SPECIAL ISSUE (MT-SPECL)
OCTOBER 1999

By the Association for the Study of Language In Prehistory
Meanwhile, offer a new starting point for attempts at decipherment. Even the ‘Para-Munda’ thesis, whilst it has yet to convince me, deserves to be tried on the Indus inscriptions. There is no a prior reason why an Austroasiatic language can not have been spoken in the Indus Valley.

Unfortunately, without the full list of substrate words and their meanings I am in no position to offer an alternative affiliation for them. This is equally true of the other intriguing cases Witzel raises, such as Trata, Vedda and Khobar. Of the handful of words listed I was unable to match a single one to any other family — and there are two compelling reasons why I might not have been any more successful even with the full lists.

There are two ways in which evidence for the original languages of South Asia may have been preserved. Firstly, the languages themselves may survive. Nihali and Kusunda are examples of this, and in each case it is the core vocabulary (body parts etc.) that survives beneath layers of more recent borrowing. It is within this core vocabulary that the evidence for deeper relationships may be sought, as I attempted to show in my recent article (Mother Tongue Ill).

Secondly, aboriginal communities may adopt the languages of more recent arrivals but incorporate some elements of their ancestral languages in the process. In these cases, however, the core vocabulary of the resulting language is that of the recent arrivals, while the retainments are limited both in number and in range. Generally the latter are cultural items reflecting different stages of lifestyle or age. This accounts for the presence of the non-Indo-European agricultural vocabulary in cultural items are also the items most prone to borrowing. Witzel presents some very interesting historically very important (particularly what it says about the coming of the Dravida), but what it doesn’t do is illuminate the wider genetic relationships of the languages concerned.

The second serious limitation on our knowledge of South Asian linguistic prehistory is the complexity of the pre-agricultural scene. Though Witzel himself also uses the word ‘complex’ to describe this, I believe he underestimates the likely degree of that complexity. It is inconceivable that the precursors figures would be twice that. Such a time scale allows for extreme diversity to develop, and I consider it that found in New Guinea — but on a larger scale. In other words, there could have been hundreds of scores of similarly disparate families. This means that the chances of a substrate language being closely related to any other surviving ‘Old Indian’ language are extremely slim, while the chances of being able to identify more distant relationships are, in the absence of residual core vocabulary, slimmer yet. It further means that for the most part the earliest linguistic prehistory of South Asia is lost to us for ever.

But we must not allow such pessimistic considerations to prevent us looking. Witzel mentions several languages in which substratum vocabulary has been identified, and gives ground for hope that there may be many more. As long as we have the data (preferably complete lists of words and their meanings) there is hope that some glimmer may yet reach us from the depths of this otherwise impenetrable darkness. This is important because, by virtue of its geographical position, South Asia and its languages are crucial to our understanding of the Human language family as a whole.

On the Austroasiatic Indus Theory

George van Driem
Leiden University

The problem in a nutshell

The idea that the ancient Austroasiatic homeland lay somewhere in India is an old one. Linguistic facts as Austroasiatic have always supported an Indian homeland for Austroasiatic. First of all, as Heinz-Jürgen Pinnow observed, ‘the Munda languages undoubtedly are more similar to Proto-Austroasiatic than the other members of the family’ (1963: 150), which suggests that the Munda peoples, who reside in India, have been subjected to less upheaval through migration than have other Austroasiatic groups, such as the Nicobarese and the various Mon-Khmer groups. Secondly, toponymical evidence has been adduced in support of the hypothesis that the Austroasiatic Urheimat lay in South Asia or, at least, that the Austroasiatic linguistic area was once far more widespread in South Asia than it is today. Toponyms and particularly river names in the Himalayan region, such as Gandaki which may derive from a Munda word for ‘river’, have suggested to researchers such as Hermann Berger and Manfred Mayrhofer that Austroasiatic is an old ethnic substrate in the north of the Indian Subcontinent and that its presence antedates the advent of Tibeto-Burman peoples in the Himalayan region. Thirdly, Przyluski (1922, 1923), Lévi (1923), Bloch (1925, 1930), Lévi, Przyluski and Bloch (1929) and Kuiper (1948, 1950, 1954, 1955, 1991) advocated the idea that an Austroasiatic substrate existed in Vedic Sanskrit, but perhaps not all of the evidence adduced to date is equally cogent (cf. Emeneau 1954: 291-292).

