EARLY INDO-URALIC LINGUISTIC RELATIONSHIPS: REAL KINSHIP AND IMAGINED CONTACTS [*]

1. This paper is based on one simple and natural assumption: there may be many riddles but no wonders in linguistic prehistory. Therefore the problem of early relationships between Indo-European and Uralic in its various aspects should be treated within the framework of the existing, sufficiently rich experience of historical linguistics, and the explanatory scenarios of prehistoric events suggested on the base of linguistic evidence must fit the criteria of experiential plausibility. It seems necessary to put this trivial claim forward again: all too often the early prehistory of languages is viewed as a *terra incognita* with its own unknown rules (which are therefore invented by some scholars freely and with vivid imagination).

But the leading role of linguists in the study of prehistoric epochs, and even the very fact of their professional participation in such study, are justified only if the phenomenon called «human language» underwent no principal changes at least since Upper Palaeolithic / Mesolithic times, so that:

— its structural types lay within the typological limits of variation attested in the contemporary world;

— the processes of its evolution did not differ from those attested in the contemporary world.

There are many reasons for assuming the validity of these restrictive clauses. The currently available typological and diachronic documentation covers geographically and culturally totally distinct regions, including those where the internal development has never reached the Neolithic stage, like aboriginal Australia or Amazonia [1]. The time span of ca. 15,000 years (or even of ca. 6,000 years, if we deal only with Proto-Indo-European or Proto-Uralic issues) is also relatively short in comparison with even the minimal estimates of the age of Homo sapiens and of its language (ca. 50,000 years). It must be stressed in this connection that the problem of language origins (rise of articulatory organs, formation of grammatical means, creation of roots…) belongs,

---

[*] Расширенный текст доклада на международном симпозиуме по доисторическим контактам индоевропейцев и уральцев (Tvärminne, Финляндия, 8–10.01.1999). Статья публикуется в «Proceedings» симпозиума (один из очередных томов «Mémoires de la Société Finno-Ougrienne»).

[1] Not long ago it could still be claimed that e. g. the Australian languages exhibit abnormal patterns of historic development hardly accountable for by the comparative reconstruction procedure, or that they are strongly aberrant in some other respects (e. g. [Boretzky 1984] or my own, perhaps too cautious, comments in [Хелимский 1986д: 258]). It becomes clear in the light of more recent research that «Australian languages change in a regular fashion, in the same way as Indo-European and other families» [Dixon 1990: 398].

in my opinion, to biology and anthropology rather than to linguistics, and even within this latter (cf. [Décsy 1977–1981]) it has nothing to do with the discipline called comparative and historical linguistics.

The inductive principle in prehistoric research advocated here means in practice:

(a) accepting only such claims and explanations which can be confirmed with exact analogies from the current experience of diachronic studies;

(b) disregarding any assumptions which cannot be confirmed by this experience (even if they cannot be definitely refuted) [2].

2. The experience of diachronic studies provides us with examples of only two types of language kinship.

2.1. The first and undoubtedly the most common type is d i r e c t k i n s h i p determined by d i v e r g e n t e v o l u t i o n. It is always created according to the same universal pattern, for which the development of the Romance language serves, due to complete historical attestation of all stages, as the best illustration:

— one relatively uniform language spoken on a relatively small territory

(Pre-Classical Latin of Latium, then Classical Latin of Central Italy);

— special historical premises favouring the spread of this language far beyond the original territory

(geopolitical energy and partly also technological superiority of Rome, its policy of conquests which led to the creation of the Roman Empire);

— export of this language to new territories in its dialectally and sociolectally differentiated forms constituting together a linguistic continuum

(Vulgar Latin = Proto-Romance in its regional variants);

— new historical situation interrupting or undermining the former socio-political and (relative) linguistic unity

(fall of the Roman Empire, new states in Pax Romanica);

— independent evolution of each of regional variants

(rise of French, Italian, Spanish, Rumanian, etc.).

