjok* 'to lay in wait for prey', 232 *biṭṭhāwi* 'union, horizon', 255 *buddi* 'to set (sun)', etc.

In the Drav. Nilgiri languages (Zvelebil 1990: 63-72) there are a few isolated geminating words that go back to a pre-Drav. substrate, e.g. Irula *mattu* 'lip', *ḍḍekkada* 'panther', *mutṭ(u)ri* 'butterfly', *vutta* 'crossbar in a house'.

The Vedda substrate contains the same type of words: *cappi* 'bird', *potti* 'a kind of bee', *panni* 'worm' (de Silva 1972: 16).

It can be stated, therefore, that the substrate languages outside of the extreme northwest indicate broad evidence for original geminates. Differently from IA (cf. below, on Turner's reconstructions), these words have not been pushed through the 'filter' of MIA, that means their original consonant clusters have not been 'simplified' (e.g. *kt* > *tt*, *kṣ* > *kkh*, etc.) The tendency probably has worked on IA from the beginning, as for example in the early example AV *kṣullaka* < *kṣudraka*. In Drav. various consonant groups are allowed, including geminates (Zvelebil 1990: 10 sqq.): e.g., *kakku* 'to vomit', *kaccu* 'to bite', *katttu* 'to tie', *kattu* 'to screech', *kappu* 'to overspread', *kammu* 'to become hoarse'; (cf. also the interchange *p*- :: *-pp*-/*-v*- :: *-p*-/*-u*).

One can therefore put the question whether this old substrate tendency has already influenced the Para-Munda of the RV. In Munda itself, such geminates are very rare (cf. Kuiper 1991: 53), and open syllables are common. However, there is a tendency in the Munda languages to eliminate consonant groups caused by vowel loss in prefixes (Pinnow 1959: 457); this does not cause geminates in such cases but is in line with the similar developments from Old to Middle and New IA (e.g. *akṣi* 'eye' > *akkhi* > *ākh*, *rakta* 'colored, red' > *ratta* > *rāt*, etc.). One may therefore explain many of the 'foreign' words with geminates in Vedic and post-Vedic, excluding Drav. loans, in the same way.

For the same area that is covered by Masica's language "X", and for N. India in general, one may also adduce the many words in NIA that are not attested in Vedic, Classical Skt. or the various MIA languages such as Pali but that occur only in their NIA form. They have been collected and reconstructed by V. Turner in his CDIAL. These include the starred forms, appearing in their reconstructed OIA form, and those words that do not appear in Ved. but are more or less accidentally attested in late Skt. texts, and the substrate words dealt with by Turner. They have a typical, often non-IA structure, including the very common clusters *-ṇḍ-* and *-ṭṭ-*. Their root structure follows the following pattern. (C = any consonant, ə any vowel)

\**Cəkkh*, *Cəg*, *Cəgg*, *Cəcc*, *Cəcch*, *Cəjj*, *Cəñc*, *Cəṭ*, *Cəṭṭ*, *Cəṇṭh*, *Cəḍ*, *Cəḍḍ*, *Cəḍg*, *Cəṇḍ*, *Cəḍḍ*, *Cəṇ*, *Cəpp*, *Cəmp*, *Cəbb*, *Cəmm*, *Cəṛ*, *CəṛC*, *Cəḷ*, *Cəll*, *Cəv*, *Cəs*, *Cəsś*, *Cəh*.

In Turner's CDIAL there are only a few forms such as \**Crək*, *Crəc*, *Crəŋt*, *Crəll*, *Cləkk*; this does not surprise as all reconstructed words have passed through the filter of MIA and have lost such clusters, -- except in the extreme northwest (Lahnda and Dardic).

Double consonants at the end of roots may go back to complicated clusters that can no longer be reconstructed, for example \**Cəkkh* < \*\**Cəkṣ* (cf. RV *kṣvīnkā* 'an animal, vulture?', *ikṣvāku* 'name of a person, tribe' (class. Skt. 'bitter pumpkin'), and compare Ved. clusters such as *matkuṇa* 'bed bug', *matkōṭaka* 'white ant', *kruñc* 'curlew'). Consonant clusters with various realizations in pronunciation may also be hidden in many Vedic loan words (Kuiper 1991: 51 sqq., Ved. cases p. 67 sqq.)

### Prehistoric Semitic loan-words?

In passing, a few notes on Cyrus Gordon's and Liny Srinivasan's discussion of Semitic loans in NIA, MT 1, 203-206. Most of them are 'disguised' derivatives of earlier stages of IA, a warning to be heeded when comparing S. Asian words with their long literary tradition with other languages (see above, introduction and cf. P. K. Benedict, MT III 93). I had a talk with C. Gordon about that time, but unfortunately we only discussed the Mitanni Aryan words (see MT I 206.)

I briefly list all their words (except for a handful that I could not yet explain) that *cannot* be derived from a Canaaitic source but stem from *earlier* forms of IA (Vedic, Class. Skt., Prakrit, NIA). Turner's CDIAL discusses the stages of development from OIA > NIA.

- *ṣar* 'bull' < MIA *saṃḍa(ka)* < Ved. *sāṇḍa* MS, *ṣaṇḍha* CDIAL 13331
- *sita* 'winter' ~ Ved. *śīta(la)* 'cold', etc. CDIAL 12485-8; *-t-* in *sita* remains a problem; it requires a compound with *śīta-*
- *gol* 'round' < Ved. *golikā* 'little ball', Skt. *gola* 'ball'; origin unknown, CDIAL 4321
- *mita* (= *mitā*) 'friend' < Ved. *mitra* CDIAL 10124
- *celi* 'purple red ritual garment' < Ved. *cela* 'clothes', Skt. *celika* 'bodice'
- *folā* 'swelling' < Ved. *phala* 'fruit', etc. CDIAL 9051 and PHAL 'burst', note Bengali pronunciation of *a* [ɔ], *ph* [f].
- *tham* 'to stop' cf. MIA *ṭhape-*, *ṭhava-* < *sthāpaya-* 'cause to stand, establish', MIA *thāma* < Ved. *sthāman* 'station', cf. Gujarati *ṭhām* 'place' CDIAL 13756-65
- *hoi*, *haya* 'is, are' < MIA *bhavāi*, *hōi* < Ved. *bhavati* 'to become, be' CDIAL 9416
- *bagan* 'garden' < NIA, Hindi *bagīcā* < Persian
- *bas* 'cloth' < Ved. *vas* 'to clothe', *vasana* 'dress' CDIAL 11437
- *thoka* 'drive in a nail' ~ cf. MIA *thaddha* < Ved. *stabdha* 'firmly fixed' CDIAL 13676; the form requires OIA \**sthabdha-ka*, like CDIAL 13675 *stabaka* 'tuft' > MIA *thavaya-* > Beng. *thok*
- *Abhira* : these tribes (or unknown origin) appear in S. Asia only at the beginning of our era; but the connection of Mitanni Aryan speech with pre-Vedic Skt. is beyond doubt (p. 204).
- *banā*, *banā* 'build' < Ved. *vana-* 'to desire, gain, make ready' CDIAL 11260
- *dha* 'run quickly' (for Beng. *dhāoyā?*, Oriya *dhāi-bā*) < Ved. *dhāva-* 'to run' CDIAL 6802
- *tola* 'draw up water' < Ved. *tolaya-* 'lift up' CDIAL 5979

- *gada* 'cause mental anxiety'; cf. CDIAL 3960 MIA *gamja-* 'to oppress, rebuke' < \**gañj* 'to press, ram'??
- *dhakal* 'trouble, misery' CDIAL 5581 < \**dhalati* 'bend over, fall'? (note extension with *-kk-* in Hindi *dhalaknā* 'to lean over', Beng. *dhalakā* 'to get loose'; or rather CDIAL 6701 \**dhakk* 'push, strike', Hindi *dhakelnā*).
- *chalak* 'smart, diplomatic' probably ~ Ved. *cal*, *cālaya-* 'to move' CDIAL 4772 (with common *-ka* suffix).
- *chamara* 'a low caste' < Epic Skt. *carma-kāra* 'leather-worker'
- *dhapas* 'fat, inert', cf. CDIAL 5580 \**dhappa* etc. 'lump', Beng. *dhepsā* 'swollen', Marathi *dhēbūs* 'lump'
- *karat* 'saw' < MIA *karapatta* < Ved. *kara-pattra* 'saw' CDIAL 2795
- *pala* 'to flee' < Ved. *palāya-* 'go away, flee'
- *pa-char* 'to disperse' < *pra-cala-* 'move forward?' CDIAL 8489
- *sach* (= *sāc*, *sācā*) 'pure, true' < MIA *sacca* < Ved. *satya* 'true' CDIAL 13112
- *kena* 'to buy' < MIA *kiṇa-*, *kiṇa-* < Ved. *krīṇā-* 'to buy' CDIAL 3594

### §9. Tibeto-Burmese

Still, this is not all as far as the Gangetic plains are concerned. The eastern section of the North Indian plains (E. Uttar Pradesh and N. Bihar) provides some indications of Tib.-Burm. settlements. The name of the Avadh (Oudh) area north of Benares in late Vedic texts is *Kosala*; this form should not appear in Vedic/Skt.; it should have been \**Koṣala* or \**Kośala* (as is indeed found in the Epics). The word clearly is foreign, and should belong, together with the slightly more eastern river name *Kauśikī* (post-Vedic, mod. *Kosi*) to a Tib.-Burmese language. Such designations for 'river' are indeed found in eastern Himalayish: R. *Kosi*, many Rai river names in *-ku*, *-gu*, in medieval Newari (*kho*, *khu*, *khwa*; *ko* 'river' in the unpublished Newari Amarakośa) and modern Newari (*khu*, *khusi* 'streamlet, creak') in and near the Kathmandu Valley, where it is already found in Licchavi time inscriptions, 464-750 CE, as: Cullaṃ-*khu*, Theñ-*khu*, Japti-*khū*, Huḍi-*khū*, Pi-*khu-*, Vihliṃ-*kho-srota*, Ripsiṃ-*ko-setu*. It is perhaps derived from Tib.-Burm. \**kluñ* (details in Witzel 1993).

Perhaps one may add the name of the tribe around Benares (*Kāśī*) whose older, Vedic form is *Kāśi* (AV), and its western neighbor, the *Kūśāmba*, *Kauśāmbi* (the later town *Kauśāmbī*, mod. village of *Kosam* near Allahabad). R. Shafer (1954) has a host of names, taken from the list of peoples in the much later Mahābhārata Epic that must be taken with caution (redaction only c. 500 CE, where even the Huns are included with *Hūṇa*, *Harahūṇa*, - they have become a Rajput clan!)

Indeed, early evidence for mountain tribes which might have been Tib.-Burm. is found in the Vedic texts all along the Himalayas. These mountain tribes, probably of Himachal Pradesh and Western Nepal, lived on the border of the Vedic settlement. They are first encountered in AV (1200 BCE) under the names *Kirāta*, in the western Himalayas where they appear as herb collecting mountain girls (*kairatikā kumarikā* PS 16.16.4, ŚS 10.4.14., *kailāta* PS 8.2.5). The more eastern text VS 30.16 has them as living in caves; cf. also the popular form *Kilāta* PB, JB, ŚB; (for details see Witzel 1993, 1999, and cf. KEWA I 211, EWA I 352, and also EWA I 311, s.v. *KAR*, and Prakṛt *Cilada*).