The competing theory that the Austroasiatic homeland lay in Southeast Asia was put forth by Robert von Heine-Geldern, whose theory was an interpretation of archaeological findings based on anthropological findings and the modern geographical distribution of Austroasiatic speaking peoples. He interpreted the Munda peoples as the result of ‘die Einwanderung mongoldier austroasiatischer Sümme in Vorderindien’ and of their ‘Mischung mit Dravida und Urbevölkerungselementen’. The original inhabitants of India were a ‘mehrassige Urbevölkerung’ which inhabited the Subcontinent in palaeolithic times (1928, 1932). This theory has remained influential to the present day. It must be kept in mind, however, that, in interpreting the archaeological record with the idea of reconstructing an ancient linguistic intrusion, the linguistic evidence holds primacy above the archaeological evidence.

In his fascinating and exciting paper ‘Early sources for South Asian substrate languages’, Michael Witzel provides new evidence for a special variant of the Indian homeland hypothesis and even goes as far as to suggest that the people behind the Indus Valley civilisation could have been Austroasiatic. In so doing, Witzel suggests that the old theory of a Munda substrate in the Himalayas has been denied by various people including myself (p. 48). This is not altogether precise. What I disbelieve is that the verbal agreement morphology observed in Tibeto-Burman languages of the Himalayas, such as the Kiranti languages in eastern Nepal, the recently discovered Gongduak language in central Bhutan, Dimal in the Nepalese Terai, etc., can be attributed to a Munda substrate, as some scholars had been inclined to believe ever since Wilhelm Schmidt misidentified such languages as ‘tibetoburmanisch-austroasiatische Misc-begriffe’ (1906). James John Bauman’s study definitively put this idea to rest, at least as far as Kiranti verbal flexion is concerned (1975). Subsequent studies have borne out that the desinences and individual agreement etyma, a subset of which are grammaticalized pronominal elements, are reconstructible to the Tibeto-Burman level and demonstrably native to that family in the sense...
that they are well reflected in far-flung branches of the family, even to the northeast of the Himalayan divide, e.g. van Driem (1990, 1991a, 1991b, 1992, 1993a, 1993b, 1993c, 1994a, 1994b, 1995, 1997b, 1997c), Rutgers (1993), Kepping (1994), Turin (1998). In passing, I cannot help but wonder from which dialect Witzel draws his modern Tibetan forms, e.g. [ye] ‘eight’ for literary Tibetan brygyad (p. 39), since he provides a proper transliteration of the native spelling, this point is not material to his argument.

It is fair and fitting that I state my own point of view on the subject at the outset. Although I believe in the antiquity of an Austroasiatic presence in the north of the Subcontinent, I do not currently subscribe to the hypothesis of an Austroasiatic Indus. None the less, I strongly feel that Michael Witzel’s new evidence for the Indian homeland hypothesis for Austroasiatic merits serious consideration and that his Austroasiatic Indus is a highly intriguing hypothesis. His past studies on toponyms, particularly on Himalayan hydronymy are of the greatest interest in this context. Last month I had occasion to discuss Michael Witzel’s fascinating article with my friend Asko Parpola, whom I visited in Kyōto. Parpola, author of the lovely study Deciphering the Indus Script (Cambridge, 1994), has given much thought to the material presented by Witzel, and he kindly shared his thoughts with me in Japan. Parpola’s objections deal mainly with the weakness of a number of the etymologies proposed by Witzel, and I shall not deal with this topic here.

