

Linguistic Aspects of the Indo-European Urheimat Question

by Dr. Koenraad ELST

1. Introduction

1.1. Evidence sweeping all before it

When evidence from archaeology and Sanskrit text studies seems to contradict the theory of the entry of the Indo-Aryan branch of the Indo-European (IE) language family in India through the so-called "Aryan Invasion" (Aryan Invasion Theory, AIT), we are usually reassured that "there is of course the linguistic evidence" for this invasion, or at least for the non-Indian origin of the IE family.

Thus, F.E. Pargiter had shown how the Puranas locate Aryan origins in the Ganga basin and found "the earliest connexion of the Vedas to be with the eastern region and not with the Panjab"¹, but then he allowed the unnamed linguistic evidence to overrule his own findings: "We know from the evidence of language that the Aryans entered India very early".² (His solution is to relocate the point of entry of the Aryans from the western Khyber pass to the eastern Himalaya: Kathmandu or thereabouts.)

At the same time, the linguists themselves are often quite aware that the AIT is just a successful theory, not a proven fact. Those who try to take the scientific pretences of their discipline seriously, are not all that over-confident about the AIT. Many are willing to be modest and concede that so far it has merely been the most successful hypothesis. In fact, when quizzing linguists about the AIT, I came away with the impression that they too are not very sure of their case. By now, most of them have been trained entirely within the AIT framework, which was taken for granted and consequently not sought to be proven anymore. One of them told me that he had never bothered about a linguistic justification for the AIT framework, because there was, after all, "the well-known archaeological evidence"!

But for the rest, "the linguistic evidence" is still the magic mantra to silence all doubts about the AIT. At any rate, it is time that we take a look for ourselves at this fabled linguistic evidence.

1.2. Down with the linguistic evidence

A common reaction among Indians against this state of affairs is to dismiss linguistics altogether, calling it a "pseudo-science". Thus, Prof. N.S. Rajaram describes 19th-century comparative and historical linguistics, which generated the Aryan Invasion Theory (AIT), as "a scholarly discipline that had none of the checks and balances of a real science"³, in which "a conjecture is turned into a hypothesis to be later treated as a fact in support of a new theory".⁴

Likewise, N.R. Waradpande questions the very existence of an Indo-European language family and rejects the genetic kinship model, arguing very briefly that

similarities between Greek and Sanskrit must be due to very early borrowing.⁵ He argues that "the linguists have not been able to establish that the similarities in the Aryan or Indo-European languages are genetic, i.e. due to their having a common ancestry". He alleges that "the view that the South-Indian languages have an origin different from that of the North-Indian languages is based on irresponsible, ignorant and motivated utterances of a missionary".⁶ The "missionary" meant is the 19th-century pioneer of Dravidology, Bishop Robert Caldwell.

This rejection of linguistics by critics of the AIT creates the impression that their own pet theory, which makes the Aryans into natives of India rather than invaders, is not resistant to the test of linguistics. However, the fact that people fail to challenge the linguistic evidence, preferring simply to excommunicate it from the debate, does not by itself validate this body of evidence. Prof. Rajaram's remark that hypotheses are treated by scholars as facts, as arguments capable of overruling other hypotheses, is definitely valid for much of the humanities, including linguistics. To be sure, it doesn't follow that linguistics is a pseudo-science, merely that linguists in their reasoning have often fallen short of the scientific standard.

2. Origin of the linguistic argument

2.1. Linguistic and geographical distance from the origins

In the 18th century, when comparative IE linguistics started, the majority opinion was that the original homeland (or *Urheimat*) of the IE language family had to be India. This had an ideological reason, viz. that Enlightenment philosophers such as Voltaire were eager to replace Biblical tradition with a more distant Oriental source of inspiration for European culture.⁷ China was a popular candidate, but India had the advantage of being linguistically and even racially more akin to Europe; making it the homeland of the European languages or even of the European peoples, would be helpful in the dethronement of Biblical authority, but by no means far-fetched.

Moreover, there was also a seemingly good linguistic reason for choosing India as the *Urheimat*: the ancient Indian language, Sanskrit, was apparently the closest to the hypothetical Proto-Indo-European (PIE) language from which all existing members of the language family descended. It had all the grammatical categories of Latin and Greek in the most complete form, plus a few more, e.g. three numbers including a dualis in declension and conjugation, and all eight declension cases. Apparently, Sanskrit was very close to if not identical with PIE, and this was taken to support the case for India as the *Urheimat*.

In reality, there is no necessary relation between the linguistic antiquity of a language and its proximity to the *Urheimat*. Thus, among the North-Germanic languages, the one closest to Proto-North-Germanic is Icelandic, yet Iceland was most definitely not its *Urheimat*. The relative antiquity of Sanskrit vis-à-vis PIE does not determine its proximity to the *Urheimat*. Conversely, the subsequent dethronement of Sanskrit, and the progressive desanskritization of reconstructed PIE does not imply a geographical remoteness of India from the *Urheimat*. Yet, this mistaken inference has been quite common, though more often silent and implicit than explicit.

2.2. Kentum/satem

The first major element creating a distance between PIE and Sanskrit was the *kentum/satem* divide. It was assumed, in my view correctly (but denied by Indian scholars like Satya Swarup Mishra)⁸, that palatalization is a one-way process transforming velars (k,g) into palatals (c,j) but never the reverse; so that the velar or "kentum" (Latin for "hundred", from PIE **kmtom*) forms had to be the original and the palatal or "satem" (Avestan for "hundred") forms the evolved variants.

However, it would be erroneous to infer from this that the kentum area, i.e. Western and Southern Europe, was the homeland. On the contrary, it is altogether more likely that the Urheimat was in satem territory. The alternative from the angle of an Indian Urheimat theory (IUT) would be that India had originally had the kentum form, that the dialects which first emigrated (Hittite, Italo-Celtic, Germanic, Tokharic) retained the kentum form and took it to the geographical borderlands of the IE expanse (Europe, Anatolia, China), while the dialects which emigrated later (Baltic, Thracian, Phrygian) were at a halfway stage and the last-emigrated dialects (Slavic, Armenian, Iranian) plus the staybehind Indo-Aryan languages had adopted the satem form. This would satisfy the claim of the so-called Lateral Theory that the most conservative forms are to be found at the outskirts rather than in the metropolis.

Moreover, Indian scholars have pointed out that the discovery of a small and extinct kentum language inside India (Proto-Bangani, with *koto* as its word for "hundred"), surviving as a sizable substratum in the Himalayan language Bangani, tends to support the hypothesis that the older kentum form was originally present in India as well.⁹ This discovery had been made by the German linguist Claus Peter Zoller, who does not explain it through an Indian Urheimat Theory but as a left-over of a pre-Vedic Indo-European immigration into India.¹⁰ He claims that the local people have a tradition of their immigration from Afghanistan.

However, in a recent survey among Bangani speakers, George van Driem (Netherlands) and Suhnu R. Sharma have found the hypothesis of a kentum Proto-Bangani to be erroneous: the supposed kentum words turned out to be misreadings of quite ordinary modern Bangani words or phrases.¹¹ Then again, an even more recent survey on the spot by Anvita Abbi (Jawaharlal Nehru University) and her students has almost entirely confirmed Zoller's list of kentum substratum words in Bangani.¹² As the trite phrase goes: this calls for more research.

2.3. Sanskrit and PIE vowels

The second element in the progressive separation of Sanskrit from PIE was the impression that the [a/e/o] differentiation in Latin and Greek was original, and that their reduction to [a] in Sanskrit was a subsequent development (as in Greek *genos* corresponding to Sanskrit *jana*). Satya Swarup Mishra argues that it may just as well have been the other way around, and unlike the palatalization process, this vowel shift is indeed possible in either direction.¹³ Mishra cites examples from the Gypsy language,

but we need look no farther than English, where [a] has practically become [e] in "back" and "bake", and [o] in "ball".

There are, however, excellent reasons to stick to the conventional view that the [a/e/o] distinctness is original and their coalescence into [a] a later development. Firstly, the reduction to [a] is typical of just one branch, viz. Indo-Iranian, whereas a differentiation starting from [a] would have been a change uniformly affecting all the branches except one, which is less probable. Secondly, the different treatment of the velar consonants in reduplicated Sanskrit verb forms like *jagâma* or *cakâra* suggests a difference in subsequent vowel, with only the first vowel having a palatalizing impact on the preceding velar: *jegâma* < *gegâma*, *cekâra* < *kekâra*.