An alternate form of the name, *Kīra*, may have been retained in Kashmir, at 500 CE (see above). Since the RV, tribal names are found have the suffix *-ta/-ṭa* (Witzel 1999), e.g. *Kīkaṭa*, *Bekanāta* (certainly a non-IA name: *b-*, *-ṭ-*), *Marāṭa* PS 5.21.3, 12.2.1, *Kīrāta* AV, PS, *Āraṭ(ṭ)a* BŚS (cf. Sumer. *Aratta*, an Eastern country!), *Kulūṭa*, *Kulūta* (MBh), *Kulū-ta(ka)*, (but also: *Kolūta*, *Kaulūta*, *Kuluṭa*, and even *Ulūṭa*, *Ulūta*, see Kuiper 1991: 38 (cf. Pinnow 1959: 198f., cf. S. Lévy, JA 203, 1923, 52 sqq. = Bagchi 1929: 119 sqq.), finally *Kuḷu* in W. Pahari, CDIAL 3348, with the typical prefix change of Munda; *Virāṭa*, a king of the Matsya (Mbh) and a country in Bṛhatsaṃhitā, Pkt. *Virāḍa*, mod. Berar.

However, names in *-ta* (and *-nda*) are restricted to the Himalayan mountains while those with *-ṭa* (and *-ṇḍa*) occur all over the northern Indian plains (Witzel 1999). As for the origin of the suffix *-ṭa*, compare the plural suffix *-ṭo* in Nahali (Berger 1959, Mundlay *MT* II, 1996, 5, cf. Kuiper, 1991: 45 on 'Dravidian' *-ṭa*).

Beyond this, the early texts do not allow us to decide on the language and appearance of the *Kīrāta*. (The Epic calls them gold-colored). However, MS and ŚB list them with the Asura ('demons') *Kilāta-Akuli*.

Apart from these Vedic sources for (possible) early Tibeto-Burmese, the earliest datable, and so far not utilized evidence is found in Nepalese inscriptions (467 CE+)<sup>16</sup>. The inscriptions are in classical Sanskrit, but contain a host of place names, some personal and tribal names, and even a number of non-Sanskritic, traditional local names for government offices which must be considerably older than c. 200 CE.

A note on the transcription of 'foreign' words in Sanskrit and in Indian alphabets is in order here. Just as in the case of adaptation of 'foreign words' to the R̥gvedic phonetical pattern, the local words of the Kathmandu Valley had to be adapted to the possibilities of Sanskrit pronunciation and of spelling them in the Gupta (Nāgarī style) alphabet.

- several vowels are used intermittently: *i/e*, *i/ī*, *u/ū/o* (also *va/o*), *ṛ/ri/o* [ə,ɔ];
- there is variation in some consonants as well, notably:  
d/ḍ (no retroflex!), tt/ḍ, k/kh, b/bh, ll/ l, s/ś (no ś ?); jñ (common N. Indian pronunciation: *gy?*); note aspirated m, n, r [hm, hn, hr].

Typical is the spelling of the government office *śolla/śullī/śulī* or of the name of the town of Bhaktapur in Licchavi inscriptions: *Khr̥puñ*, *Khopr̥ñ* [*khɔpr̥iñ*], (*Mā-*)*kho-*, > medieval *Khvapo*, *Khvapva(m)*, *Khvapa*, *Khapva*, *Khopva* [*khɔpa*] > mod. *Khvapyā* [*khɔpe*], (for medieval names see Witzel 1999, 1993).

Of importance is a variation (just as in Kanauri) that indicates implosive (unreleased) consonants: *co/cok/cokh*. On the other hand, final *-k* must, at least in part, still have been pronounced in the late middle ages as it has been taken over into Nepali during the 17th and 18th centuries, e.g. *Jama-cok*, *Pul-cok*, or cf. the Patan toponym Nep. *Nugal* < New. *Nugaḥ* < O.New. *Nogvala*, *Nogola*, *Nogala*; or the Nep. loan word *jhyāl* "window". -- For all such variant spellings in the Licchavi inscriptions, see Witzel 1980: 327, n. 60,69, 72, 74, 75, 87, 1993: 240 sqq., 248, n. 171-3, and 1993, n. 120, 152.

<sup>16</sup> Now there is one still older inscription which indicates Sanskritization of the valley already around the time of Jayavarman, c. 200 CE (see Tamot and Alsop *Asian Arts*, July 10, 1996, at: [www.asianart.com/index.html](http://www.asianart.com/index.html)).

The actual attribution of the locally spoken language and its substrate found in the Licchavi inscriptions remains in the balance. It may be early Newari or a predecessor, the Kirāta language of the so-called Kirāta dynasty (see below) that reigned in the valley well before 200 CE and has left us with names of government offices such as *śulli*, *kuthera*.

If it is indeed early Newari, it is a very archaic form, characterized by a large numbers of *initial* clusters (*Cr-*, etc.), which differ even from the oldest attested Newari texts (names, occasional words or phrases in early Newari in documents, of 983 CE.) Such consonant clusters are very rare in medieval and certainly in modern Newari.

A clear case for Tib.-Burm. is *ti* 'water'; I have compared (1980 n. 90, n. 94) *co(kh)-*, *bu-*, *dol/dul*, *khu*, *gal/gvala* of the Licchavi inscriptions with mod. New. words: *-co* "hill, mountain top", mod. New. *cwa*, *cwak-*, cf. Kaike *chwang*, Khaling *cong*; (note also *cuk* "mountain range" in Gilyak); *-bu*, 'land'; O.New. *bu/bru*, cf. Tamang *pū*; *-gaa* '\*village'? cf. Mod. New. "classifier for round objects, part of Kathmandu", O.New. *gvala(m)*, but note Skt. *gola(ka)*, 'ball, globe'; perhaps cognate with Tib.-Burm. (Benedict, 1972: 444) *\*r-wa / \*g-wa*; cf. 91 *\*wal* 'round'; *-ko* 'slope', *kwa*, *kwaa* 'down'; *pā-kā* 'slope of a hill'; cf. Thakāli *koh-plen*. (K. P. Malla has explained some of such place names as being of Newari origin (1981: 17).

In the following list of names, place names are not specially marked, words ending in - continue with Skt. words such as *-adhikāra* 'office', *-kara* 'tax', *-grāma* 'village', *-draṅga* 'fortress', *-nadī* 'river', *-pāñcālī* 'association', *-vāstu* 'area'.

ajika-(monastery), aśiñ-ko (area) (*ko* 'river? or *ko* 'slope?'), ānlābaka-(association), uṭṭane, uḍanehuśa, uḍra, etañ- (village), kañkā-vaṭṭikhā (Skt. *vāṭika* 'garden?'), kañku-lam (area) (*lam* 'road?'), kaḍam-priñ (area) (*priñ = prñ*), kapiśā (river), kampro-yambī (see *jamayambī*), kambīlampra, karvaṭa (hamlet), kalopi- (village), kādalaka- (village), kāduñ- (village), kici-priciñ- (village), kuthera-(office), kurpāsi- (village), kuhmuṃ- (area) (see *hāhmuñ*), kekha, keṭumbāṭa (name of a Kirāta official), koñ-ko (village), kolī (cf. *dakṣiṇa-koli-grāma* 'Southern Koli village'), kośī (river), khakam-priñ, khaḍuka, kharjurikā-(monastery), khaḍabraṃśai, khādyām, khārika, khārevāḷga-co (*co*, *cok* 'pass'), khāhri-co, khudū-(deity), khula-priñ- (village), khrpuñ- (village), khainaspu (area), kho-prñ- (village) (see *mākho-*), gaṃprondiñ (village), gaṇiduñ (village), gaṇḍakī (river), gamme (area), gañśul (village), gī-gval- (association), giṭā- (association), gīnuñ, guñḍī-maka (village) "hill-water"?, guḍan-dul (area) (*dul* 'river?'), gumpadvṛs (area), gullataṃga- (village) (see *gollam*), gechimjaka, gollam (river), gohala- (village), gvalam (~*gollam*?), (catur-)bhalatāsana- (monastery), citalāñ, ciśimaṇḍa (*tila-maka*), cu-priñ- (river), custuñ- (village), custun- (river), cuhuṅgapeḍā, cullam-khu (~ *collam*?), cokh-parā (*cokh* 'mountain pass'), chūma-kūti, cho-guṃ- (village), jama-yambī (s. *kampro*), jayapallikā- (village), jaya-lambha (cf. *lambā*), jajje- (association), joñjondiñ- (village), jol-priñ- (village), jñātikhṛn (river), ṭistuṅga, ṭe-gval (village) (see *tegval*), ḍolājī-prathā, tam-brū? (cf. *Lalita-b(r)umā* name of Patan; cf. *-bū?*), tāṭṭanakam, tavecekha, tāñ-, tim-brū (cf. *ma/mittam-brū*, *prīti-brū*), tīla-(river) (= *tila-maka*? 'irrigation channel?'), tīla-maka, (śrī)-tukhāṇa, tuñ-catcatu- (village), the-khuṃ-dul (river) (cf. Tib.-Burm. *dul* 'dust', local meaning 'sandy river?'), tegvañ-, te-gvala (see *tegval*), te-gvala-(association), teñ-khu, tepula (office), testuñ- (village), thumtuṃ-rī- (fortress), thasam-prin- (deity),