Witzel points out that three chronological layers can be identified in the Rgveda. This division into an ‘early’, ‘middle’ and ‘late’ layer is well accepted. Roughly speaking, the classical division is that Books 2 through 7 and Book 9 represent the old portion. Book 8 and the first part of Book 1 represent the middle layer, followed by the second part of Book 1. Book 10 represents the late layer. On the basis of recent investigations, Witzel has now assigned Books 3 and 7 to the middle layer, and he has reassigned to the late layer the first part of Book 1, the second part of Book 8 and the late accretion in the first part of Book 8 running from lines 49 through 59. Witzel remains tacit on the status of Book 9. It is generally accepted that editorial revisions to the Rgveda, particularly in composition, were introduced later, probably on the Gangetic Plain, where later texts such as the Brāhmaṇas were most probably composed. Witzel points out that there is little Dravidian in the oldest layer of the Rgveda, which he dates to the period 1700-1500 BC, perhaps arguably none, and that Dravidian loans are only to be found in the later strata of the Rgveda.

The crux and simultaneously the Achilles’ heel of Witzel’s argument is that he maintains that the Rgveda was first composed in the Punjab and later on the Gangetic Plain. The idea that much of the Rgveda took shape in the Punjab is relatively well accepted because the geography reflected in the text involves rivers flowing from north to south. However, a good number of scholars, including Asko Parpola, believe that the oldest layers of the Rgveda were composed in more northerly areas, perhaps even as far north as modern Afghanistan. Witzel’s argument hinges upon his ability to convincingly demonstrate that the oldest hymns of the Rgveda were composed in the Punjab and not in more northwesterly parts of what today is Pakistan and Afghanistan. Little new compelling evidence has been adduced to substantiate this claim. The presence of a non-Dravidian non-Aryan component in the oldest layer of the Rgveda has long been recognized, and Witzel, with his mastery of a large corpus of data, drives home this point strongly. This fact does not, however, necessarily militate against the conventional theory of a Dravidian Indus.

Conventionally, the ancient Indo-Iranians are identified with the Andronovo culture, a blanket term for a number of similar local cultures which occupied the entire west Asiatic steppe from the Ural river to the Yenisei between 2000 and 900 BC. The split between the Indo-Aryans and the Proto-Iranians is believed to originally have been a north-south split with the Proto-Iranians to the north of the Indo-Aryans, who led the vanguard south down through Bactria and Margiana. Thence the Indo-Aryans spread both east into the Indus Valley as well as west to the Fertile Crescent, where they became the Mitanni ruling class of an ancient kingdom in the Jazīrah in upper Mesopotamia in the xvth and xivth centuries BC, and succeeded in imposing their Indo-Aryan religion and deities upon their Hurrian subjects, although it was the language of the subjected Hurrians which ultimately prevailed. In the east, the Indo-Aryans were to be more successful, and the languages which derive from their original tongue are today the major languages of northern India, Bangladesh and Pakistan. The Indo-Aryans were followed by the Proto-Iranians.

Three weaknesses and an alternative explanation

I believe that there are at least three reasons why the facts adduced by Witzel more aptly argue against an Austroasiatic Indus and for a Dravidian Indus. First of all, Alexander Lubotsky has recently examined lexical items which qualify as Indo-Iranian isolates in that they are attest ed in Iranian and Indo-Aryan but in no other branches of Indo-European. This lexical component in Indo-Iranian is evidently a loan layer characterized by a shared phonological and morphological shape which is uncharacteristic for words of Indo-European stock, such as the presence of voiceless aspirates and long middle syllables. These loans show sound correspondences which are, in part, irregular between Iranian and Indo-Aryan. This points either to earlier borrowing of these items between Indo-Aryan and Iranian or to borrowing of these items from the substrate language into both Indo-Aryan and Iranian at different times. These Indo-Iranian isolates appear at this early stage of investigation to correspond to the old non-Dravidian loan layer in the Rgveda. Lubotsky’s findings indicate that this oldest loan layer must therefore antedate the arrival of the Indo-Aryans on the fluvial plains of the Indus and the large, largely dry river bed of the Ghaggar-Hakra. The ancient language from which this substrate layer was borrowed was probably spoken on both sides of the Hindu Kush and may, in fact, represent the language which the Indo-Iranians encountered as they descended from the steppes onto the people who inhabited the Bactria Margiana Archaeological Complex. Significantly, some of the isolated Indo-Iranian vocabulary is religious in nature and includes terms dealing with the soma or Ṛṣabha cult, which Viktor Sarianidi claims was a feature of the urban civilisation of Margiana (1991, 1998a, 1998b). Lubotsky’s findings suggest that the ancient Indo-Iranians were first infected with their religion or, at least, with certain key elements thereof as they passed through Bactria and Margiana on their march to the south.