So, there is no reason to reject the conventional view that Greek vowels are closer to the PIE original than the Sanskrit vowels are. But here again, we also see no reason to make geographical deductions from this. India may as well have been the homeland of Proto-Greek, which left before the shift from [a/e/o] to [a] took place.

2.4. Indo-Hittite

A third element which increased the distance between reconstructed PIE and Sanskrit dramatically was the discovery of Hittite. Though Hittite displayed a very large intake of lexical and other elements from non-IE languages, some of its features were deemed to be older than their Sanskrit counterparts, e.g. the Hittite *genus commune* as opposed to Sanskrit's contrast between masculine and feminine genders, or the

It is by no means universally accepted that these features of Hittite are indeed PIE. Thus, the erosion of grammatical gender is a common phenomenon in IE languages, especially those suddenly exposed to an overdose of foreign influence, notably Persian when it was overwhelmed by Arabic, and English when it was overwhelmed by French influence (and this in spite of the fact that both French and Arabic have grammatical gender themselves). So, it is arguable that Hittite underwent the same development when it had to absorb large doses of Hattic or other pre-IE influence. The laryngeals have been explained by competent scholars as being due to South-Caucasian or Semitic influence.

But for our purposes there is no need to align ourselves with these dissident opinions. Even if we go with the dominant opinion and accept these elements as PIE, that is still no reason why the Urheimat should be in the historical location of Hittite or at least outside India. As the first emigrant dialect, Hittite could have taken from India some linguistic features (*genus commune*, laryngeals) which were about to disappear in the dialects emigrating only later or staying behind.

3. Direct geographical clues

3.1. Geographical asymmetry in expansion

In the 19th century, as India went out of favour, a number of European countries started competing for the honour of being the Urheimat. Ukraine and Russia gained the

upper hand with the archaeological discovery of the so-called Kurgan culture, dated to the 5th to 3rd millennium, and apparently the source of migrations into central and western Europe. This area also fell neatly in the middle of the expansion area of IE, a fact which some took as an element in support of the Kurgan culture's Urheimat claim. However, unless IE differs in this respect from other languages and language families, this central location argues more against than in favour of the Kurgan culture's Urheimat claim. Indeed, we find very few examples of languages expanding symmetrically: Chinese spread from the Yellow River basin southward, Russian from Ukraine eastward, Arabic from Arabia northwestward. There is consequently nothing against an IE migration starting from India and continuing almost exclusively in a westward direction.

The reason for this observed tendency to asymmetry is that the two opposite directions from a given region are only symmetrical in a geometrical sense: climatologically, economically and demographically, the two are usually very different,

e.g. the region north of the Yellow River is much less fertile and hospitable than the regions to its south. From the viewpoint of Kurgan culture emigrants, there was hardly a symmetry between the European West and the Indian Southeast: India was densely inhabited, technologically advanced and politically organized, Europe much less so. Europe could be overrun and culturally revolutionized by immigrants, while in India even large groups of immigrants were bound to be assimilated by the established civilization.

India satisfied the conditions for making the spectacular expansion of IE possible: like Europe in the colonial period, it had a demographic surplus and a technological edge over its neighbours. Food crises and political conflicts must have led to emigrations which were small by Indian standards but sizable for the less populated countries to India's northwest. Since these emigrants, increasingly mingled with the populations they encountered along the way, retained their technological edge vis-a-vis every next population to its west (esp. in the use of horse and chariot), the expansion in western direction continued until the Atlantic Ocean stopped it. Processes of elite dominance led to the linguistic assimilation of ever more westerly populations.

It is easy to see how and why the tendency to asymmetric expansion in the case of other languages also applies to India as the Urheimat of IE. On the road to the northwest, every next region was useful for the Indo-Europeans in terms of their established lifestyle and ways of food production. The mountainous regions to the north and west of India were much less interesting, as were the mountainous areas in the Indian interior. In India, Aryan expansion was long confined to the riverine plains with economic conditions similar to those in the middle basin of the Indus, Saraswati and Ganga rivers; the Vindhya and Himalaya mountains formed a natural frontier (the Vindhya mountains were first bypassed by sea, with landings on the Malabar coast). To the northwest, by contrast, after crossing the mountains of Afghanistan, emigrants could move from one riverine plain into the next: Oxus and Jaxartes, Wolga, Dniepr, Dniestr, Don, Danube, and into the European plain stretching from Poland to Holland. Only in the southwest of Europe, a more complex geography and a denser and more advanced native population slowed IE expansion down, and a number of pre-IE languages survived there into the Roman period, Basque even till today.

3.2. Geographical distribution

Another aspect of geographical distribution is the allocation of larger and smaller stretches of territory to the different branches of the IE family. We find the Iranian (covering the whole of Central Asia before 1000 AD) and Indo-Aryan branches each covering a territory as large as all the European branches (at least in the pre-colonial era) combined. We also find the Indo-Aryan branch by itself having, from antiquity till today, more speakers on the Eurasian continent (now nearing 900 million) than all other branches combined. This state of affairs could help us to see the Indo-Aryan branch as the centre and the other branches as wayward satellites; but so far, philologists have made exactly the opposite inference. It is said that this is the typical contrast between a homeland and its colony: a fragmented homeland where languages have small territories, and a large but linguistically more homogeneous colony (cfr. English, which shares its little home island with some Celtic languages, but has much larger stretches of land in North America and Australia all to itself, and with less dialect variation than in Britain; or cfr. Spanish, likewise).

It is also argued that Indo-Aryan must be a late-comer to India, for otherwise it would have been divided by now in several subfamilies as distinct from each other as, say, Celtic from Slavic. To this, we must remark first of all that the linguistic unity of Indo-Aryan should not be exaggerated. Native speakers of Indo-Aryan languages tell me that the difference between Bengali and Sindhi is bigger than that between, say, any two of the Romance languages. Further, to the extent that Indo-Aryan has preserved its unity, this may be attributed to the following factors, which have played to a larger extent and for longer periods in India than in Europe: a geographical unity from Sindh to Bengal (a continuous riverine plain) facilitating interaction between the regions, unlike the much more fragmented geography of Europe; long-time inclusion in common political units (e.g. Maurya, Gupta and Moghul empires); and continuous inclusion in a common cultural space with the common stabilizing influence of Sanskrit.

From the viewpoint of an Indian *Urheimat* hypothesis, the most important factor explaining the high fragmentation of IE in Europe as compared to its relative homogeneity in North India is the way in which an emigration from India to Europe must be imagined. Tribes left India and mixed with the non-IE-speaking tribes of their respective corners of Central Asia and Europe. This happens to be the fastest way of making two dialects of a single language grow apart and develop distinctive new characteristics: make them mingle with different foreign languages.

Thus, in the Romance family, we find little difference between Catalan, Occitan and Italian, three languages which have organically grown without much outside influence except for a short period of Germanic influence which was common to them; by contrast, Spanish and Rumanian have grown far apart (lexically, phonetically and grammatically), and this is largely due to the fact that the former has been influenced by Germanic and Arabic, while the latter was influenced by Greek and Slavic. Similarly, under the impact of languages they encountered (now mostly extinct and beyond the reach of our searchlight), and whose speakers they took over, the dialects of the IE emigrants from India differentiated much faster from each other than the dialects of Indo-Aryan.

3.3. Linguistic paleontology's failure

One of the main reasons for 19th-century philologists to exclude India as a candidate for Urheimat status was the findings of a fledgling new method called *linguistic paleontology*. The idea was that from the reconstructed vocabulary, one could deduce which flora, fauna and artefacts were familiar to the speakers of the proto-language, hence also their geographical area of habitation. The presence in the common vocabulary of words denoting northern animals like the bear, wolf, elk, otter and beaver seemed to indicate a northern Urheimat; likewise, the absence of terms for the lion or elephant seemed to exclude tropical countries like India.