thambi-dul, tham-bū- (village) (Tib.-Burm. *bū* 'rice paddy'? Benedict 135, *bu(d)/pu* "open land" Ben. 260), theñ-co, (dakṣiṇa-)koli- (village), (dakṣiṇa-)tilaḍu-ku (*tila*='grain?'), dañkhuttā- (street, tax), daṇḍaṅ-guṃ, dumprañ- (village), dumlañ- (village), duprañ, dommāna, dolā-(śikhā) (deity), dova- (village), dhelan-tī (river) (cf. Tib.-Burm. *tī* 'water'), nara-priñ- (village), nālaṅga- (village), nim-brū (Tib.-Burm. *bū* "rice paddy?"), nīli-śāla, pañkuṭi, panapphu (area), pan-(river), parikhā, palāṇḍu- (tax), pākhuśi, pā-guṃ-maka, pāsīnkhya, pikañkulaka (area), pikhū- (village), pīṭalja-(office), pumdaṭṭa (village), puṭham-priṅga, puṇḍri-(palace), puttī- (river), puttī-(deity), punu- (association), pṛcchi-brū, poṇḍi-(shrine?), prañ-priñ, praṇāli-dī-maka (Tib.-Burm. *tī* 'water?'), prayiṭṭikhā (area), pritu-brū (Tib.-Burm. *bū* 'rice paddy?'), promjñam-bu, pronniprañ, pronprovāñ, phāṃsinpral (river), phalanju (corvee), phavadrañ-(village), phrṭhula (area), phathula (area), pherañ-(fort), bugā-yūmi- (village), bunlu- (river), bemmā (area), brahmuñ (office), brā-dul (river), brem-guṃ-co (pass), bhumbhukkikā- (deity), bhel-bū, bhoṭṭa- (corvee), bremgu-co, mā-kho-duluṃ, mā-kho-prñ (fortress), mā-gvala, māṅga- (village), mātin-(temple), māthañ- (village), māp-cok-(office) cf. -*co(k/kh)* 'pass', māśa, miñ-ko, mittim-brū?, miṇḍi-co (pass), mekañ-ḍi-dul (tilamaka) (Tib.-Burm. *tī* water?), mo-guṃ-co (*co* 'pass'), yaku, yā-priñ- (village), yāvi- (village), yū-gvala- (association) (cf. *gola*), yū- (village), yūvisāmā- (village), yebramkhara, roṭā- (association), rogamācau (watchman), lakha-maka, laṃkhulāṃ uṭṭane, lañja-gval- (association), laḍitama-(deity), liñ-gvala- (office), lunśrī-(area), lumbañ-co, lulju (river), leṃdu (area), lembaṭi-(fortress), lo-priñ- (village), lo-priñ- (association), vāditra- (association), valasokṣi-(temple), viñvocā-(shrine), vilivikṣa (area), vihliñ, vihliñ-kho (river, Tib.-Burm. *khu* 'river?'), vṛjika, vṛjika- (street, highway), vempra- (village), (vaidya-)madgudi (village), vottarino?, voddī- (province), śaktibāta (corvee), śaṅgā- (village), śaṭammi (area), śalaṃkhā, (śiva)-gal- (temple), śiṭāṭi, śulhmuñ (office), śolla (office), śulli (office), śuli (office), sañ-ko, sattvaumā-lambā (area) (cf. *lambha*), sapelā- (association), salam-bū (palace), siṃ- (tax), sindrira (watchmen), subraṃ-ko (area), surisimbattī, stharu- (fortress), hasvimavallī- (village), hāhmuñ- (place), hima-? (river), huḍi-khū (Tib.-Burm. *khu* 'river'), hus-prinduñ (village), hnā-guṃ, hmas-priñ- (village), hnu-priñ, hrīm-ko (area).

There also is a traditional list of local kings, 32 in all, called the Kirāta dynasty (transmitted only in a manuscript of 1389 CE, the Gopālarājavamaṣāvalī), it runs:

Elam, Pelam, Melam, Caṃmiṃ, Dhaskem, Valuṃca, Huṃtiṃ, Huramā, Tuske, Prasaphuṃ, Pavam, Dāstī, Camba, Kaṃkam, Svananda, Phukoṃ, Śiṃghu, Julam, Lukam, Thoram, Thuko, Varmma, Guṃjam, Puska, Tyapami, Mugamam, Śasaru, Guṃṇam, Khimbuṃ, Girijaṃ, Khuraṃja, Khigu.

Some of these names obviously are Sanskritic: *giri-ja*- 'born in the mountains'; Svananda (*sva-nanda* or *sv-ānanda*); *Varmma* (*varmā*, the designation of Kṣatriyas and kings); note that a new Kathmandu inscription of c. 200 CE is already one of *Jaya-varman*; consequently, this list will go back to at least 200 CE. The rest looks Tib.-Burm; note the initial clusters *pr-*, the internal clusters *-sk-*, *-st* and the final nasals: these features agree with the phonetic shape and the syllable structure of the place names recorded in the Licchavi inscriptions.

All these data have not yet been exploited for Tib.-Burm. linguistics. (For place names, see Witzel 1980, 1993; for relations between the eastern Himalayan languages and Munda, s. Kuiper 1962: 42, with Nahali, p. 46f; cf. Laufer 1916-18, 403 sqq.).

The Kathmandu Valley, however, seems to have has its own strange substrate, below this Tib.-Burm. level. It is visible in some place names which definitely do not look Tib.-Burm. Some of them are characterized by the geminates studied above: *gamme*, *gullatam̐ga*, *gollam̐*, *jajje-*, *dommāna*, *daṅkhuttā-*, *bemmā*, cf. also *bhumbhukkikā* (onomatopoetic with double consonant < \**bhumbhum-ki-kā*?)

Our task would be simplified if we had an etymological dictionary of Newari, but so far we only have a limited dictionary of O.New. (by H. Jørgensen, A dictionary of the Classical Newarī, København 1936) and an equally limited one of modern Newari by Th. Manandhar (Newari-English Dictionary, Delhi 1986); most of the older New. texts have not even been edited. The next step would be to eliminate all Skt. loan words; (they are often difficult to determine due to diverse and far reaching sound changes, and to telescoping: who would derive *punhī* 'full moon' < Skt. *pūrṇimā*, or *yege(ṁ)* 'offering' < Skt. *yajña*, *ācāgu* 'yearly meeting of all Kathmandu Bajrāchārya priests' from Skt. *ācārya-goṣṭhikā*?) Only then, we can be sure as to what is Newari, and what not and can proceed to eliminate Tib.-Burm. and other loan words in order to trace the substrate language of the Kathmandu Valley. A shortcut, for the time being, is provided by those untypical words with geminates mentioned above.



## §10. Himalayan Languages

### §10.1. Early IA settlements in the Himalayas

D. D. Sharma, Old-Indo-Aryan element in Kinnauri (in: R.K. Sharma et al. (eds.), *Dr. B. R. Sharma felicitation Volume*, Tirupati 1986, 149-155) describes older elements in the Kōchī dialect (of LSI 9, 4: 613-6), classified by Grierson as a subgroup of the Kīūṃṭhali Group (= Simla dialects, p. 549 sqq.). It is spoken in the western part of the former state of Bashahr, along the upper Satlej River (Sharma's Lower Kinnaur, including the Kalpa, Nichar and Sangla Valleys), while in the eastern part (Sharma's Upper Kinnaur), up to the Tibetan border, the Tib.-Burm. language Kanauri (Kanāw<sup>a</sup>rī) is used.

Sharma states that Lower Kinnauri has 'a substantial portion of Tibeto-Himalayan vocabulary' and that it has 'various linguistic elements in its sub-strata ... whose origin is sought elsewhere'. One of these substrata is, in his opinion, an Aryan element that cannot have come from the (Vedic) OIA of the plains (*Madhyadeśa*), since he regards linguistic and other contact impossible ("unthinkable"; this can of course, be doubted, as we now have early iron age civilizations in these hills and contacts with the plains).

He links his 'Lower Kinnauri' with "an independent group of Aryans in the prehistoric days" ... the "Khaśa and Yakṣa" who spoke the Aryan language of the pre-Vedic period." He also sees a "conspicuous absence of Brahmans and Brahmanical culture" (which, incidentally, is quite typical for most of the upper Himalayan regions with NIA languages). Or, this form of Aryan is seen by Sharma as that of the Khaśa people "who form the bulk of the populace of this region now-a-days.... The existing OIA elements are the remnants of language of these Khaśas". He thus is on a trail quite similar to that taken later on by Zoller (see below).

The vocabulary given by Sharma, however, shows traces of OIA, MIA and NIA -- as might have been expected. One curious feature of L.Kin. is the division of nouns in animate (suffix *-s*) and inanimate (suffix *-ñ*) which he compares to that of the Munda languages, while he links the endings to OIA masc. *-s*, neuter *-m*.

However, his materials represent a mixture of OIA, MIA and NIA forms that have to be separated. Typically, we find OIA *kvath* 'to boil' preserved as *kwath* or *grāma* 'village' as *grāma-ñ* (as opposed to NIA *gaū/gaō* etc.); next, forms which represent a MIA stage such as *sappa-s* 'snake' < *sarpa*, and NIA forms such as *bāyā* 'brother' < *bhrātā*, *tau* 'heat' < *tāpa*, *dauya-ñ* 'curds' < *dadhi*, *ana-ñ* 'food' < *anna*, or *māmā* 'maternal uncle.'

Then, there are earlier and later loans directly from Sanskrit (*tatsama*). The earlier ones have undergone some sound changes typical for this NIA language. Direct loans include *śāstra-ñ* < *śāstra*, *rākṣa-s* < *rākṣasa*, *bāga-ñ* 'part, share' < *bhāga*; older ones must be: *dhaura* 'religion' < *dharma* has preserved *dh* (otherwise > *d*, a typical Dardic trait), or *akhaura* < *akṣara* 'letter'. There are several cases of "Gāndhārī metathesis" as well: *trāma-ñ* 'copper' < *tāmra*, cf. *grota-ñ* 'cow urine' < *gomūtra* etc.

In short, several layers have to be distinguished very carefully; Dardic influences and medieval loan words from Skt. have to be separated, and finally, true OIA survivals must be isolated, -- all of which cannot be done here. This Pahari language thus contains many loan words from the levels of OIA and MIA and NIA.

The case is of interest as it shows, just as that of early Burushaski, the interaction of plains and mountain people (cf. also, below, on Bangani). The present case also provides

some indication of the early date of such interaction between IA and Tib.-Burm. speakers; this may be reflected even in AV, if the *Kirāta* indeed are Tib.-Burm. speakers, and if the name has not been passed on from an unknown earlier population (cf. the Kashmiri *Piśāca*, *Nāga* traditions, above) to Tib.-Burm. speakers.

However that may be, from at least 1100 CE onwards, we see an increasing Aryanization of the western Himalayas and W. Nepal with the spread of the *Khaśa* tribe (found already in Manu's law book); by 1150 CE they are still mentioned in the *Rājatarāṅgiṇī* as settling southwest of the Kashmir Valley. *Khas kurā* is the self-designation of what was called the "language of the Gurkhas" (in Newari called *khaṃy* < *khas*); they have substituted the name Nepali only in this century. By 1150 CE they had established the W. Nepal/C. Tibetan Malla kingdom; by 1769 they had conquered the Kathmandu Valley; and by 1900 they had settled, mixed with Gurung, Magar, and other Tib.-Burm. tribes speaking Nepali as lingua franca, in Darjeeling, Sikkim, S. Bhutan and some parts of Assam. This movement is indicated by their renaming of river names all across the Himalayas (Witzel 1993).

Some part of the Himalayas may also have been occupied by the pre-Tibetan language of W. and Central Tibet, Zhang Zhung. (See the list of Zhang Zhung words, Thomas 1933, Beckwith 1987; for recent archaeological discoveries of Zhang Zhung settlements in the area before the spread of Buddhism in the 7th c., see Bellezza, 12/17/98, at: [www.asianart.com/index.html](http://www.asianart.com/index.html).)

The history of the settlement of the Himalayas is far from clear. (For some details, based especially on hydronymy, see Witzel 1993, and cf. now van Driem <http://iiias.leidenuniv.nl/host/himalaya/>). For example, the Thāmi tribe who live higher up in the Tāma Kosi valley east of Kathmandu belong, as their language shows according to Shafer (1964: 3 n.1), to the Western Himalayish group of the Bodic division of Tibeto-Burmese (Kanauri, etc.); cf. however, Starostin 1989. Indeed, the Thāmi claim to have immigrated from Humla in northwest Nepal. This is one indication among others (Witzel 1993) that there was a west-east flow of population and languages, similar to the much later one of the Nepālī speaking Khas tribe.