Secondly, the sociolinguistically most obvious and foremost thing for a conquering people to do is not necessarily to borrow words extensively from the language of a subjugated populace. Borrowing is more likely to have increased as the subjugated populace was assimilated and a form of coexistence and peaceful interaction had come into being. Although some borrowing may already have occurred in the earliest phases of contact, it is reasonable to assume that it would have taken some time for such alien words to enter the elevated, formal language of native oral tradition. If the hypothesis of a Dravidian Indus is correct, then the Indo-Aryans would not necessarily have encountered the Dravidians until they had descended from the mountains and actually entered the Punjab in what today is Pakistan, and the later date of the Dravidian loans is precisely what we should expect.

Thirdly, any solution to the Austroasiatic homeland problem must satisfy several criteria. One of these is what Jim Mallory calls the ‘total relationship’ principle, whereby the origins for any single Austroasiatic group cannot be resolved independently of other Austroasiatic groups. This criterion is often overlooked in the case of Austroasiatic, for scholars have put the Austroasiatic homeland as far east as the eastern seaboard of China and as far west as the Punjab. The origins of Munda cannot be resolved without taking into account the linguistic ancestors of the
Nicobarese and speakers of Mon-Khmer languages as far flung as Vietnamese. The Punjab is not only far away from the geographical centre of gravity of modern Austroasiatic language communities, the Punjab in the far northwest is beyond the range of any modern or historically attested Austroasiatic language community. The present distribution of Austroasiatic language communities makes an early Austroasiatic presence in the north of the Subcontinent plausible, but the distribution of Mon-Khmer language communities in mainland Southeast Asia (including Khasi in the Meghálaya) and Nicobarese in the Andamans suggests that an Austroasiatic homeland lay in the northeast along the Brahmaputra and around the Bay of Bengal, perhaps extending as far west as the Gangetic Plain. The presence of Aryanized ‘scheduled castes’ and possibly originally Austroasiatic groups such as the Bhils, Tháru and Musáhár could suggest an Austroasiatic presence even further west, but the original linguistic affinity of these groups is, to be precise, unknown. A real problem is that there are too many miles and too many peoples between the Hindu Kush and the Bay of Bengal, and the distinctness of the neolithic assemblage in eastern and northeastern India and the neolithic traditions elsewhere in the Subcontinent is a very well-established in Indian archaeology.

In summary, some indologists hold that the oldest layers of the Rigveda were composed in areas to the northwest of the Punjab, and compelling evidence has not yet been adduced to demonstrate that this is not the case. Labotsky’s findings indicate that the non-Dravidian loan layer in the Rigveda is too early to be traced to the Punjab and that the same source language is already reflected in Indo-Iranian. The fact that Dravidian loans are to be found in the later layers of the Rigveda is precisely what we should expect if we entertain the hypothesis of a Dravidian Indus. Finally, the geographical distribution of Austroasiatic language communities and the well-established distinctness of the neolithic assemblage in eastern India from the neolithic traditions elsewhere in the Subcontinent renders the hypothesis of an Austroasiatic Indus implausible. In this connection, Witzel’s para-Munda source becomes meaningful, for the hypothesis of a lost western branch of Austroasiatic is an intriguing possibility. However, much is contingent on the soundness of the proposed etymologies and on how much leeway is permitted by the necessarily nebulous nature of para-Munda. In view of the ethnonomological composition of the north of the Subcontinent, Kusunda or para-Kusunda might be a more obvious place to look for the source of early loan layer in Vedic. On the other hand, it may be that the language from which the early loan layer entered into Vedic and, for that matter, Indo-Iranian may have been lost forever in the sands of Bactria and Margiana. Therefore, even if Witzel’s Austroasiatic Indus theory is incorrect, the large body of analysed data which he has adduced will be of lasting value to Vedic studies.