It should be realized that virtually all IE-speaking areas are familiar with the cold climate and its concomitant flora and fauna. Even in hot countries, the mountainous areas provide islands of cold climate, e.g. the foothills of the Himalaya have pine trees rather than palm trees, apples (though these were imported) rather than mangoes. Indians are therefore quite familiar with a range of flora and fauna usually associated with the north, including bears (Sanskrit *rksha*, cfr. Greek *arktos*), otters (*udra*, Hindi *ûd/ûdbilâw*) and wolves (*vrka*). Elks and beavers do not live in India, yet the words exist, albeit with a different but related meaning: *rsha* means a male antelope, *babhru* a mongoose. The shift of meaning may have taken place in either direction: it is perfectly possible that emigrants from India transferred their term for "mongoose" to the first beavers which they encountered in Russia or other mongoose-free territory.

While the commonly-assumed northern location of PIE is at least disputable even on linguistic-paleontological grounds, as just shown, the derivation of its western location on the basis of the famous "beech" argument is undisputably flawed. The tree name *beech/fagus/bhegos* exists only in the Italic, Celtic and Germanic languages with that meaning, while in Greek (spoken in a beechless country) its meaning has shifted to "a type of oak". More easterly languages do not have this word, and their speakers are not naturally familiar with this tree, which only exists in western and central Europe. Somehow, our 19th-century predecessors deduced from this that PIE was spoken in the beech-growing part of Europe.

But in that case, one might have expected that at least some of the easterly languages had taken the word with them on their eastward exodus, applying it to other but somewhat similar trees (as Greek effectively did on its journey from central to southern Europe, a journey which it made in both the European and the Indian Urheimat scenarios). The distribution of the "beech" term is much better explained by assuming that it was an Old-European term adopted by the IE newcomers, and never known to those IE-speakers who stayed to the east of Central Europe. Few people now take the once-decisive "beech" argument seriously anymore.

3.4. Positive evidence from linguistic paleontology

It is one thing to show that the fauna terms provide no proof for a northern Urheimat. I believe that this can be done, so that the positive evidence from linguistic paleontology for a northern Urheimat is effectively refuted. Thomas Gamkrelidze and Vyacheslav Ivanov, in their bid to prove their Anatolian Urheimat theory, have gone a step further and tried to find terms for hot-climate fauna in the common IE vocabulary.¹⁴

Thus, they relate Sanskrit *pr̥daku* with Greek *pardos* and Hittite *parsana*, all meaning "leopard", an IE term lost in some northern regions devoid of leopards. The word "lion" is found as a native word, in regular phonetic correspondence, in Greek, Italic, Germanic and Hittite, and with a vaguer meaning "beast", in Slavic and Tokharic. Moreover, it is not unreasonable to give it deeper roots in IE by linking it with a verb, Sanskrit *rav-*, "howl, roar", considering that alternation *r/l* is common in Sanskrit (e.g. the double form *plavaga/pravaga*, "monkey", or the noun *plava*, "frog" related to the verb *pravate*, "jump").

A word for "monkey" is common to Greek (*kepos*) and Sanskrit (*kapi*), and Gamkrelidze and Ivanov argue for its connection with the Germanic and Celtic word "ape", which does not have the initial [k], for such *k/mute* alternation (which they derive from a pre-existing laryngeal) is also found in other IE words, e.g. Greek *kapros* next to Latin *aper*, Dutch *ever*, "boar". For "elephant", they even found two distinct IE words: Sanskrit *ibha*, "male elephant", corresponding to Latin *ebur*, "ivory, elephant"; and Greek *elephant-* corresponding to Gothic *ulbandus*, Tokharic **alpi*, "camel". In the second case, the "camel" meaning may be the original one, if we assume a migration through camel-rich Central Asia to Greece, where trade contacts with Egypt made the elephant known; the word may be a derivative from a word meaning "deer", e.g. Greek *elaphos*. In the case of *ibha/ebur*, however, we have a linguistic-paleontological argument for an Urheimat with elephants (Gamkrelidze and Ivanov also suggest a connection with Hebrew *shen-habbim*, "tusk-of-elephant", "ivory").

With this, we have briefly entered the game of linguistic paleontology, but not without retaining a measure of skepticism before the whole idea of reconstructing an environment of a proto-language from the vocabulary of its much younger daughter-languages. As Stefan Zimmer has written: "The long dispute about the reliability of this 'linguistic paleontology' is not yet finished, but approaching its inevitable end -- with a negative result, of course."¹⁵ This cornerstone of the European Urheimat theory is now largely discredited. At any rate, we believe we have shown that even if valid, the findings of linguistic paleontology would be neatly compatible with an Indian Urheimat.

4. Exchanges with other language families

4.1. Souvenirs of language contacts

One of the best keys to the geographical itinerary of a language is the exchange of lexical and other elements with other languages. Two types of language contact should be distinguished. The first type of language contact is the exchange of vocabulary and other linguistic traits, whether by long-distance trade contact, by contiguity or by substratum influence, between languages which are not necessarily otherwise related. A well-known example is the transmission of terms in the sphere of cattle-breeding from IE (mostly Tokharic) to Chinese: terms for dog, horse, cow, milk, honey. This doesn't add new information on the Urheimat question but neatly confirms the long-suspected presence of Tokharic in Western China since at least the 2nd millennium BC. It also tells us a lot about the relations between the tea-drinking Chinese farmers (till today, milk is a rarity in the Chinese diet) and the milk-drinking cattle-rearing "barbarians" on the northwestern borders.

A more surprising example is the apparent influence of Hamitic on Irish (as in the unusual word order in Irish sentences): it would seem that after the Ice Age, the European west coast was repopulated from the southwest, by Basque and even Hamitic-speaking peoples, who were assimilated into the IE and esp. the Celtic speech community, but smuggled some of their language traits into their newly adopted language. The example is interesting but does not provide information on the Urheimat, except to confirm that it was not in Celtic Western Europe.

Often, substratum elements are not identifiable with any known language. Thus, while IE has a neat decimal counting system, the Albanian and French languages show traces of a pre-IE, Old European counting system with base twenty, e.g. in French, 76 is *soixante-seize*, "60 + 16" (but in Belgian French, *septante-six*, "70 + 6", the normal IE form), or 80 is *quatre-vingts*, "4 x 20". The most likely explanation is that this was the prevalent system in parts of Europe in the pre-IE period, and that the people retained this system at least in part even after adopting an IE dialect as their language. This way, we find glimpses of pre-IE heritage in odd corners of the IE linguistic landscape.

4.2. Sumerian

A few terms exchanged with Sumerian, esp. *karpasa/kapazum*, "cotton", and possibly *ager/agar*, "field", and *go/gu*, "cow" (to cite some suggestions from Gamkrelidze and Ivanov's *magnum opus*), would confirm the presence of IE (though not necessarily of its PIE ancestor if Sumerian was the borrowing language) in an area conducting trade with Sumeria in the 3rd millennium or earlier. The main candidates would be Anatolia (Gamkrelidze and Ivanov's Urheimat choice) and the Indus basin.

But being the main language of civilization in ca. 3000 BC, one could not exclude contact through long-distance trade with the Kurgan area. Note however that the trade links between Sumeria and the Harappan civilization ("Meluhha" in Mesopotamian texts) are well-attested, e.g. the names *Arisena* and *Somasena* in a tablet from Akkad dating to ca.2200 BC.¹⁶ No such attestation exists for similar contacts with the Kurgan people.

4.3. Uralic

A case of contact on a rather large scale which is taken as providing crucial information on the Urheimat question is between early IE and Uralic. It was a one-way traffic, imparting some Tokharic, dozens of Iranian and also a few seemingly Indo-Aryan terms to either Proto-Uralic or Proto-Finno-Ugric (i.e. mainstream Uralic after Samoyedic split off). Among the loans from Indo-Iranian or Indo-Aryan, we note *sapta*, "seven, week", *asura*, "lord", *sasar*, "sister", *shata*, "hundred".¹⁷ At first sight, this would seem to confirm the European Urheimat theory: on their way from Europe, the Indo-Iranian and Tokharic tribes encountered the Uralic people in the Ural region and imparted some vocabulary to them. This would even remain possible if, as leading scholars of Uralic suggest, the Uralic languages themselves came from farther east, from the Irtysh river and Balkhash lake area.