### §10.2. Bangani

The intriguing question of Bangani has not been entirely resolved. Bangani is spoken just east of Kinnauri, in the western-most tip of Garhwal, Uttar Pradesh. Zoller (1988, 1989) has reported a non-IA substrate in this otherwise typical NIA language found high up in the western Himalayas. Surprisingly, this substrate is a strange western variety of IE with words such as *ɔgnɔ*~ 'unborn' (not Skt. *a-ja*) and *gɔnɔ* 'give birth' (not Skt. *jan*), *kɔtrɔ* 'fight' (not Skt. *śatru*), *dɔkru* 'tear' (not Skt. *aśru*); the initial *d-* is W. IE, cf. Greek *dakru*, Engl. *tear*, as opposed to E. IE : Skt. *aśru*, Avest. *asru*, Lithuanian *ašara*. This claim has been disputed by G. van Driem (1996, 1997), but has been sustained by research carried out in Bangani by Anvita Abbi of Delhi University (see H.H. Hock [On Bangani] <http://www-personal.umich.edu/~pehook/bangani.html>, with further discussion). Anvita Abbi recognizes three layers in Bangani: words of the type *dɔkru*, *lɔktɔ*, *gɔsti*, the general NIA Pahari level, and recent loans from Hindi, etc.

In principle, bands or tribes who have 'lost their way' and turn up in unexpected areas are not altogether unknown. Tokharian, the easternmost IE language, has western

characteristics (*känt*, *känte* '100'), and the North Iranian Alani, ancestors of the Ossetes, traveled all the way through Central Europe, Spain and North Africa with the Germanic Vandals, to settle in Tunisia.

### §10.3. Kusunda

Tib.-Burm. is, however, not the first language in the Central Himalayas. In Nepal it has been preceded by the language isolate of Kusunda. (The language of the nomadic hunter-gatherer group, the Raute, is Tib.-Burm., though; see D.B. Bista, 1976, J. Reinhard 1974). Kusunda has recently been treated at length in MT II and III (cf. Shafer, 1966: 145; 1954: 10 sqq.), and I can therefore be brief. Personally, I deeply regret not having investigated the language when I worked at Kathmandu (1972-8), at a time when this was still possible; I thought that the Summer Institute of Linguistics (T. Toba 1970) and J. Reinhard (1969, 1970) were on the trail. At that time, some Kusunda still lived in the Mahabharata Range west of Kathmandu, in the village Satobati on the river Kar Khola, west of Gorkha. This is, however, a fairly recent settlement, and we should investigate whether other sections of the tribe have survived elsewhere. One of my Nepalese friends, J.R. Acharya, tells me that some decades ago, Kusunda used to come to his village near Tanahuṃ. Children were threatened by their mothers with exclamations such as "Kusūṇḍā ayo!" ("the Kusunda have come... they will take you away!" It is important to note that Reinhard reports them from another area, the Dāng plains (south of Gorkha). Their possible survival should therefore be investigated urgently -- though the question remains whether various groups of hunters called Kusūṇḍā in Nepal (such as those at Tanahuṃ) do/did indeed speak this language. The language is reported to have died out by now. The Summer Institute's web site ([www.sil.org/ethnologue/](http://www.sil.org/ethnologue/)), misclassifying them as Tib.-Burm., says: "KUSANDA (KUSUNDA) ... Tanahun District, Gandaki Zone, western hills, Satto Bhatti west of Chepetar and possibly jungle south of Ambhu. Kireni, near Kumhali. ... Last speaker died recently (1985). Extinct." Therefore, Reinhard's taped material of c. 500 words and sentences, deposited in the Phonogramm-Archiv, Vienna, should be (re-)investigated.

It is also important to point out the difference between Hodgson's (1848, 1880) and Reinhard's (1969, 1970) Kusunda, a point also mentioned by P. Whitehouse MT III : 31; however, these differences extend beyond the grammatical forms cited to the basic vocabulary, e.g. *gipan* 'hand' H(odgson) : *āibi* R(einhard); *ing gai* 'star/night' H : *sā'nām* R (cf. *ing*, *ing ying* 'sun'); *jum* 'moon' H : *nihō* R; cf. also smaller variations: *toho* 'tooth' H : *uhu* R; *gitān* 'skin' H *gitat* R.

It goes without saying that, for a thorough investigation of Kusunda, the loans it has received from Nepali and some of the neighboring Tib.-Burm. languages such as (Kham-)Magari, Gurung, Chepang, Newari, etc. must be taken into account, and that its relation to the nearby substrate in Tharu (and Masica's "Language X") needs to be evaluated.

S. M. Joshi's dictionary (*Paryācavācī Śabda Koś*, 1974) unfortunately has no Kusunda lists; on p. *kha* of the introduction he says (here translated from Nepali, with my notes enclosed in []): "... there also exist two leftovers of these families: Jhangar of the Dravida family, and Santhali or Sartar of the Āgneya [= Munda] family. [Note that the 1961 Census has both Dhargar = Kurukh in Dhanusha Dst., and Jhangar, see Zvelebil 1990: xxiv, n.24; Santali and Sartar are both spoken in the extreme southeast of Nepal]. Again, there are also

such languages about which certainty of their language family has not been reached, such as the Kusundā language. In the Census the language of a tribe, wandering about in small numbers and either living in village houses or not, and of other languages have been separated. Thus, the language of the Kusundā (a tribe found here and there in the Gandaki district) and of the Raute people (found in Rapti, Bheri, Karnali and Seti districts) [Tib.-Burm.] cannot be seen [in this dictionary]. But, from the point of view of anthropology and linguistics, the language of some such tribes is important."

#### §10.4. A Munda substrate in the Himalayas?

In passing, the old theory of a Munda substrate in the Himalayas should be revisited. It goes back to S. Konow 1905, 117-125. This has been denied by P.K. Benedict 1972 7, n. 23 and G. van Driem, Rutgers 1993, J.J. Bauman (1975), Turin 1998 (see website: <http://iias.leidenuniv.nl/host/himalaya/individ/kirmor.html>).

Nevertheless, it must be remembered that the name of the R. Gaṇḍakī can be traced back to Munda. It is found all over Central Nepal, where the major rivers are called "the seven Gaṇḍakī". How far into the Nepalese hills did the settlements of a Munda speaking people reach? Even in exclusively Nepali speaking W. Nepal, the common hydronomical 'suffix' *gāḍ* denoting 'river' may be connected with the Munda word *da'k*, *ganda'k* (Witzel 1993, 1999; further materials in Kuiper 1962: 10, with lit.; and already B. H. Hodgson (1880, 1848).

A further hint may be provided by the implosives (unreleased stops) found in the substrate of the Kathmandu Valley (*cokh/cok/co*, see above) and in Kanauri (see Grierson, LSI on Kanauri). We may see here an areal feature of implosives that has influenced both the Tib.-Burm. languages in Kinaur (Kanauri) in the western Himalaya and in the Kathmandu Valley. Apart from Munda and Sindhi, this feature is otherwise not found in S. Asia. There are indications in the eastern Himalayas of a pre-Tib.-Burm. population (Witzel 1993). Anecdotally, it may be mentioned that the Kulunge Rai, a Tib.-Burm. tribe in E. Nepal has legends about the earlier settlers of the Hongu valley, the *Rungsiupa*.

Even today, the Munda languages Satar and Santali are actually spoken in the extreme south-east of Nepal (probably, like the Kurukh, recent imports). Other Munda speakers are, after all, found south of the Ganges, only about a hundred miles south of Eastern Nepal.

#### §10.5. The Tharu substrate

Finally, there are the various Tharu tribes who live in the foothills of the Himalayas, from the Rāmaṅgā river in U.P. (India) to the eastern border of Nepal, and in some bordering hill tracts, such as in the Rāptī Valley (Chitawan, just 50 miles SW of Kathmandu).

They practice slash-and-burn agriculture and nowadays speak a form of one of the neighboring NIA languages, just like the Nahali or Vedda (see below); however, I believe that we can find, again, a so far unstudied substrate from a pre-IA, Pre-Munda language.

Although often referred to as an archaic, remnant group, they have been little studied (cf. the bibliography in Leal 1972, see now Krauskopf (1989). G. Grierson (Linguistic Survey of India, 5.2: 311) reports the opinion of W. Crooke (1906) that they were formerly Dravidians who intermarried with Himalayan people. Indeed, rarely, some of the

vocabulary looks Tib.-Burm.: for example Tib.-Burm. *ti*- 'water' in Tharu *suitī* 'small river.' (For *-ti* in Himalayan river names, see Witzel 1993).

And indeed, D. N. Majumdar (1944) reports blood group types 'predominantly Mongoloid.' This is now supported by recent, more advanced genetic studies. The Tharu are very isolated within S. Asia (L. Cavalli-Sforza 1994: 84, 239 with fig. 4.14.1).

In Nepal, the Tharu have also consistently been reported to be immune against malaria. Their area was heavily infested until the use of DDT, in the Sixties and early Seventies; and no non-Tharu traveler stopped in this "8 kos" jungle belt overnight as to avoid catching the "mountain fever". L. Cavalli-Sforza 1995:125 gives a genetic reason for immunity. The anecdotally high Tharu immunity rate should be compared to the generally low Indian 'immunity gene' rate.

As for the suspected substrate, D. Leal (1972), provides an example of the influence of their original non-NIA language, i.e. the difficulty the Chitaun Tharu have to pronounce aspirated mediae (*bh* > *bəh*; cf. above, on the Kathmandu Valley substrate) and mentions as another ("Dravidian") substrate influence the simplification of the possessive case suffix Hindi *-kā*, *-ke*, *-ki*, Nepali *-ko*, *-kā* to *-k*.

I list some examples of suspected substrate evidence from the Tharu word list in S. M. Joshi (1974); this contains lists of 2914 words, starting from the Nepali entry. As in Grierson and Leal, most Tharu words in this dictionary are close to Bhojpuri and Nepali; a cursory check has resulted in the following words (cf. Witzel 1999, n. 43) which are neither related to the surrounding IA languages (Nepali, Awadhi, Bhojpuri, Maithili) nor to the nearby Tib.-Burm. ones (Magar, Chepang, Newari, Tamang). I propose a more detailed study in a later issue of MT. Some interesting words are:

- *ubbā* 'small box'
- *koñhilā* 'tiger'
- *khūdi* 'sugar cane' (cf. Ved. *ikṣu* 'sugar cane', MIA *ikkhu*, *icchu*, Maithili, Bhojpuri *ūkhi*, + Tib.-Burm. *ti* 'water' ??)
- *gukhā* 'shaman'
- *gulagula* 'mild'
- *gētṭī* 'splinter'
- *jhemjhemiyā* 'small cymbal or drum'
- *ṭippā* 'mountain top' (probably NIA)
- *ta* 'small'
- *tira* 'afterbirth'
- *tilvā* 'whore house'
- *nimak* 'salt'
- *bhubhui* 'white scurf'
- *yedi* 'brick'.

But the agricultural terms are NIA: *bājrā* 'millet', *dhān* 'rice', *makai* 'maize', *gehūṃ* 'wheat', as well as most of their basic vocabulary.

All these cases indicate that we probably can discover more substrates if more work along these lines would be done. But we lack etymological dictionaries for most NIA languages (apart from Turner's great work, CDIAL), not to speak of Munda (in preparation

by D. Stampe et al.) and Tib.-Burm.; (see, however, those on the internet: Starostin et al., accessible from: <http://starling.rinet.ru/>). For example, it may very well be that the Bihari languages have more Tib.-Burmese substrate words. There is, after all, *cāmal* 'cooked rice' in Nepali, *cāwal* in Hindi, etc. which can be connected with Tib.-Burm. \**dza* 'to eat', Newari *jā* 'cooked rice, etc.' Yet, nobody in Indian Studies is looking for such substrate material.