Related issues

Finally, I shall address a number of issues germane to Witzel’s argument. The first and foremost question is just how Austroasiatic is the source language for the early borrowings seen in the Rgveda and — as Alexander Labotsky has recently argued — in Indo-Iranian. We know that Frans Kuiper already thought that items in the early loan layer exhibited elements which he identified with Austroasiatic prefixes, only relics of which he believed could be found in Munda but many of which were still found intact in Malay. Here we have arrived at another core problem, for which Michael Witzel can in no way be held accountable, i.e. the state of the art in Austroasiatic linguistics.

Malay is one of many Austroasiatic languages, whereas the Munda, Nicobarese and Mon-Khmer languages together make up the Austroasiatic language family. Wilhelm Schmidt is the father of the old Austric theory, which postulated a language family consisting of Austroasiatic and Austroasian. Schmidt’s Austric was a very inclusive group, which later even included Japanese as a predominantly Austric ‘Mischsprache’ consisting of an ‘austroasiatische’ and an ‘ural-altaische’ layer (1906, 1930). The late Paul Benedict, who himself still entertained the Austric theory during the war, facetiously pronounced this proto-language ‘extinct’ in 1991, but his pronouncement was premature. Gérard Diffloth has found that the lexical evidence for Austric is largely negative (1994), but Lawry Reid has kept the idea alive by advancing a meagre but tantalizing handful of Austric morphemes (1994). Reid relates the Proto-Austroasiatic causative morphemes *<pa- > - *<ka- > to the Proto-Austroasian causative prefixes *<p- > - *<k- >, and *<paka- >, the Proto-Austroasian agitative marker *<pma- > - *<pma- >, the Proto-Austroasian instrumental agentive *<pma- > - *<pma- >, and the Proto-Austroasian instrumental infixes *<pa- > - *<pma- > and *<ins- > to the Malay-Polynesian instrumental prefix *<pna- > and Proto-Austroasian instrumental morpheme *<p= - *<pma- >. Finally, Reid proposes a not very convincing correspondence between a Nancowry Nicobarese nominalizer suffix *<a- > and a Proto-Austroasian ‘objective’ suffix *<o- >. There are several problems with the proposed morphological parallels: Most Austroasiatic languages are grammatically hardly documented, and the epistemological basis for Austroasian reconstructions is feeble. Internal reconstruction and informed comparison require detailed grammatical descriptions and a sound understanding of morphology and phonology. The inflexion of segments containing liquids and nasals is such a widespread phenomenon in Austroasiatic that it is easy to find apparent formal parallels elsewhere for these semantically still poorly defined grammatical categories. Finally, Reid draws heavily upon Nicobarese for his morphological parallels and not on the grammatically more conservative Munda languages. Comparing Malay prefixes with a hypothetical Austric source language for early loans in the Rgveda is therefore a fanciful exercise. But even if we dispense with ‘Austric’ and just deal exclusively with Austroasian, or more particularly with Munda, as Michael Witzel has judiciously chosen to do, we are still a long way from an historical grammar of Austroasiatic, and this is a severe limitation on such work.

My four remaining remarks deal not so much with linguistics, but more with archaeology and population history. First, in relation to the Northern or Kashmir Neolithic, Witzel says that the influence of the Indus civilization ‘is strong and long-lasting’ (p. 5). This is not the view held by archaeologists. In fact, one of the remarkable features of the Kashmir Neolithic is that this culture, to put it concisely, ‘is distinct and stands aloof from that of the rest of India’ (Ramachandran 1989: 52). In particular, archaeologists recognize that the Kashmir Neolithic represents a separate and independent tradition from that of the Indus civilization despite its geographical proximity to the latter. Certainly, there was trade, and the stray find of imported Kot Dijian pottery appearing quite out of context at one neolithic site in Kashmir has not diminished the view that the Kashmir Neolithic as an archaeological assemblage is closely affiliated with the Māljājāy culture in Gānsā and with sites such as mKhar-ro south of Cham­bodo in eastern Tibet, and not with the Indus tradition. I have discussed the archaeological context and the likely antecedents of the Northern or Kashmir Neolithic elsewhere (1997a, 1998).