The question of the Uralic homeland obviously has consequences. Karoly Rédei reports on the work of a fellow Hungarian scholar, Peter Hajdu (1950s and 60s): "According to Hajdu, the Uralic Urheimat may have been in western Siberia. The defect of this theory is that it gives no explanation for the chronological and geographical conditions of its contacts between Uralians (Finno-Ugrians) and Indo-Europeans (Proto-Aryans)."¹⁸ Not at all: Hajdu's theory explains nicely how these contacts may have taken place in Central Asia rather than in eastern Europe, and with Indo-Iranian rather than with the Western branches of IE. After the westward trek of the first IE-speaking tribes, it was the turn of the Iranians and the Uralic speakers to undertake parallel migrations to South Russia and North (European) Russia, respectively.

V.V. Napolskikh has supported the Siberian Urheimat theory of Uralic with different types of evidence from that given by Hajdu.¹⁹ The case against this Siberian Urheimat often rests precisely on a European Urheimat theory of IE, as Rédei's objection to Hajdu's position illustrates. So, if we drop the European Urheimat assumption for IE, we need not maintain it for Uralic either.

In that case, two alternative explanations are equally sustainable. Imagine the first waves of emigrants from India, taking most of the ancestor-dialects of the various branches of the IE family with them, through the Oxus valley to the Wolga plain and beyond. With the exception of Tokharic which remained in the area, they did not come in contact with Uralic, or when they did, they linguistically swallowed this marginal Uralic-speaking population without allowing it much substratal influence. Only the Slavic branch of IE shows some substratal influence from Uralic (and even this is disputed), a fact which is neatly compatible with an India-to-Europe migration: an Uralic-speaking tribe in the peri-Caspian region got assimilated in the westwardly expanding IE-speaking population.

It was the Iranians who came in contact with Uralic on a large scale, partly because they filled up the whole of Central Asia and (in the Scythian expansion) even Eastern Europe as far as Western Ukraine and Belarus, where an older Slavic population subsisted and adopted a lot of Iranian vocabulary, just as the Uralic population to its northeast did; and partly because the Uralic-speaking people were moving westward through the Urals region in a movement parallel to the Iranian westward expansion. At any rate, the Iranian influence is uncontroversial and easily compatible with any IE Urheimat scenario.

But how do the seemingly Indo-Aryan words fit in? One possibility is that these words were imparted to Uralic by non-Iranian, Indo-Aryan-speaking emigrants from India at the time of the great catastrophe in about 2000 BC, when the Saraswati river dried up and many of the Harappan cities were abandoned. This catastrophe triggered migrations in all directions: to the Malabar coast, to India's interior and east, to West Asia by sea (the Kassite dynasty in Babylon in ca. 1600 BC venerated some of the Vedic gods)²⁰, and to Central Asia. The Sanskrit terms in the Mitannic language attested in Kurdistan in the 15th century BC seem to be a leftover of an Indo-Aryan presence in West Asia, which presupposes an earlier Indo-Aryan migration through (an already predominantly Iranian-speaking) Central Asia. A similar emigrant group may have ended up in an Uralic-speaking environment, imparting some of its own terminology but getting assimilated over time, just like their Mitannic cousins. The Uralic term *orya*, "slave", from

either Iranian *airya* or Sanskrit *arya*, may indicate that their position was not as dignified as that of the Mitannic horse trainers.

An alternative possibility is that the linguistic exchange between Proto-Uralic and Iranian took place at a much earlier stage, before Iranian had grown distinct from Indo-Aryan. It is by no means a new suggestion that these seemingly Indo-Aryan words are in fact Indo-Iranian, i.e. dating back to before the separation of Iranian from Indo-Aryan, or in effect, before the development of typical Iranianisms such as the softening of [s] to [h]. This would mean that the vanguard of the Iranian emigration from India had not yet changed *asura* and *sapta* into *ahura* and *hafta*, and that Iranian developed its typical features (some of which it shares with Armenian and Greek, most notably the [s]>[h] shift) outside India. This tallies with the fact (admittedly only an argument *e silentio*) that the Vedic reports on struggles with Iranian tribes such as the Dasas and the Panis (attested in Greco-Roman sources as the East-Iranian tribes *Dahae* and *Parnoi*), the Pakthas (Pathans?), Parshus (Persians?), Prthus (Parthians?) and Bhalanas (Baluchis?) never mention any term or phrase or name with typically Iranian features.²¹

Even the stage before Indo-Iranian unity, viz. when Indo-Iranian had not yet replaced the PIE *kentum* forms with its own *satem* forms, may already have witnessed some lexical exchanges with Uralic: as Asko Parpola has pointed out, among the IE loans in Uralic, we find a few terms in *kentum* form which are exclusively attested in the Indo-Iranian branch of IE, e.g. Finnish *kehrä*, "spindle", from PIE **kettra*, attested in Sanskrit as *cattrā*.²² It is of course also possible that words like **kettra* once did exist in branches other than Indo-Iranian but disappeared in the intervening period along with so many other original PIE words which were replaced by non-IE loans or new IE formations. If *kettra* was indeed transmitted to Uralic by early Indo-Iranian, it may have been as a result of trade instead of migration, for the Indus basin was an advanced manufacturing centre which exported goods deep into Central Asia.

This leads us to a third possibility, viz. that the seemingly Indo-Aryan words in Uralic were transmitted by long-distance traders, regardless of migrations, possibly even at a fairly late date. They may have been pure Indo-Aryan, as distinct from Iranian, normally spoken only in India itself, but brought to the Uralic people by means of long-distance trade, regardless of which languages were spoken in the territory in between, somewhat like the entry of Arabic and Persian words in European languages during the Middle Ages (e.g. *tariff*, *cheque*, *bazar*, *douane*, *chess*). If we see India in the 3rd millennium BC as the mighty metropolis whose influence radiated deep into Central Asia (as archaeology suggests)²³, this cannot be ruled out. At any rate, we believe we have shown enough possible ways to reasonably reconcile the lexical exchange between the eastern IE languages and Uralic with an Indian Urheimat scenario.

4.4. "Nostratic"

Isoglosses with other languages may be due to historical contact between the languages, but also to deep kinship: just as Portuguese and Italian have both developed out of Latin (partly by absorbing each its own dose of foreign elements), and just as both Latin and Tokharic have evolved out of a common ancestor-language provisionally called PIE, so PIE must have evolved from an even earlier language, which may at the same time have been the ancestor of other language families as well.

The most important theory in this line of research is the *Nostratic* superfamily theory, postulating a common origin for Eskimo-Aleut, Altaic, Uralic, IE, Afro-Asiatic, Dravidian and possibly South-Caucasian. Some people make fun of this theory, and refer it jokingly to the "nostratosphere", yet its basic postulate makes perfect sense: differentiation of ancestor-languages, as attested in detail in the case of Latin and the Romance language family, must have happened at earlier stages of history as well. Whether the present superfamily theory and the methods actually used for reconstructing the supposed Nostratic vocabulary are at all acceptable, is a different matter.

The state of the art is that we just don't know very much yet about the ancestry of PIE, especially when even the location of PIE in its heyday is still the object of debate. But just to be on the safe side in case of a breakthrough of the Nostratic theory, we do

want to remark that the distribution of the alleged Nostratic language families at their earliest date of appearance, with most of them within travelling distance from the Indus-Saraswati basin (Uralic in the Ob-Irtysh basin, Altaic in Mongolia, Semitic in Mesopotamia, Elamite in Iran, Dravidian on the Indian coast), is certainly compatible with a Northwest-Indian Urheimat of IE, more than with a European Urheimat. For the rest, it is best to leave these proto-proto-languages alone and concentrate on real language families.

4.5. Semitic

Semitic (and by implication also the Chadic, Kushitic and Hamitic branches of the Afro-Asiatic family, assumed to be the result of a pre-4th-millennium immigration of early agriculturists from West Asia into North Africa) is suspected to spring from a common ancestor with IE, even by scholars skeptical of Nostratic adventures. The commonality of some elementary lexical items is striking, and includes the numerals 6 and 7 (Hebrew *shisha*, *shiva*, Arabic *sitta*, *sab'a*, conceivably borrowed at the time when counting was extended beyond the fingers of a single hand for the first time), arguably even all the first seven numerals.