#### §11. Nahali (Nihali).

Turning further South, it may very well be that Rajasthani has a strong Bhili (and Nahali) substrate; Koppers (1948: 23, Kuiper 1962, 1966, 1991) and Shafer (1940, 1954: 10) thought that the Bhils once spoke Nahali as well. The Bhils are now widely spread between the Ārāvalā (Aravalli) Mountains, the Vindhya Mts. and the Tapti River (Khandesh area); they now speak Gujarati-like IA.

Again, as extensively treated in *MT* II and III, it should be underlined that Kuiper (1962: 51) distinguished 4 levels in this isolate language: some 25% substrate, then a Munda, Dravidian and finally a NIA layer. The vocabulary given by Mundlay in *MT* II should be reinvestigated by specialists of IA, Drav., and Munda. From the point of view of IA some words stand out, even if they have not come directly from Marathi or a Hindi dialect or not, whether they have been marked as L (loan-word), as the case of *akkal-kāyni* (above) shows: *MT* II p. 45 no. 161, p. 70 no. 10 *sanu* 'younger brother', belongs to CDIAL 12732 Ved. *ślakṣṇa* 'slippery, tender' > NIA: Panjabi *nannhā* 'small, young', Nepali *sānu* 'small', *nāni* 'little girl', Oriya *sāna* 'small, youngest', Hindi *nanh* 'small, light', Marathi *sānā* 'small' etc., or *MT* II: 36 no. 1274 *parayn* 'river', is other than maintained in *MT* II: 64 no. 17, a simple look-alike of Nostratic \**bihra*', as *parayn* (Kuiper 1962: 96, 1966: 78) is a borrowing from its neighboring language, Marathi *parhyā* 'streamlet, brook', and its dialect Konkani *parāy*, < *Skt. parivāha* 'overflow of a tank, water channel' CDIAL 7878, cf. *MT* II: 36.

Berger (1959) was of the opinion that the Nahals were identical with the well known Niṣāda of the Chambal, Malwa and Bandelkhand areas. He discussed their mythology as found in the Mahābhārata; however the *Niṣāda* (and once, the variant *Niṣadha*) are found already in the Middle Vedic texts (see below).

The people called Nihāl or Nāhal are found (Berger 1959: 35) in many medieval texts, such as in Hemacandra's Grammar (c. 1200 CE) as *lāhala*; in Padma Pur. *nāhalaka*, together with the *Bhilla*, as mountain/jungle tribe; in Puṣpadanta's Harivaṃśapurāṇa as *ṇāhala*, synonym of *bhilla*, *savara* (another jungle tribe : modern Saora); also in Vikramaṅkadevacaritra of Bilhaṇa (c. 1150 CE), and in Rājasekhara's drama *Balarāmaṇa* (on the R. Narmadā). Berger wanted to identify them with the *ḍahāla* as well; they are found in inscriptions of the Kalacuri dynasty of Tripurī and in Albiruni (1030 CE). All of their territories are c. 400 km away from the modern eastern Nahalis near *Nimar*.

He thus derived Nahal/Nihal from a form such as \**neṣad* reflected by Ved. *Niṣāda*. Indeed, the word is found in early post-RV texts: KS, MS, and with the typical sound changes in 'foreign' words: *Niṣāda* : \**Niṣidha* : ŚB 2.3.2.1-2 *Naḍa Naiṣidha*, (apparently the Vedic 'ancestor' of the Epic *Nala Naiṣadha*); thus, *d*: *dh* (as in *Magadha* : *Pra-magandha*, etc.). The name certainly is a popular etymology (however, the modern self-designation of the Nahals is *kalṭo*, du. *kalṭih-ṭel*, pl. *kalṭṭa*; < stem \**kalṭ-o*, s. Kuiper 1962: 82, 17, 27,

Mundlay MT II 5-7, no. 858 *kalṭo*, pl. *kolṭa*). The Niṣāda are described in Vedic texts (first MS 2.9.5 =KS 17.13, TS 4.5.4.2, VS 16.27) as being “neither wilderness (*aranya*) nor 'wagon trek, settlement' (*grāma*)” who are “given over to the earth:” (*asyām eva parītāḥ*), next to *jana* '(foreign) tribe' PB, other non-Brahmins (JB), and *samānājana* “one's own people” (cf. PB 16.6.7-9); cf. also KB 25.15, LŚS 8.2.8 on temporary residence in a *naiṣāda* settlement.

Similarly, MS 2.9.5 describes the Niṣāda, among the names of the fearsome god Rudra and his people, together with hunters and other low caste people (=KS 17.13, TS 4.5.4.2, VS 16.27); -- AB 8.11 as robbers in the wilderness; similarly the *dasyu* JB 2.423:§168, where the text insists on Kṣatriya accompaniment during travel, necessary to keep the Dasyu at bay and turn them *madhu* 'sweet' , cf. AB 8.11 where the *dasyu* rob a wealthy man or a caravan in the wilderness.

Acculturation is seen at MS 2.2.4, where their chief (*sthapati*) is allowed to offer sacrifices, cf. KŚS 1.1.12. The inclusion of the headman of the Niṣāda reflects the well-known process of upward social movement, called “Sanskritization.” (Witzel 1997a)

Their Vedic designation obviously is a popular etymology "those who sit at home." However, they are more frequently described as robbers (still a favorite occupation of the Nahals in early British times) -- against whom one had to guard when traveling through uninhabited territory. Their chieftains (*sthapati*), however, were allowed into the Aryan fold and could perform solemn Vedic sacrifices, clearly an early form of Sanskritization.

In passing, as has been first seen by Shafer and Kuiper, Nahali has connections with Ainu, etc. (see now Bengtson, *MT* II 51-55), remnants of the earliest substratum of modern *Homo Sapiens sapiens*' move from the Near East all the way to E. Asia (and S.E. Asia, Australia); note however, the differing views of V. Blažek, H. Fleming, and I. Peiros in *MT* II.

## §12. Dravidians in the Vindhya Range

Both North Dravidian languages, Kurukh (Oraon, on the borders of Bihar/Orissa/Madhya Pradesh; the settlement in Nepal and Assam is recent) and Malto (on the bend of the Ganges in S.E. Bihar) are late-comers to Munda territory as many loans from Munda languages indicate. Brahui in Baluchistan has returned to E. Iran only a few hundred years ago (Elfenbein 1987); it has no older Iranian loans (from Avestan or Pashto, just from their symbiotic neighbors, the Baluch).

In the Vindhya Mountains we find such names as the following: the *Vidarbha* people, in the area around Nagpur, (the mod. *Barhād*, *Berar* < *Virāṭa*, Mbh) are mentioned (JB), along with their fierce *mācala* dogs 'that kill even tigers' (note that this is an area with early iron and horses). *Vidarbha* seems to be a popular etymology *vi-darbha* 'with widely spread *darbha* (grass)', especially if connected with Munda *da'b* 'to thatch' (Pinnow 1959: 69), cf. *vi-bhindu* in the Gangetic plains (above). The name of the Vibhindus is related to that of the *Bainda* tribe (derived from \**bind*) that still survives in the Vindhyas today, and names such as *Ku-sur(u)-binda* (above). The very name of the Vindhya (post-Vedic) can be related, with typical Sanskritizing interchange of *d* : *dh*, as in *Pra-maganda* : *Magadha*, (above). East of these mountains, we have the *Kaliṅga* (cf. *Triliṅga* south of Orissa) and *Aṅga*, *Vaṅga*. All of these are names that hardly have a Drav. etymology, but which look Austro-Asiatic because of their prefix changes.

However, all around Vidarbha, the first Drav. river names are met with : the *Pūrṇā* (< \**pēṇ*) west of it, the *Vēn-gaṅgā* east of it, and the *Pain-gaṅgā* south of it. They all are adaptations of a Drav. term for rivers, DEDR 4160a. \**pēṇ*-: \**pēṇ*-V- 'to twine, twist'. It seems that the area which still has a Munda name in the Vedic middle period (*Vidarbha*) has also received a Dravidian overlay. This is confirmed by Drav. place names in *-oli* in Maharashtra and in *-palli*, *-valli*, *-pal* in Bastar, just east of the Vidarbha area (now southernmost Madhya Pradesh) where they range from 21% in the south to only 0-4% as one approaches the Raypur plains. The south and southwest of Bastar is occupied by Gonds, all other regions by Chattisgarhi Hindi speakers. (For an overview of studies in (South) Indian place names see the paper by M.N. Nampoothiry 1987: 1-47, --- including a good bibliography, also of unpublished Indian theses).

### §13. The South

The South is frequently supposed to have been Dravidian from times immemorial. However, in the refuge area of Nilgiris with their isolated Drav. tribes (Toda, etc.), we find a substrate, see Zvelebil 1990, 63-70. Isolated words indicating this pre-Drav. substrate (Zvelebil 1990: 69f., Zvelebil 1979: 71f.) include the Irula words

- *mattu* 'lip',
- *ḍökēṇē*, *ḍēkēne*, *ḍēkena*, *ḍēkkada* 'panther',
- *ovarakāṅku*, *ōrakāṅku*, *ōraṅgeku*, *ōraṅge*, *ōrapodu* 'tomorrow' (unless DEDR 707 Tam. *urāṅku* 'to sleep'),
- *bundri* 'grass hopper' (unless DEDR 4169),
- *mutṭ(u)ri* 'butterfly' (unless DEDR 4850 miṭṭ 'locust'),
- *vutta* 'crossbar in a house'.

These instances should encourage Drav. specialists to look for substrates in Tamil, Telugu, Kannada, etc. However, just like the propagators of indigenous "Aryans" in the North, Dravidians of the South frequently think that they are autochthonous.

### §14. Vedda

Finally, in Sri Lanka, the remnant population of the Vedda now speaks Sinhala. (De Silva 1972).

The substrate that they may have preserved is in urgent need of thorough study, carried out by comparing Pali, Sinhala and Tamil words. Some typical words, interestingly many with geminates, that cannot be linked either to Sinhala or to Tamil are:

- *cappi* 'bird'
- *mundi* 'monitor lizard'
- *potti* 'a kind of bee'
- *panni* 'worm'
- *rukula* 'home, cavity'

(see de Silva 1972 : 16; his vocabulary, pp. 69-96, does not contain etymologies).

Finally there is Andamanese, but unlike the Austro-As. Nicobarese, so isolated that it can only be compared in long-range fashion, something entirely beyond my competence.

### §15. Indo-Iranian substrates in Central Asia and Iran



By way of addition, attention may be drawn to the northwestern borders of the subcontinent. Nuristani or Kafiri, as it was formerly called, is (differently from the older handbooks which lump it together with the Dardic branch of IA) a third branch of the Indo-Iranians (G. Morgenstierne, *Irano-Dardica*. Wiesbaden 1973). It has survived in the mountains of East Afghanistan and in neighboring Chitral (N.W. Pakistan). The Kalasha (Chitral) subgroup have even preserved their ancient non-Hindu and non-Iranian religion. Nuristani has preserved such sounds as IIr.  $\varepsilon$  that has been changed even in the RV  $> \acute{s}$  (c. 1500 BCE) and in Old Iranian  $> s$ . It has transmitted at least one loan word into Vedic, Nur. *\*kat's'a*  $>$  Ved. *kāca* 'shining piece of jewelry' (K. Hoffmann 1986, EWA I 335).