A second point is word of caution regarding cultivated plants, especially in connexion with millet. Any grass with round edible seeds is called millet in English, and cultivated millets belong to a variety of distinct genera and have very different geographical origins. As an example, let us take the cultivar that is called ‘millet’ in Nepal, where it goes by the Nepali name of kodo. Whereas Sesuaria and Panicum millets were first cultivated in the Yellow River basin, the ‘ragi’ or finger millet Eleusine coracana cultivated by Tibeto-Burman peoples in Nepal ultimately originates from Africa, where the wild tetraploid form, which crosses freely with the cultivated variety, is still to be found. Although the latter is called kodo in Nepali, this is an altogether different plant from what in Hindi is known as kodo or kódá and therefore in English as ‘kodo millet’, i.e. Paspalum scrobiculatum, viz. ‘ditch millet’ or ‘birds’ millet’. The
stories on individual millets are often complex, and names for such cultivars should be treated circumspectly.

A third issue of immediate relevance to the population history of the Subcontinent is, or was, the Bangâni enigma. On this point Witzel gingerly says that the question ‘has not been entirely resolved’ (p. 47), and he mentions a website with postings about Bangâni, an Indo-Aryan language ostensibly bearing an early substrate layer from a kentum language. In good conscience, I can say that the question of a kentum substrate in the western Himalayas has been entirely resolved. The only kentum language to leave indelible traces in India today is English. I looked over the website mentioned once and, although this was some time ago, I saw nothing there that I should choose to dignify with a response or comment. What Suhni Râm Sharma and I have said on the topic is readily available in institutional libraries, and it will suffice here to refer to our published reports (van Driem and Sharma 1996, 1997).

In the theory of a Dravidian Indus, the Brahui are conventionally treated as a remnant of the original Dravidian population of the northwest. Jules Bloch once expressed skepticism about the northern provenance of the Brahui, but Georg Morgenstierne cogently argued against a southern provenance for the Brahui (1932: 5-7). Bloch’s old hypothesis of a northbound migration by the forebears of the Brahui along this coastal route was revived and defended by Josef Elfenbein. Elfenbein’s argument involved native traditions of the Kurukh and Malto. In advocating a southern provenance for the Brahui, Elfenbein was hard pressed to dismiss the evidence for an early presence of Brahui speakers in Kalat and Baluchistan. Moreover, Elfenbein acknowledged that attempts to identify earlier Iranian loans in Brahui are ‘greatly bedevilled by the nature of Balochi, extremely archaic and conservative in its phonology as it is’, so that by consequence ‘borrowings from Middle Iranian into Brahui are bound very often to be indistinguishable from borrowings from Balochi’. Elfenbein attempted to explain away the older Iranian loans in Brahui listed by Georg Morgenstierne as being ‘representatives of dialect forms of Balochi’. None the less, even Elfenbein accepted an etymology, first proposed by Denys Bray (1934, III: 74), whereby *bîrîna* ‘womb’ could represent a genuine Middle Persian survival in Brahui, for whereas Brahui preserves the older meaning, ‘the Modern Persian descendant, *bîr nurturing, means “hole, crevice”’ (1987: 219). I deal with this question in greater detail in my handbook Languages of the Himalayas, but this question really is far from resolved. At any rate, I am presently disinclined to believe Elfenbein’s theory about a southern provenance for the linguistic ancestors of the Brahui. 

References


Uvi, Sylvain. 1923. 'Pre-aryen et pre-dravidien dans l’Inde', Journal Asiatique, CCXIII: 1-57.