Contact with Akkadian (the Semitic language of Mesopotamia in the third millennium BC) and even Proto-Semitic is attested by a good handful of words, esp. some terms for utensils and animals. This includes two terms for "axe": PIE **peleku*, Greek *pelekus*, Ossetic *faeraet*, Sanskrit *parashu*, "axe", related (one way or the other)

to Akkadian *pilaqqu*, "axe", cfr. Arabic *falaqa*, "to split apart"; and PIE **sekwr*, Latin *securis*, "axe", *secula*, "hatchet", Old Slavic *sekyra*, "hatchet", related to a Semitic root yielding Akkadian *shukurru*, "javelin", Hebrew *segor*, "axe". Some terms are in common only with the Western IE languages, e.g. Semitic *gedi*, still recognizable in English *goat*.

This testimony is too slender, though, for concluding that the Western Indo-Europeans had come from the East and encountered the Semites on their way to the West.

Even more remarkable are the common fundamental grammatical traits, which indicate a common genetic origin rather than an influence from the one language family on the other. Semitic, like IE, has grammatically functional vowel changes, grammatical gender, declension, conjugational categories including participles and medial and passive modes, and a range of phonemes which in Proto-Semitic was almost entirely in common with PIE, even more so if we assume PIE laryngeals to match Semitic *aleph*, *he* and *'ayn*. Many of these grammatical elements are shared only by Semitic (or Afro-Asiatic) and IE, setting them off as a pair against all other language families. If any language family has a chance of being the sister of the IE family, it is Semitic.

One way to imagine how Semitic and IE went their separate ways has been offered by Bernard Sergent, who is strongly convinced of the two families' common origin. He combines the linguistic evidence with archaeological and anthropological indications that the (supposedly PIE-speaking) Kurgan people in the North-Caspian area of ca. 4000 BC came from the southeast, a finding which might otherwise be cited in support of their Indian origin. Thus, the Kurgan people's typical grain was millet, not the rye and wheat cultivated by the Old Europeans, and in ca. 5000 BC, millet had been cultivated in what is now Turkmenistan (it apparently originates in China), particularly in the mesolithic culture of Jebel. From there on, the archaeological traces become really tenuous, but Sergent claims to discern a link with the Zarzian culture of Kurdistan 10,000 to 8500 BC. Short, he suggests that the Kurgan people had come along the eastern coast of the Caspian Sea, not from the southeast (India) but the southwest, in or near Mesopotamia, where PIE may have had a common homeland with Semitic.²⁴

However, those who interpret the archaeological data concerning the genesis of agriculture in the Indus site of Mehrgarh as being the effect of a diffusion from West Asia, may well interpret an eventual kinship of IE with Semitic as proving their own point: along with its material culture, Mehrgarh's language may have been an offshoot of a metropolitan model, viz. a Proto-Semitic-speaking culture in West Asia. This would mean that the Indus area was indeed the homeland of the original PIE, but that in the preceding millennia, PIE had been created by the interaction of Proto-Semitic-speaking colonists from West Asia with locals. On the other hand, now that the case for an independent genesis of the Neolithic revolution (i.e. the development of agriculture) in Mehrgarh is getting stronger, we may have to reconsider the direction of such a process.

4.6. Dravidian substratum elements

Apart from contact between different languages which have continued to exist, there can also be influence from a disappearing language on a surviving language, often in the form of a substratum: people take to speaking a new (mostly the elite's) language, and drop their old language all while preserving some lexical items, some phonetic

propensities, some grammatical ways of organizing information. The alleged presence of a large dose of "pre-Aryan" substratum features in Sanskrit and the other Indo-Aryan languages, notably from now-extinct Dravidian languages once spoken in northern India, was historically one of the important reasons for deciding against India as the *Urheimat*.

In the 19th century, it was not yet realized how the European branches of IE are all full of substratum elements, mostly from extinct Old European languages. For Latin, this includes such elementary terms as *altus* and *urbs*, borrowed from a substratum language tentatively described as "Urbian". For Germanic, it includes some 30% of the acknowledged "Germanic" vocabulary, including such core lexical items as *sheep* and *drink*. For Greek, it amounts to some 40% of the vocabulary, both from extinct branches of the Anatolian (Hittite-related) family and from non-IE languages. The branch least affected by foreign elements is Slavic, but this need not be taken as proof of a South-Russian homeland: in an Indian *Urheimat* scenario, the way for Slavic would have been cleared by forerunners, chiefly Celtic and Germanic, and though these languages would absorb many Old-European elements as substratum features, they also eliminated the Old-European languages as such and prevented them from further influencing Slavic.

Even if we accept as non-IE all the elements in Sanskrit described as such by various scholars, the non-IE contribution is still not greater than in some of the European branches of IE.²⁵ And, as Shrikant Talageri has shown, a large part of this so-called Dravidian contribution is highly questionable: many words routinely described as

Dravidian-originated have been analyzed as pure IE.²⁶ Numerous supposed loanwords have no counterpart in Dravidian and Munda, or when they do, there is often no reason to assume that the direction of borrowing was into rather than out of Indo-Aryan, especially when you consider that Dravidian is attested in writing at least 1500 years after (and at a distance of 2000 km) the Sanskrit sources, and Munda has not been committed to writing until the 19th century.

The observation had been made earlier by Western scholars: the convergence of Indo-Aryan and Dravidian (as well as Munda and to an extent Burushaski) in lexical and grammatical features need not be due to a Dravidian substratum, for which there are in fact no compelling indications.²⁷ At any rate, there has been so much interaction of Indo-Aryan with Dravidian, including exchange of people and goods, that a Dravidian contribution (as a neighbourly or *adstratum* influence) is perfectly normal; this contribution remains in any case much smaller than the well-known Indo-Aryan influence on the Dravidian languages, which no one tries to explain as a substratum effect.

In this respect, the testimony of the place-names may be useful. In the Hindi belt and most of Panjab, there is no evidence of a Dravidian substratum in the toponyms. By contrast, in Sindh and Gujarat, Dravidian toponyms are fairly common, e.g. names ending in *valli/palli*, "village". In Sindhi, and more so in Gujarati and Marathi, Dravidian influence is discernible, e.g. in the existence of two pronouns for *we*, an inclusive one (including the speaker as well as the person addressed) and an exclusive one (including only the speaker and his group, like in the French expression *nous autres*). By contrast, Hindi has much fewer Dravidian elements, even "losing" (or just never having had) a number of loanwords which had been adopted in Sanskrit. There is no reason to assume a Dravidian presence in North India, but it seems to have been there in the coastal area.

This would fit in with David McAlpin's Elamo-Dravidian theory, which puts Proto-Elamo-Dravidian on the coast of Iran, spreading westwards to Mesopotamia (Elam) and eastwards to Sindh and along the Indian coast southward.²⁸ This theory is supported by the similarities between the undeciphered early Elamite script and the Harappan script, and by the survival of the Brahui Dravidian speech pocket in Baluchistan. It would make the Harappan culture area bi- and possibly multi-lingual: a perfectly normal situation, comparable with multi-lingual Mesopotamia or with Latin-Greek bilingualism in the Roman Empire.

But in that case, Indo-Aryan influence on Dravidian may be much older than usually assumed, and date back well into the heyday of Harappan culture. However, the Dravidians influenced by Indo-Aryan in Gujarat and Maharashtra may have been a dead-end in the history of Dravidian, losing their language altogether. There is no trace of Harappans migrating south, whether to save their Dravidian language from Indo-Aryan contamination or for other, more likely reasons.

Either way, Indo-Aryan influence on Dravidian is certainly more profound than generally thought. Apart from the *tatsama* (literally adopted) Sanskrit words which make up more than half of Telugu or Kannada vocabulary, and which are attributed to the influence of Brahmin families settling in South India since the turn of the Christian era, many apparent members of the Dravidian core vocabulary as attested in Sangam Tamil are actually very ancient *tadbhava* (evolved and sometimes unrecognizably changed) loans from Sanskrit or Prakrit, e.g. *âkâyam*, "sky" (< *âkâsha*); *âyutham*, "weapon" (< *âyudha*); *tavam*, "penance" (< *tapas*); *tîvu*, "island" (< *dwîpa*); *chetti*, "foreman, merchant" (< *shreshthi*), *tîru*, term of respectful address (< *shrî*).²⁹ It is not impossible that there ever was a pure Dravidian language in South India, but in the oldest texts already, we find a Dravidian written in a Brahmi-derived script and influenced by Sanskrit.