Beyond this, in Proto-IIr., there is a host of unstudied words found both in IA and Old Iranian that do not have an IE etymology and must represent old, Bactria-Margiana (BMAC culture 2100-1900 BCE), or other Central Asian substrate(s). They include plants, animals, and material culture, such as found in Ved. /Avestan:

- *uṣtra* / *uṣtra* 'camel', middle and new Akkadian *udru* "Bactrian camel" is a loan from Iran, see EWA I 238, KEWA III 652, cf. Diakonoff in *JAOS* 105, 1985, 600; the camel was introduced into the BMAC area from Central Asia only in the late 3rd mill. BCE;
- *khara* / *xara* 'donkey', cf. Toch. B *ker-ca-po*  $<$  *\*karca-bha?*, with the common Indian animal suffix *-bha* (as in *garda-bha*, *śara-bha*, *ṛṣa-bha*); the word ultimately may be a late 3rd mill. Near Eastern loan, cf. Akkadian (Mari) *hārum*, *ajarum* 'male donkey', EWA I 447. Note also the overlap with Dravidian (denied by EWA 473): Drav. *\*garda*  $>$  Tamil *kalutai*, etc., one of the few possible links of a Central Asian substrate with Dravidian (and with Vedic);
- *iṣṭi*, *iṣṭikā* / *iṣṭiia* 'brick', *zəmōištuua* 'clay brick'; OP *iṣṭi*, MP., NP. *xišt*; cf. Toch. *išcem* 'clay'?
- *sthūna* / *stūnā*, *stunā*, OP *stūnā* 'pillar', unless it belongs to Ved. *sthūra* 'tall, thick', Avest. *-stura*, Khot. *stura* (thus EWA II 768);
- *yavyā* / O.P. *yauviyā* 'channel',  $>$  MP., NP. *jō*, *jōy* 'stream, channel', Parachi *žī* 'rivulet', EWA II 405; both words, typical for loans, do not go back to exactly the same source;
- *godhūma* / *gantuma* 'wheat' from a Near Eastern language, cf. P-Semitic *\*ḥnt*, Hitt. *kant* and Egyptian *xnd* (EWA II 499, Kuiper *III* 34, 1991, 119)
- *parša* / *parša* 'sheaf', see EWA II 101;
- *bīja* / OIran. *\*bīza* (in names), 'seed, semen', Buddh. Sogdian *byz'k*, Parachi *biz* 'grains';
- *śaṇa* / *kana*- 'hemp', MP. *šan* 'hemp' (with northwestern interchange *k/ś*, see above), Khot. *kaṃha*, Osset. *gœn*, *gœnœ*, Russ. Church Slavic *konoplja*, Gr. *kánnabis*, itself a loan from Scythian, as also the early loans into Germanic (before *\*k*  $>$  *h*): Old High German *hanaf*, Dutch *hennep*  $<$  *\*kanap*;
- *bhaṅga* / *banga* 'hemp, hashish', if the word does not belong to *bhañj* 'to break';
- *\*sinšap* 'mustard': Ved. *saṣarpa* 'mustard', Khot. *śśaśvāna*, Parthian *šyṣ-d'n*, Sogdian *šywšp-δn*, MP. *span-dān* 'mustard seed'; Greek *sinapi*;  $<$  pre-Iran. *\*sinšapa*  $<$  *\*\*sinsap* (Henning *s1ens2ap*); cf. also: Malay *sawi*, *səsawi*, or Austro-As. *\*sapi*, *sV(r)-*

- sapi*; further EWA 712, 727: *śimśāpā* RV+ 'Dalbergia sissoo' NP. *šīšam*, Pashto *šəwa* < \**śīšampā*, CDIAL 12424), Elam. *še-iš-šá-ba-ut* = /*šeššap*/;
- *kaśyapa* / *kasiapa* 'turtle', Sogdian *kyšph*, NP. *kašaf*, *kaš(a)p* 'tortoise'; cf. Kashaf Rūd, a river in Turkmenistan and Khorasan;
  - *pard/pandh* 'spotted animal, panther' : Ved. *pr̥dāku* 'snake' RV, *pr̥dakū* AV, *pr̥dākhu* BSS (EWA II 163), with Para-Munda prefix *pər?*; Khovar *purdum* < \**pr̥dhūma?* KEWA II 335, CDIAL 8362; Bur. (Yasin) *phúrdum* 'adder, snake'; later Skt. 'tiger, panther'; NP. *palang* 'leopard' < O.Iran. \**pard-*, Greek *párdalis*, *párdos*, *léo-pardos* 'leopard' (EWA II 163), all < \*\**pard* 'spotted, wild animal?'; Henning reconstructs \*\**parθ* (but note Greek *pánthēr*), which may have been close to the Central Asian form;
  - \**kar(t)ka* 'rhinoceros', Ved. *khadga* 'rhinoceros' MS+, EWA 443, cf. N.P. *karka-dān*, Arab. *karkaddan*, Aelianus *kartázōnos* (\**kargazōnos*) 'Indian rhinoceros', all from a pre-Aryan source; however, cf. Kuiper 1948: 136 sqq.
  - *bheśaja* / *baēsaziia* 'healing'; Iir \**bhiš-aj* > Ved. *bhiš-aj*; the root \**bhiš* may be a loan word (cf. EWA s.v.),
  - *vīnā* 'lute': Ved. *vīnā* Khot. *bīna* 'harp, lute', Sogdian *wyn* 'lute', MP. *win* 'lute', Armenian *vin* 'lute', unless loans from India, cf. EWA II 568;
  - \**kapauta* 'blue': Ved. *kapota* 'pigeon', O.P. *kapauta* 'blue'; Khot. *kavūta* 'blue', MP. *kabōd* 'grey-blue', *kabōtar* 'pigeon'; EWA I 303, Kuiper 1991;
  - \**kadru* 'brown': Ved. *kadru* 'red-brown', *Kadrū* 'a snake deity', Avest. *kadruua.aspa* 'with brown horses', NP. *kahar* 'light brown';

The following words may be of still older origin and may have been taken over either in E. Europe or in Northern Central Asia:

- \**medh/melit* 'sweet, honey': IE. \**medhu* 'sweet' is found in Ved. *madhu* 'sweet, honey, mead', Avest. *mađu*, Sogdian *mδw* 'wine', (cf. Bur. *mel* 'wine, from grapes'), Toch. B *mit* 'honey', Gr. *méthhu* 'wine' etc.; it has spread to Uralic \**mese*, *mete*, Finnish *mete*, Hungarian *méz* 'honey'; Chin. *mi* < \**m<sup>i</sup>et*, Sino-Korean *mil*, Jpn. *mitsu* < \**mit(u)*; Iran. \**mađu* > Turkic, Mongolian *bal* 'honey'; Arabic *mādī?*, and to > Toch. B *mot* 'intoxicating drink'. --- From another source \**melit*, Greek *mélit-*, Hitt. *milit*, Latin *mel*, *mell-*, Gothic *milið*. In Nostratic (Illich-Svitych, *Opyt* II, Moskva 1976: 38sq.) both forms are united under \**majl*Δ > \*Uralic *majδ*'Δ, Drav. *mat̪*, *mit̪*, Altaic /*m/ala*, *bala*; cf. also, still further afield, in Polynesia: Samoan *meli*, Hawaiian *mele*, *meli*; *mele*, *melemele* 'yellow', Maori *miere*; Tongan *melie* 'sweetness, sweet, delicious', Rarotongan *meli* 'honey', Mangareva *mere* 'honey'.
- \**sengha/singha* 'lion' : Ved. *siṃha* 'lion' < \**sinj'ha* < \**sing'ha* differs from Proto-Iran. \**sarg*: Khorasmian *sary*, Parthian *šarg*, Khot. *sarau*; Henning reconstructs \*\**s<sub>1</sub>engha*; -- loans into nearby languages, such as Toch. A *śiśäk*, B *šecake* 'lion'; Tib. *seṅge*, Chin. \**suân-ŋei* (Henning, EWA), note, however, Karlgren 1923, no. 893 Arch. Chin. \*,*ši*, Jpn. \**si* > *shi(-shi)*; cf. perhaps Armenian *inc*, *inj* EWA II 727, KEWA III 447; the western IE languages have received the 'lion' word from a different source, Gr. *lis*, *leon(t)-*, Latin *leon-*.

In short, western and central Iran must have been inhabited by archaeologically well attested people of non-Ir speech. However, their languages have left few remains in Iranian. Apparently, Elamian was spoken up to Simaški (Kerman/Bandar Abbas area), while Aratta (Sistan) and Marhaši (W. Baluchistan, Bampur region) apparently had other language(s), (Vallat 1980); note also the loan word links between Sumerian and Drav. (above, §6.) All of these data need to be studied in greater detail, especially the early Ir substrate language(s).

## §16 Conclusion

In short, the early linguistic picture of South Asia in the second and first millennium BCE is as complex as (or even more so) than its modern counterpart. Some of the examples adduced above indeed indicate that we are in for surprises, once more information is received. The RV hapax *akhkhalī kr̥* 'to speak haltingly, to bleat' would have remained one -- if not for A. Mundlay's list in *MT* II, 17 with Nahali *akkal-(kāyni)* '(to cry) loudly in anguish'. This excludes other etymologies recorded in EWA, ingenious as they may be. This example also indicates that even the oldest literary tradition has retained important information on the (lost) substrates. What may we still find in the Tamil Sangam texts? Yet, as expressed above: nobody is looking!

On the other hand, it is important to know the location and time frame of the first occurrence of substrate words in order to evaluate them properly, and to avoid comparing accidental look-alikes by using derivatives that may have been possible, e.g., only a thousand years later. The Canaaitic words or Nahali *parayn* and *sanu* (see above) are cases in point. P. Benedict's warning (*MT* III: 93) on EFPs needs to be heeded.

The few etymological dictionaries available so far do not provide geographical and historical information, though Mayrhofer's EWA now gives a general idea, for the specialist, of the historical levels, but hardly of the geographical spread. DEDR does not have any such information yet, and we need to check the on-line dictionary at Cologne ([http://www.uni-koeln.de/phil-fak/indologie/tamil/otl\\_search.html](http://www.uni-koeln.de/phil-fak/indologie/tamil/otl_search.html)); and the KWIC Concordance of Classical Tamil texts (<http://www.uni-koeln.de/cgi-bin/SFgate>). A Munda etymological dictionary is still under preparation.