Many scholars now assume that there was a third language in northwestern India, which acted as a buffer between Dravidian and Indo-Aryan before being eliminated by the latter. Words looking like Dravidian loans in Indo-Aryan could then in fact have been borrowed from this third language into both Indo-Aryan and Dravidian. To Indian critics of linguistics as a "pseudoscience", such a ghost language is a perfect proof of the purely speculative nature of our science. Yet, it is an entirely reasonable proposition: even Sumerian, one of the great vehicles of civilization, died out, and we have reason to assume that the Bhil tribals originally spoke a different language, possibly related to the isolated tribal Nahali language still spoken in a few villages in Madhya Pradesh.

Such a buffer language would at any rate explain, in an Indian Urheimat theory, why there is no Dravidian influence on IE as a whole, merely on Indo-Aryan and to a very small extent on Iranian (though it is remarkable that some of the words transmitted from Indo-Iranian to Uralic are usually credited with a Dravidian origin, e.g. *shishu*, "child", and *kota*, "house"; if correct, this would be a modest argument for an Indian Urheimat). By the time the buffer language had been swallowed and Dravidian-IE interaction began, most of the IE proto-languages had already left India.

4.7. Sino-Tibetan

To prove an Asian homeland for IE, it is not good enough to diminish the connections between IE and more westerly language families. To anchor IE in Asia, the strongest argument would be genetic kinship with an East-Asian language family.

There have been very early contacts between IE and Chinese, fossilized in IE loan-words in Chinese, e.g. *ma* (< **mra*, cfr. *mare*, Sanskrit *marka*, "swift"), "horse"; *quan*, "hound"; *sun*, "grandson" (cfr. *son*); *mi*, "honey" (cfr. *mead*, Sanskrit *medhu*); *gu*, "bull", and *niu*, "cow" (through **ngiu*, from IE **gwou-*); and, more recently, *shi*, "lion" (Iranian *sher*). Chang Tsung-tung has pleaded that there were linguistic and cultural

contacts between Indo-Europeans from Inner Asia and late-neolithic Chinese peasants, who learned cattle-breeding from them.³⁰ These loans generally came through Tokharic, which we know was the northwestern neighbour of Chinese for many centuries, at least since the turn of the 1st millennium BC when the Tokhars are mentioned in records of the Western Zhou dynasty, and until the mid-1st millennium AD.

The contact between Tokharic and Chinese adds little to our knowledge of the *Urheimat* but merely confirms that the Tokharic people lived that far east. The adoption of almost the whole range of domesticated cattle-names from Tokharic into Chinese also emphasizes a fact insufficiently realized, viz. how innovative the cattle-breeding culture of the early IE tribes really was. They ranked as powerful and capable, and their prestige helped them to assimilate large populations culturally and linguistically. But for *Urheimat*-related trails, we must look elsewhere.

Vedic Sanskrit and ancient Greek, and therefore perhaps also PIE, had a pitch accent, a typical feature of Proto-Sino-Tibetan, preserved in Chinese and in a smaller way in Tibetan. True, the behaviour of this pitch accent is completely different in Vedic from what it is in Sino-Tibetan. But that is only what you would expect after millennia of separate development; after all, the behaviour of the pitch accent is completely different between some of the Sino-Tibetan languages as well. Picking up this hint from a similarity in accentuation, scholars have looked around for other "deep", structural similarities, e.g. the presumed fact that all PIE roots were monosyllabic.³¹ Edwin Pulleyblank claims to have reconstructed a number of rather abstract similarities in the phonetics and morphology of PIE and Sino-Tibetan.

Though he fails to back it up with any (even a single) lexical similarity, he confidently dismisses as a "prejudice" the phenomenon that "for a variety of reasons, the possibility of a genetic relationship between these two language families strikes most people as inherently most improbable." He believes that "there is no compelling reason from the point of view of either linguistics or archaeology to rule out the possibility of a genetic connection between Sino-Tibetan and Indo-European. Such a connection is certainly inconsistent with a European or Anatolian homeland for the Indo-Europeans but it is much less so with the Kurgan theory", esp. considering that the Kurgan culture "was not the result of local evolution in that region but had its source in an intrusion from an earlier culture farther east".³² This is of course very interesting, but: "It will be

necessary to demonstrate the existence of a considerable number of cognates linked by regular sound correspondences. To do so in a way that will convince the doubters on both sides of the equation will be a formidable task."³³

Apart from Pulleyblank's vision of a deep, Nostratic-type connection between Sino-Tibetan and PIE, we should also consider the question of influence, especially the interaction with neighbouring Tibetan. There is of course a mass of Buddhistic loan-words which crept into Tibetan during the Middle Ages, but they tell us nothing about origins.

As Prof. Ulrich Libbrecht writes, the Tibetans were not native to their present habitat, and immigrated there in the historical period: "The general ethnic movement of the Sinitic-speaking peoples was southward. The immigration of Tai- and Tibeto-Burman-speaking languages in Indochina has entirely taken place within the historical period. The same is true of the Chinese-speaking peoples from the middle part of the Yellow River basin towards the southern and eastern coast. Indications from Greek geographers and in Tibetan traditions teach us that the early centre of these peoples lay more to the north than present-day Tibet, viz. in the upper Yangzi basin. It is suspected that the centre of dispersion of the Sinitic languages was near the Koko-nor lake, at the borders of China proper, Tibet and Mongolia. From there, one branch spread eastward and formed the Chinese language; another went southward to form the Tibeto-Burman subgroup. The cause of this dispersal may well be found in the periodic droughts affecting Inner Asia in prehistoric and historical periods."³⁴

So, unless PIE came from China, there may have been thousands of years without any substantial contact between IE and Sino-Tibetan, the first contact being the Tokharian settlement on the Chinese border. No evidence of contact has been identified for the PIE period.

4.8. Austronesian

A language family with unexpected similarities to IE, similarities which provide a strong geographical clue, is Austronesian. This family of languages is the one with the second greatest geographical spread after IE: from Madagascar through Malaysia and Indonesia, Taiwan and the Philippines, to Melanesia and Polynesia, as far south as New Zealand. So, what is the relation of Austronesian to Indo-Aryan and to PIE?

According to Franklin Southworth: "The presence of other ethnic groups, speaking other languages [*than IE, Dravidian or Munda*], must be assumed (...) numerous examples can be found to suggest early contact with language groups now unrepresented in the subcontinent. A single example will be noted here. The word for 'mother' in several of the Dardic languages, as well as in Nepali, Assamese, Bengali, Oriya, Gujarati, and Marathi (...) is *âî* (or a similar form). The source of this is clearly the same as that of classical Tamil *ây*, 'mother'. These words are apparently connected with a widespread group of words found in Malayo-Polynesian (cf. Proto-Austronesian **bayi* ...) and elsewhere. The distribution of this word in Indo-Aryan suggests that it must

have entered Old Indo-Aryan very early (presumably as a nursery word, and thus not likely to appear in religious texts), before the movement of Indo-Aryan speakers out of the Panjab. In Dravidian, this word is well-represented in all branches (though *amma* is perhaps an older word) and thus, if it is a borrowing, it must be a very early one."³⁵

Next to *âyî*, "mother", Marathi has the form *bâî*, "lady", as in *Târâ-bâî*, *Lakshmî-bâî* etc.; the same two forms are attested in Austronesian. So, we have a nearly pan-Indian word, attested from Nepal and Kashmir to Maharashtra and Tamil Nadu, and seemingly related to Austronesian. For another example: "Malayo-Polynesian shares cognate forms of a few [words which are attested in both Indo-Aryan and Dravidian], notably Old Indo-Aryan *phala*-['fruit'], Dravidian *palam* ['ripe fruit'], etc. (cf. Proto-Austronesian **palam*, 'to ripen a fruit artificially'...), and the words for rice."³⁶