We need much more philological and linguistic study in a number of areas for further comparisons inside and outside South Asia. In this undertaking, the ancient Vedic and Tamil texts still hold out a lot of important and interesting data, but they have not yet been tapped properly. Even in the well-studied IA sector we do not yet have enough reliable information on the geographical spread and time frame of the texts (except for the Veda, see Witzel 1987, 1989, 1997). The various levels and the geography of the Pali and Epic texts still need much more sorting out. In the Dravidian field, we need, especially, a detailed historical grammar and dictionary of Tamil that covers the past two millennia or so in a comprehensive fashion. In Munda, a new reconstruction that pays more attention to S. Munda is eagerly awaited, not to speak of a comparative or etymological dictionary of the various languages and dialects involved. For the remnant languages such as Burushaski, Nahali, Kusunda, and the various substrates the lesser said the better. Even the extensive new Burushaski dictionary of Berger (1998) contains few etymological notes, and they are restricted to the northwestern languages and to Urdu. All major Indian languages, north or south, are lack historical and etymological dictionaries. Even in the well researched field of

Indo-Aryan, Turner's CDIAL and Mayrhofer's EWA are only of limited help for our purpose, restricted as they are, to words derived from OIA or reconstructable as OIA. Mayrhofer's 'unexplained, difficult, unclear' words and Turners 'starred' words may be a help, at least, to highlight possible loan and substrate words; even then, Mayrhofer's marked tendency to explain virtually everything as IE needs constant attention. Unfortunately, in similar vein, Burrow-Emeneau's DEDR only compares only inside Drav., and outside the family refers only IA and not to Munda or to other S. Asian languages, so that "their dictionary, by omitting all references to Munda, sometimes inevitably creates a false perspective from a Pan-Indic point of view" (Kuiper 1991: 53).

Even then, it is my hope that this brief survey will induce comparative linguists to pay closer attention to the rich materials found in the early Indian texts, and that even this still rather limited list will provide some useful materials for further study. More lists are in preparation.

In sum, not only is the linguistic situation of northern South Asia in the second millennium BCE much more complex than usually admitted, the materials adduced above also indicate that, even with the addition of the modern descendants of Proto-Burushaski, -Nahali and -Kusunda, we have to reckon with, and make use of, a number of other substrate languages such languages as Tharu, Masica's "Language X", the substrate of the Kathmandu Valley, and the Panjab and the Sindh varieties of the Indus language.

However, except for the few items pointed out for Vedda and the Nilgiri languages, the prehistoric linguistic situation of South India (before Dravidian) is entirely unclear: in this respect, a lot of spade work needs to be done by Dravidian specialists; the same applies to the reconstruction of Munda and the possible substrates of the eastern and central parts of India; yet, just as in the modern North Indian languages, no progress has been made in all these respects over the past few decades. The Himalayan languages that are finally studied in greater detail by the Linguistic Survey sponsored by the German Research Association and by the Himalayan Languages Project at Leiden, may still surprise us with remnants of pre-Tib.-Burm. substrates.

All of this, and to a small degree even the summaries of substrata given above, provide a multitude of data for the many waves of immigration and amalgamation that have swept over the Indian subcontinent. Ultimately, these substrates will hint at the first wave of immigrant groups of Homo Sapiens sapiens, which may have left us some remnants in the deep substratum of languages such as Nahali, Vedda and Kusunda.

\*\*\* \*\*

## 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         | Gr̥hyasūtra(s) (5)                              |
| Guj.       | Gujarati  |
| HŚS        | Hiranyakeśi 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         | Jaiminiya Brāhmaṇa (4 S)                        |
| Jpn.       | Japanese  |
| Kan.       | Kannada, Canarese                               |
| Kaśm.      | Kashmiri  |
| KaṭhĀ      | Kaṭha Āraṇyaka (4 W)                            |
| KauśS.     | Kauśika Sūtra (5 C)                             |
| KB         | Kauśitaki 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      | Ṣadviṃśa Brāhmaṇa (4 W)                   |
| Samh.      | Saṃhitā(s)                                |
| Sant.      | Santali                                   |
| ŚĀ         | Śāṅkhāyana Āraṇyaka (4 C)                 |
| ṢB         | Ṣadviṃśa 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.         | Videvdād  |
| VādhB      | Vādhūla Brāhmaṇa (Anvākhyāna) (4 C)                     |
| Ved.       | Vedic   |
| Ved. Index | Macdonell - Keith 1912                                  |
| VS         | Vājasaneyi Samhitā (2 E)                                |
| YV         | Yajurveda (-Samhitā) (2)                                |
| ZDMG       | Zeitschrift der deutschen morgenländischen Gesellschaft |

## BIBLIOGRAPHY

- Allchin, F. R. and N. Hammond, *The Archaeology of Afghanistan from the earliest times to the Timurid period*. London, New York: Academic Press 1978.
- 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
- Bagchi, P. C. (ed.), *Pre-Aryan and Pre-Dravidian in Sanskrit*. Calcutta : University of Calcutta 1929
- Bartholomae, Christian. *Altiranisches Wörterbuch*. repr. Berlin 1961
- C. Beckwith, *The Tibetan Empire in Central Asia*. Princeton University Press: 1987
- Bedigian, D. and J. H. Harlan. Evidence for the cultivation of sesame in the ancient world. *Economic Botany* 1985
- Bellezza, J.V. New Archeological Discoveries in Tibet. *Asian Arts*, 12/17/98, at: [www.asianart.com/index.html](http://www.asianart.com/index.html)
- Benedict, P. K. *Sino-Tibetan. A Conspectus*, Cambridge 1972
- , *Japanese/Austro-Thai*. Ann Arbor: Karoma 1990.
- Bengtson, J. Nihali and Ainu. *MT II*, 1996, 51-55
- Berger, H. Deutung einiger alter Stammesnamen der Bhil aus der vorarischen Mythologie des Epos und der Purāṇa. *WZKSOA* 3, 1959, 34-82
- , review of: K. H. Pinnow, Versuch einer historischen Lautlehre der Kharia-Sprache. (Wiesbaden 1959), *ZDMG* 112, 1963, 416-421
- , *Die Burushaski-Sprache von Hunza und Nager*. Wiesbaden: Harrassowitz, 1998.
- Bhattacharya S. Some Munda etymologies, in: N. H. Zide (ed.), *Studies in Comparative Austroasiatic Linguistics*. London, The Hague: Mouton 1966 : 28-40
- Bista, D. B. Encounter with the Raute: The last hunting nomads of Nepal. *Kailash* 4, 1976, 317-327
- Blažek, V. and C. Boisson, The Diffusion of Agricultural Terms from Mesopotamia. *Archív Orientalní* 60, 1992, 16-37
- Bomhard, A. On the Origin of Sumerian. *MT III*, 1997, 75-92
- 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
- , *The Sanskrit Language*. (3rd ed.) London.
- , 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)
- , *Dravidian Borrowings from Indo-Aryan*. Berkeley : University of California Press 1962

- Cavalli-Sforza, L.L., P. Menozzi, A. Piazza. *The history and geography of human genes*. Princeton: Princeton University Press 1994.
- and F.Cavalli-Sforza, *The Great Human Diasporas. The History of Diversity and Evolution*. Reading MA : Helix Books 1995
- Crooke, W. *The Tribes and Castes of the North-west provinces and Oudh*. Calcutta: Office of the Superintendent of Government Printing 1906
- Das, R. P. The hunt for foreign words in the R̥gveda. *IJ* 38, 1995, 207-238
- De Silva, M.W. Sugathapala, *Vedda language of Ceylon; texts and lexicon*. Münchener Studien zur Sprachwissenschaft. Beiheft n.F. 7. München: R. Kitzinger, 1972
- Diakonoff, I. M. Hurro-Urartian Borrowings in Old Armenian. *JAOS* 105, 1985, 597- 604
- , External Connections of the Sumerian Language. *MT* III, 1997, 54-62
- 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
- Deshpande, M. M. and P.E. Hook (eds.), *Aryan and Non-Aryan in India*, Ann Arbor: Center for South and South-East Asian Studies, University of Michigan 1979
- Emeneau, M. B. India as a linguistic area. *Language* 32, 1956, 3-16
- , and Th. Burrow, *Dravidian Borrowings from Indo-Aryan*. Berkeley : University of California Press 1962
- 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 R̥gveda: 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 R̥gveda. 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.
- Funayama, T. Remarks on Religious Predominance in Kashmir; Hindu or Buddhist? In: Y. Ikari, *A study of the Nilamata*. Kyoto 1994, 367-375
- Fussman, G. *Atlas linguistique des parlers dardes et kafirs*. Paris 1972
- Gardner, J. R. *The Developing Terminology for the Self in Vedic India*. Ph.D.Thesis, Iowa U., 1998
- Gardner, P. Lexicostatistics and Dravidian Differentiation *in situ*. *Indian Linguistics* 41, 1980, 170-180
- Gening, V.F. Mogil'nik Sintashta i problema rannikh Indoiranskikh plemen. *Sovietskaya Arkheologiya* 1977, 53-73
- Glover, L.C. and Higham, C.F.W. New evidence for early rice cultivation in South, Southeast and East Asia. In: D. R. Harris (ed.), *The origins and spread of agriculture and pastoralism in Eurasia*. London: UCL Press 1996: 413-441
- Gordon, K. H. *Phonology of Dhangar-Kurux*, Kathmandu 1976
- Grierson, G. *Linguistic Survey of India*. Calcutta: Office of the superintendent of government printing, India 1903-22 (repr. Delhi 1967)
- Haudricourt, A.G. *Daïque (Daic)* In: Shafer 1964, 453-525
- Heine-Geldern, R. *Das Dravidaproblem*. Anzeiger der Österreichischen Akademie der Wissenschaften, Jg. 1964, no. 9, Wien 1964: H. Böhlau, pp. 187-201
- Hinüber, O. v. Die Kolophone der Gilgit-Handschriften, *Studien zur Indologie und Iranistik* 5/6, 1980, 49-82
- , Brāhmī inscriptions on the history and culture of the upper Indus valley. In: K. Jettmar et al., *Rock Inscriptions in the Indus Valley. Antiquities of Northern Pakistan. Reports and Studies*, vol. 1. Mainz: Philipp von Zabern 1989: 41-72
- Hock, H.H. Substratum influence on (Rig-Vedic) Sanskrit? *Studies in the Linguistic Sciences* 5, 1975, 76-125
- , [On Bangani] <http://www.personal.umich.edu/~pehook/bangani.html>
- Hodgson, B. H. On the Chépáng and Kúsúnda tribes of Nepál. *JASB* 17, 1848, 650-58