Austronesian seems to have very early and very profound links with IE. In the personal pronouns (e.g. Proto-Austronesian **aku*, cfr. *ego*), the first four numerals (e.g. Malay *dua* for "two") and other elementary vocabulary (e.g. the words for "water" and "land"), the similarity is too striking to be missed. Remarkable lexical similarities had been reported since at least the 1930s, and they have been presented by Isidore Dyen in 1966.³⁷ Dyen's comparisons are sometimes not too obvious but satisfy the linguistic requirement of regularity. At the same time, this lexical influence or exchange is not backed up by grammatical similarities: in contrast with the elaborate categories of IE grammar, Austronesian grammar looks very primitive, the textbook example being the Malay plural by reduplication, as in *orang*, "man", *orang-orang*, "men".³⁸

Most scholars of IE including myself know too little of Austronesian to verify this claim, and all of us tend to remind ourselves of the existence of pure coincidence when confronted with these data. At any rate, the relation would be one between the entire Austronesian and the entire Indo-European family, indicating that it pre-dates their split into daughter languages. Moreover, it concerns the very core of the vocabulary. Further, it so happens that some Austronesian languages have the typically Indian cerebral or retroflex consonants; it is possible that this was an original feature of Proto-Austronesian, which its other daughter languages have lost.

As for the language structure, to our knowledge the similarity between PIE and Proto-Austronesian is not established as being much above statistical coincidence. It is, in that case, less than that between PIE and Proto-Semitic, which latter is still not enough to convince all linguists of a genetic relationship rather than an influence through contact. At first sight, the similarities between IE and Austronesian vocabularies may therefore better be explained through contact than through a genetic relationship. In this case, we may also be dealing with a case of heavy pidginization: a mixed

population adopting lexical items from PIE but making up a grammar from scratch. Then again, genetically related languages may become completely different in language structure (e.g. English vs. Sanskrit, Chinese vs. Tibetan). Dyen therefore saw no objection to postulating a common genetic origin rather than an early large-scale borrowing.

Dyen cannot be accused of an Indian Urheimat bias either for IE or for Austronesian. For the latter, "Dyen's lexicostatistical classification of Austronesian suggested a Melanesian homeland, a conclusion at variance with all other sources of information (...) heavy borrowing and numerous shifts in and around New Guinea have obviously distorted the picture", according to Peter Bellwood.³⁹ It is *in spite of* his opinions about the Austronesian and IE homelands that he felt forced to face facts concerning IE-Austronesian similarities. Meanwhile, the dominant opinion as reported by Bellwood is that Southeast China and Taiwan are the Urheimat from where

Austronesian expanded in all seaborne directions (hence its substratum presence in Japanese, a rather hard nut to crack for an Indian Urheimat theory of Austronesian).

Yet, just as the Kurgan culture may be a secondary centre of IE dispersal, formed by immigrants from, say, India, the supposed Southeast-Chinese Urheimat of Austronesian may itself be a secondary homeland. If there is to be a point of contact between PIE and Proto-Austronesian, it is hard to imagine it in another location than India.

Bernard Sergent suggests northern China, arguing that the yellow race as a whole comes from there, and that the Chinese-Siberian border was the place of contact between white Indo-Europeans and the yellow race, including speakers of Sino-Tibetan, Austro-Asiatic (Munda, Khmer) and Austronesian.⁴⁰ But that is a *petitio principii*: just as it need not be assumed that the Proto-Indo-Europeans were blonde Nordics (as Sergent himself has forcefully argued)⁴¹, there is no ground for racial assumptions about the Austronesians. If they originated in India, they may have been brown-skinned (as most of them still are) rather than yellow. Moreover, even if it is assumed that Austronesian came from southern China, there is no need to trace it further back to northern China; and if its very thin connection to northern China is sufficient for an impressive amount of IE-Austronesian isoglosses, how come there aren't even more IE-Chinese isoglosses, as Chinese or Sino-Tibetan has a much longer certified presence in northern China on the border with the barbarians?

For another alternative: suppose the Indo-Europeans and the Austronesians shared a homeland somewhere in southern China or Southeast Asia. An entry of the Indo-Europeans into India from the east, arriving by boat from Southeast Asia, is an interesting thought experiment, if only to free ourselves from entrenched stereotypes. Why not counter the Western AIT with an Eastern AIT? Just imagine, a wayward Austronesian tribe sailed up the Ganga led by one Manu who, as related in the Puranas, started Aryan history in the mid-Ganga basin (Ayodhya, Prayag, Kashi), and whose progeny subsequently conquered the Indus basin and expanded further westward. In that case, the elaborate structure of PIE would be an innovation due to a peculiar intellectual culture (let's call it proto-brahminism) and to the influence of local languages, including perhaps a lost branch of Semitic spoken by colonists who had brought agriculture from West Asia to Indus settlements like Mehrgarh.

We will welcome any new evidence which forces us to take the southeastern scenario seriously. Until then, if there has to be a common homeland of IE and Austronesian, we consider India more likely. India, in this case, may have to be understood as including the submerged lands to its south which were inhabited perhaps as late as 5000 BC. The scenario that unfolds is of India as a major demographic growth centre, from which IE spread to the north and west and Austronesian to the southeast as far as Polynesia. Though disappearing from India, Austronesian expanded in the same period and just as spectacularly as IE. These two most impressive linguistic migrations would then have been part of one India-centred expansion movement spanning the Old World from Iceland to New Zealand.

5. Conclusion

We have just presented the pro and contra of some *prima facie* indications for language contacts which would imply an ancient IE and even PIE presence in Harappan and pre-Harappan India. In our opinion, none of these can presently be considered decisive evidence for an Indian Urheimat theory.

However, to put the strengths and weaknesses of our findings in the proper perspective, we should not forget to also evaluate the evidence from language contacts for the rivalling European Urheimat theory, which should be put to the same tests as the Indian Urheimat theory. The fact is that such evidence is very scarce, if not non-existent. The Old-European Basque language has no ancient links with IE. For the rest, all Old-European languages have disappeared and have not even survived as dead inscriptional languages providing us with material for linguistic comparison. Evidence of the type tentatively provided by isoglosses between IE and Semitic, Austronesian or Uralic, all Asian language families, is simply not available for the westerly branches of IE or for a hypothetical Europe-based PIE. On balance, the evidence from contact with once-neighbouring languages does not provide compelling evidence for an Indian Urheimat (unless the Austronesian connection is valid), but even less evidence for a European Urheimat.

It is too early to say that linguistics has proven an Indian origin for the IE family. But we can assert with confidence that the oft-invoked linguistic evidence for a European Urheimat and for an Aryan invasion of India is completely wanting. One after another, the classical proofs of the European Urheimat theory have been discredited, usually by scholars who had no knowledge of or interest in an alternative Indian Urheimat theory. In the absence of a final judgment by linguistics, other approaches deserve to be taken seriously, unhindered and uninhibited by fear of that large-looming but in fact elusive "linguistic evidence".

Footnotes:

1. F.E. Pargiter: *Ancient Indian Historical Tradition*, Motilal Banarsidass, Delhi 1962, p.302.

2. F.E. Pargiter: *Ancient Indian Historical Tradition*, p.1.

3 N.S. Rajaram: *The Politics of History*, Voice of India, Delhi 1995, p.144.

4. N.S. Rajaram: *The Politics of History*, p.217.

5. N.R. Waradpande: *The Aryan Invasion, a Myth*, Babasaheb Apte Smarak Samiti, Nagpur 1989, p.19-21.

6. N.R. Waradpande: "Fact and fiction about the Aryans", in S.B. Deo & SuryanathKamath: *The Aryan Problem*, Bharatiya Itihasa Sankalana Samiti, Pune 1993, p.14-15.

7. The classic reference for the ideological factors in the development of the Indo-European theory is Léon Poliakov: *The Aryan Myth*, London 1974.

8. Satya Swarup Mishra: *The Aryan Problem* (Delhi 1992). This palatalization is known in numerous languages, e.g. Chinese (*Yangzi-kiang* > *Yangzi-jiang*), the Bantu language Chiluba (cfr. *Ki-konko*, *Ki-swahili*, but *Chi-luba*), Arabic (*Gabriel* > *Jibrīl*), English (*kirk* > *church*), the Romance languages, Swedish etc.