- , Comparative Vocabulary of the languages of the broken tribes of Népal, *JSAB* 22, 317-427 = B. H. Hodgson, *Miscellaneous Essay relating to Indian subjects*. Vol. 1. London: Trübner 1880, 171-175. >>
- Hoffmann, Karl. *Die alt-indoarischen Wörter mit -ṇḍ-, besonders im Rgveda*. 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
- Illich-Svitych, V. M. *Opyt sravneniya nostraticheskikh yazykov*. II, Moskva: Nauka 1976
- 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.
- Jettmar, K. et al., *Rock Inscriptions in the Indus Valley. Antiquities of Northern Pakistan. Reports and Studies*, vol. 1. Mainz: Philipp von Zabern 1989
- Jørgensen, H. *A dictionary of the Classical Newari*, København 1936
- Joshi, S.M. (ed.) *Paryācavācī Śabda Koś*, Kathmandu : Nepāl Rājakīya Prajñā-Pratiṣṭhān VS 2030 (1974)
- Kajale, M. D. Current status of Indian palaeoethnobotany: introduced food plants with a discussion of the historical and evolutionary development of Indian agriculture and agricultural systems in general. In: J. Renfrew (ed.) *New Light on Early Farming*. Edinburgh: Edinburgh University Press 1991, 155-189.
- Karlgren, B. *Analytical Dictionary of Chinese and Sino-Japanese*. Paris 1923
- Kenoyer, J. M. *Ancient Cities of the Indus Valley Civilization*. Oxford: Oxford University Press/American Institute of Pakistan Studies 1998
- Konow, S. On some facts connected with the Tibeto-Burman dialect spoken in Kanawar, *ZDMG* 59, 1905, 117-125.
- Koppers, W. *Die Bhil in Zentralindien*. Horn: F. Berger 1948
- Krauskopf, G. *Maîtres et possédés; Les rites et l'ordre social chez les Tharu (Népal)*. Paris : Editions du Centre National de la Recherche Scientifique 1989
- Krishnamurti, Bh. Dravidian personal pronouns, *Studies in Indian Linguistics*, Poona and Annamalainagar, 1968, 189-205
- , Comparative Dravidian Studies since Current Trends 1969. In: V. Z. Acson and R.L. Leed (eds.), *For Gordon Fairbanks*. Honolulu: Univ. of Hawaii Press 1985, 212-231
- Kuiper, F.B. J. Beiträge zur altindischen Wortforschung. *tālah* "Geklatsch". *Zeitschrift für Indologie und Iranistik* 8,1931, 250-251
- , Ai. Mandākinī 'EN. verschiedener Flüsse'. *Acta Orientalia* 17, 1939, 17-20 = 1997: 3-6
- , *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. *IJJ* 10, 1967, 81-102
- , *Proto-Munda words in Sanskrit*. Amsterdam 1948
- , 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. *IJJ* 38, 1995, 239-247
- , *Selected writings on Indian Linguistics and Philology*. A. Lubotsky, M.S. Oort and M. Witzel (eds.). Amsterdam-Atlanta : Rodopi 1997
- Laufer, B. Loan Words in Tibetan, *T'oung Pao* 17, 1916-18, 403 ff.
- Leal, D. *Chitwan Tharu Phonemic Summary*. Kirtipur: Summer Inst. of Linguistics 1972,
- Lévy, S. Pré-Aryen et pré-Dravidien dans l'Inde. *Journal Asiatique* 203, 1923, 1-57 [transl. in: Bagchi 1929, 63-126]
- Littauer, M. and Crouwel, J. H. *Wheeled vehicles and Ridden Animals in the Ancient Near East*, Leiden : Brill 1979
- Lorimer, David L. R. *The Burushaski language*. Oslo : H. Aschenhoug 1935-38.

- , Burushaski and its alien neighbours: Problems in linguistic contagion. Transactions of the Philological Society 1937, 63-98
- Majumdar, D. N. *The Fortunes of Primitive Tribes*. Lucknow 1944
- MacDonell, A. A. and A.B. Keith, *Vedic Index of Names and Subjects*, Delhi 1967 [repr. of 1912]
- Malla, K. P. Linguistic Archaeology of the Nepal Valley. A Preliminary Report. *Kailash* 8, 1981, 5-23
- Manandhar, Th. L. *Newari-English Dictionary. Modern language of Kathmandu Valley*, ed. by Dr. Anne Vergati. Delhi: Agam Kala Prakashan 1986
- Masica, C. P. *Defining a Linguistic Area. South Asia*. Chicago: The University of Chicago Press 1971
- , 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, David W., Elamite and Dravidian: Further evidence of relationship. (With discussion by M.B. Emeneau, W.H. Jacobsen, F.B.J. Kuiper, H.H. Paper, E. Reiner, R. Stopa, F. Vallat, R.W. Wescott, and a reply by McAlpin). *Current Anthropology* 16, 1975, 105-115
- , 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. Preface to Lorimer 1935, vii-xxx
- , Notes on Burushaski phonology. *Norsk Tidsskrift for Sprogvidenskap* 14, 1947, 61-95
- , *Irano-Dardica*. Wiesbaden 1973
- Mundlay, Asha. Nihali lexicon. *MT* II, 1996, 17-40
- Nampoothiry, M.N. Indian Toponymy. A critical evaluation of the work done in this field in India with a bibliography in: P. Ramachandran 1987: 1-47
- Nurse, D. A Hypothesis of the origin of Swahili. *Azania* 18, 1983, 127-150.
- Oberlies, Th. *Review Article: F.B. J. Kuiper: Aryans in the Rigveda*. *IJ* 37, 1994, 333-349.
- 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
- 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
- Pokorny, J. *Indogermanisches etymologisches Wörterbuch*, Bern/München 1959
- Possehl, G. Meluhha. in: J. Reade (ed.) *The Indian Ocean in Antiquity*. London: Kegan Paul Intl. 1996a, 133-208
- , *Indus Age. The writing System*. Philadelphia: University of Pennsylvania Press 1996b
- , 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
- Przyłudski, J. *Further Notes on Non-Aryan Loans in Indo-Aryan*, in Bagchi 1929 : 145-149
- Ramachandran, Puthusseri and K. Nachimuthu (eds.) *Perspectives in Place Name Studies : Proceedings of the National Seminar on South Indian Place Names, Held at Trivandrum on 21-23 June 1985. A Festschrift to Prof. V.I. Subramoniam, On His Sixtieth Birth Day*. Trivandrum: Place Name Society, 1987

- Randhawa, M. S. *A history of agriculture in India*. New Delhi : Indian Council of Agricultural Research 1980-1986.
- Rau, W. The Earliest Literary Evidence for Permanent Vedic Settlements. In: M. Witzel (ed.), *Inside the Texts, Beyond the Texts. New Approaches to the Study of the Vedas*. Harvard Oriental Series. Opera Minora, vol. 2. Cambridge 1997, 203-206
- Reinhard, J. Aperçu sur les Kusundā. *Objets et Mondes* 9, 1969, 89-106
- , The Raute: Notes on a nomadic hunting and gathering tribe of Nepal. *Kailash* 2, 1974, 233- 271
- Reinhard, J. and Tim Toba. *A preliminary linguistic analysis and vocabulary of the Kusunda language*, Kirtipur: Summer Institute of Linguistics and Tribhuvan University 1970
- 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
- Schmid, L. Report on a survey of Dardic languages of Kashmir. *Indian Linguistics* 42, 1981, 17-21
- Shafer, R. Nahāli, A linguistic study in paleoethnography. *Harvard Journal of Asiatic Studies* 5, 1940, 346-371
- , *Ethnogeography of Ancient India*, Wiesbaden 1954
- , *Introduction to Sino-Tibetan*. Wiesbaden 1966-7.
- Shaffer, J. G. and Diane A. Lichtenstein. The concepts of "cultural tradition" and "palaeoethnicity" in South Asian archaeology. In: G. Erdosy (ed.) 1995, 126-154.
- Sharma, D.D. Old Indo-Aryan element in Kinnauri. *Dr. B.R. Sharma Felicitation Volume*, Tirupati: Kendriya Sanskrit Vidyapeetha 1986, 149-155
- Southworth, F. Linguistic stratigraphy of north India. In: F. Southworth and M.L. Apte (eds.), *Contact and Convergence in Indian Languages*, International Journal of Dravidian Linguistics 3, 1974, 201-223
- , 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, 559-668
- , 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
- Starostin, S.A. *Rekonstruktsiya drevnekitaiskoi fonologicheskoi sistemy*, Moscow 1989
- Szemerényi, O., *Einführung in die vergleichende Sprachwissenschaft*. Darmstadt 1970
- Tamot, K. and I. Alsop, The Kushan-period Sculpture from the reign of Jaya Varma-, A.D. 185, Kathmandu, Nepal: *Asian Arts*, July 10, 1996, at: [www.asianart.com/index.html](http://www.asianart.com/index.html)
- Thomas, F.W., The Zan Zun language, *JRAS* 1933, 405-410
- Tikkanen, B. On Burushaski and other ancient substrata in northwest South Asia. *Studia Orientalia* (Helsinki), 64, 1988, 303-325
- Turner, R. L. *A comparative Dictionary of the Indo-Aryan Languages*. London 1966 (CDIAL)
- Tyler, Stephen, Dravidian and Uralian: the lexical evidence. *Language* 44, 1968, 798-812.
- van Driem, G. and Suhnū Rām Sharmā. In Search of Kentum Indo-Europeans in the Himalayas. *Indogermanische Forschungen* 101, 1996, 107-146
- , Some Grammatical Observations on Baṅgānī. *Indogermanische Forschungen* 102, 1997, 179-198 (cf. <http://iiias.leidenuniv.nl/host/himalaya/individ/kirmor.html>)
- Vallat, F. *Suse et Elam*. Paris : Editions ADPF 1980
- , *Éléments 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]
- P. Whitehouse, P. The External Relationships of the Nihali and Kusunda Languages. *MT III*, 1997, 4-44
- Witzel, Michael, On the location of the Licchavi Capital of Nepal. *Festschrift für P. Thieme* (= *StII* 5/6) 1980, pp. 311-337
- , Zu den Namen einiger vedischer Schulen. *StII* 10, 1983/85, 231-237

- , On the localisation of Vedic texts and schools (Materials on Vedic Śākhās, 7). G. Pollet (ed.), *India and the Ancient world. History, Trade and Culture before A.D. 650. P.H.L. Eggermont Jubilee Volume.* Leuven 1987, 173-213
- , 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.
- , The Brahmins of Kashmir. In: Ikari, Y. (ed.) *A study of the Nilamata - Aspects of Hinduism in Ancient Kashmir.* Kyoto: Institute for Research in Humanities, Kyoto University 1994, 237-294
- , 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
- , R̥gvedic history: poets, chieftains and polities. 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, 307-352.
- , 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
- , The Development of the Vedic Canon and its Schools: The Social and Political Milieu. (Materials on Vedic Śākhās 8). In: *Inside the Texts, Beyond the Texts. New Approaches to the Study of the Vedas.* Harvard Oriental Series. Opera Minora, vol. 2. Cambridge 1997b, 257-345
- , 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
- Zide, A. and N.H. Zide, Semantic reconstruction in proto-Munda cultural vocabulary. *Indian Linguistics* 34, 1973, 1-24
- Zide, N. H. Munda and non-Munda Austroasiatic Languages, in: *Current Trends in Linguistics*, 5. The Hague: Mouton 1969, 411-430
- , Proto-Munda cultural vocabulary: evidence for early agriculture. In: Ph. N. Jenner et al., *Proceedings of the First International Austroasiatic Conference.* Honolulu: University of Hawaii Press 1976, 1295-1334
- , On Nihali. *MT II*, 1996, 93- 100
- Zoller, C. P. Bericht über besondere Archaismen im Bangani, einer Western Pahari-Sprache. *Münchener Studien zur Sprachwissenschaft* 49, 1988, 173-200
- , Bericht über grammatische Archaismen im Bangani, *Münchener Studien zur Sprachwissenschaft* 50, 1989, 159-218
- Zvelebil, K. The descent of the Dravidians. *International Journal of Dravidian Linguistics* 1-2, 1972, 56-63
- , Review of McAlpin 1981, *JAOS* 105, 1985, 364-372
- , *Dravidian Linguistics: an Introduction.* Pondicherry: Pondicherry Institute of Linguistics and Culture 1990