9. E.g. Shrikant Talageri: *The Aryan Invasion Theory, a Reappraisal*, Aditya Prakashan, Delhi 1993, p.70.

10. The "discovery" of Kentum elements in Proto-Bangani was announced to the world by Claus Peter Zoller at the 7th World Sanskrit Conference, Leiden 1987, in his paper: "On the vestiges of an old Kentum language in Garhwal (Indian Himalayas)", and elaborated further in his articles: "Bericht über besondere Archaismen im Bangani, einer Western Pahari-Sprache", *Münchener Studien zur Sprachwissenschaft*, 1988, p.173-200, and: "Bericht über grammatische Archaismen im Bangani", *ibid.*, 1989, p.159-218.

11. George van Driem and Suhnu Ram Sharma: "In search of Kentum Indo-Europeans in the Himalayas", *Indogermanische Forschungen*, 1996, p.107-146. In terms of serenity and academic factuality, the language they use to qualify Zoller's work leaves much to be desired, a fact which is sure to be used by the Indocentric school to prove their point that the AIT school is just biased.

12. Anvita Abbi: "Debate on archaism of some select Bangani words", <http://www-personal.umich.edu/~pehook/bangani.abbi2.html>, 1998.

13. Satya Swarup Mishra: *The Aryan Problem*.

much-discussed laryngeal consonants, absent in Sanskrit as in all other IE languages.

14. T. Gamkrelidze and V. Ivanov: *Indo-European and the Indo-Europeans*, Walter De Gruyter, Berlin 1995.

15. S. Zimmer: "On Indo-Europeanization", *Journal of Indo-European Studies*, spring 1990. Cited in R.S. Sharma: *Looking for the Aryans*, p.36, with reference to J. Harmatta: "The emergence of the Indo-Iranians: the Indo-Iranian languages", in A.H. Dani and

16. V.M. Masson, ed.: *History of Civilizations*, vol.1, UNESCO Publ., Paris 1992, p.374

17. A rather complete list and discussion of common IE-Uralic vocabulary is Karoly Rédei: "Die ältesten indogermanischen Lehnwörter der Uralischen Sprachen", in Denis Sinor, ed.: *The Uralic Languages: Description, History and Foreign Influences*, Brill, Leiden 1988, p.638-664.

18. Karoly Rédei: "Die ältesten indogermanischen Lehnwörter der Uralischen Sprachen", in Denis Sinor, ed.: *The Uralic Languages: Description, History and Foreign Influences*, p.641.

19. V.V. Napolskikh: "Uralic fish names and original home", *Ural-Altische Jahrbücher*, Neue Folge Band 12, Göttingen 1993, p.35-57.

20. Even according to AIT defender Prof. R.S. Sharma (*Looking for the Aryans*, p.36), Mesopotamian inscriptions from the 16th century BC "show that the Kassites spoke the Indo-European language", and mention the Vedic gods "Suryash" and "Marutash".

21. That the Dasas, Dasyus (Iranian *dahyu*, "tribe") and Panis were Iranians and not "dark-skinned pre-Aryan aboriginals" is argued by a number of Indian anti-invasionist authors but also by Asko Parpola: "The problem of the Aryans and the Soma: textual-linguistic and archaeological evidence", in G. Erdosy: *The Indo-Aryans of Ancient South Asia* (W. De Gruyter, Berlin 1995), p.367 ff. The identification of Pakthas, Parshus and other tribes encountered by the Vedic king Sudas in the "battle of the ten kings" (related in Rg-Veda VII:18, 19, 33, 83) is elaborated by Shrikant Talageri: *The Aryan Invasion Theory, a Reappraisal*, p.319 ff.

22. A. Parpola in G. Erdosy: *The Indo-Aryans of Ancient South Asia*, p.355.

23. In the margin of the 1996 South Asia Conference in Madison, Wisconsin, Prof.

J.M. Kenoyer did a slide show on beads and jewels found in Central Asia: many of them were imported from the Harappan civilization.

24. Bernard Sergent: *Les Indo-Européens*, Payot, Paris 1995, p.398 and p.432.

25. Among the highest estimates is the 5% to 9% of Dravidian loans in Vedic Sanskrit proposed by F.B.J. Kuiper: *Aryans in the Rigveda*, Rodopi, Amsterdam 1991. On p.90 ff., he gives a list of 383 "foreign words in the Rigvedic language", including such obviously IE words as *aksha*, "axle".

26. Shrikant Talageri: *Aryan Invasion Theory, a Reappraisal*, p.156-175. To this effect, Thomas Burrow (in Thomas A. Seebok: *Current Trends in Linguistics*, Mouton, The Hague/Paris, vol.5, p.18, quoted by Talageri, *op.cit.*, p.162) already wrote that "there has been a certain amount of controversy concerning the question of non-Aryan loanwords in Sanskrit, and some scholars (P. Thieme, H.W. Bailey) have adopted a sceptical position in this respect. Alternate Indo-European etymologies have been offered for words for which a Dravidian or Munda etymology had previously been proposed, in some cases successfully (...) but more dubious in other cases."

27. Summarized by Edwin Bryant: "Linguistic Substrata and the Indo-Aryan Migration Debate", read at the 1996 Atlanta conference on the Indus-Saraswati civilization; he mentions Jules Bloch and H. Hock, among others, to this effect.

28. See e.g. D. McAlpin: "Linguistic Prehistory: the Dravidian Situation", in M.M. Deshpande and P.E. Hook, eds.: *Aryan and Non-Aryan in India*, Ann Arbor 1979.

29. R. Swaminatha Aiyar: *Dravidian Theories*, Motilal Banarsidass, Delhi 1987 (but written in 1923).

30. Quoted in Stefan Zimmer: *Ursprache, Urvolk und Indoger-manisierung*, Innsbruck 1990, p.25.

31. As remarked in 1952 by Oswald Szemerényi, quoted to this effect by Edwin G. Pulleyblank: "The Typology of Indo-European", *Journal of Indo-European Studies*, spring 1993, p.63-118, spec. p.63-64.

32. Edwin Pulleyblank: "The Typology of Indo-European", *Journal of Indo-European Studies*, spring 1993, p.106-107. The article is followed by two sharply critical pieces of comment, but the focus of their criticism is not the connection between Sino-Tibetan and PIE, though the authors do not conceal their skepticism of that point too. Remark that the claim of typological similarity with PIE, here made by Pulleyblank for Sino-Tibetan, is also made by others for North-Caucasian, and that the triangle is closed by yet other argumentations for a typological (and even lexical) relation between North-Caucasian and Sino-Tibetan, e.g. S.A. Starostin: "Word-final Resonants in Sino-Caucasian", *Journal of Chinese Linguistics*, June 1996, p. 281-311.

33. Edwin Pulleyblank: "The Typology of Indo-European", *Journal of Indo-European Studies*, spring 1993, p.109.

34. U. Libbrecht: *Historische Grammatika van het Chinees*, part III, Leuven 1978, p.3-4. In my opinion, the fertile and moderate-climate Yellow River basin itself is a more likely centre of dispersal.

35. Franklin Southworth: "Indo-Aryan and Dravidian", in M. Deshpande & P.E. Hook: *Aryan & Non-Aryan in India*, Ann Arbor 1979, p.205.

36. Franklin Southworth: "Indo-Aryan and Dravidian", in M. Deshpande & P.E. Hook: *Aryan & Non-Aryan in India*, p.206.

37. I. Dyen in G. Cardona: *Indo-European and Indo-Europeans*, Philadelphia 1970, proceedings of the Third Indo-European Conference, 1966, p.431-440.

38. It goes without saying that "primitiveness" in grammar says little about the civilizational level of a language community; Chinese is spoken by a highly civilized people, but its grammar strikes native speakers of German or Russian as very childlike.

39. Peter Bellwood: "An archaeologist's view of language macrofamily relationships", *Oceanic Linguistics*, December 1994, p.391-406.

40. Bernard Sergent: *Les Indo-Européens*, p.398.

41. B. Sergent: *Les Indo-Européens*, p.435.