The numerous less abundantly documented linguistic histories repeat (to the extent they are known) this general scheme up to minute details. Cf. such groupings and families as Slavic (Common Slavic somewhere between Vistula and Dnieper — Slavs as allies of Avars, Avaro-Slavic conquests in South-Eastern and Central Europe — spread of Slavic idioms across the Avar Khaghanate — decline and fall of the Avar Khaghanate, Hungarian conquest, later also Catholic-Orthodox schism — rise of Pol-

[2] This principle has naturally also other significant consequences. For example, available experience proves that good correlation between linguistic and genetic data is rather an exception than the rule, and most surely there exist no genes or gene-combinations responsible for using a certain language. Therefore even if the contemporary fashion of the co-operation between linguists and geneticists in prehistoric studies is sensible, the results obtained in each of the disciplines should be kept strictly apart. This remark (cf. also [Janhunen 1998: 9]) is addressed first of all to the international team of «root-finders», whose variegated and often extravagant ideas have been summarised in [Künnap 1998] and who tend, on the contrary, to overestimate the relevance of genetic and even of geological and gynaecological [Ibid.: 59] research for the Uralic studies. In their practice the proclaimed co-operation leads too often to a failure to distinguish between the origins («roots») of the population of Finland and the origins of the Finnish language.

ish, Bulgarian, Russian, etc.), Mongolic (Common Mongolic in Western Mongolia — Chinggis Khan — spread of Proto-Mongolic/Middle-Mongolic dialects in Central Asia — disintegration of the Mongol Empire — rise of Khalkha, Buryat, Oirot, Monguor, etc.), Permic (Common Permic at Lower Kama — migration to Upper Kama and further northward — Proto-Permic dialectal continuum in all these regions — Volga-Bulgarian conquest and penetration of Slavs to the Upper Kama — rise of Udmurt and Komi). This list of examples can be continued with those from e. g. Germanic, Indo-Aryan, Bantu, Samoyedic, Turkic, Polynesian, etc. (to quote only the most evident and better investigated cases).

Accompanying circumstances (substratum, secondary interaction after the breach of unity but before the disappearance of mutual comprehensibility, and spread of a *koine* across the individual regions which disguises the traces of the original continuum) are no less universal, but they do not introduce any major changes into the picture of kinship. Therefore the notion of the genealogical tree (with necessary refinements) was and remains fully valid.

2.1.1. Direct kinship is in principle traceable in the daughter languages. The more distant this kinship is, the less numerous and obvious such traces are and the more they are concentrated only in the most stable fragments of the language system:

main grammatical markers — if not swept away by reduction and typological drifts;

pronouns;

numerals, if the proto-language existed after the introduction of decimal counting (in Eurasia — ca. 5,000 — 6,000 years ago);

basic vocabulary.

2.1.2. The above statement concerning the numerals calls for a few special comments. The problems related with the origin of numerals, their etymological transparency, and their role in proving linguistic kinship have been reviewed in detail and persuasively discussed in [Honti 1993: 25–37], and I can only echo the author's reserved and moderately pessimistic estimates of the perspectives of the etymological research in this domain [3].

However, the well-known fact that all Finno-Ugric languages have common numerals ('2' — '6', '10', '100', partly also '1' and '7'), while the common Samoyedic numerals are entirely different (with the exception of PUr. **kekta* '2' [4]) seems to be non-accidental. According to the prevailing (and well substantiated, both linguistically and historically) dating, PUr. time ended ca. 6,000 years ago, while PFU is about 1,000 years younger. Exactly this difference in time seems to be significant: as far as I know, there are no language groups or families of the Northern and Central Eurasia aged 5,000 and less which do not exhibit common numerals, and there are no families aged 6,000–7,000 years and more which do. It must be added that normally the common numerals of a language family or group are perfectly preserved in all daughter languages as an entire system. Later replacements affect usually only the numerals for '8' — '9', '1', '7', as well as derived and compound numerals; the situations when e. g.

[3] Cf.: «Ich bin überzeugt, daß die Zahlwörter der heute gesprochenen Sprachen und der nur aus ihrer Schriftlichkeit bekannten ältesten Sprachen des Altertums nichts mehr von der Entstehung und Entwicklung dieser Kategorie verraten» [Honti 1993: 26].

[4] I reconstruct this proto-form as disharmonic.

the words for ‘3’ and ‘5’ are original, while the word for ‘4’ is innovative, are abnormal. My observations are reflected in the following table [5]:

FAMILY / GROUP OF LANGUAGES	APPROXIMATE AGE (years)	PRESENCE OF COMMON NUMERALS
Finno-Ugric	5,000	+
Samoyedic	2,000	+
Uralic	6,000	– (but cf. **kekta* ‘2’)
Mongolic, Turkic, Tungusic	between 800 and 2,500	+ (in each of the groups)
Altaic	(much?) more than 6,000	–
Late Indo-European	5,000 – 6,000	+
Indo-Hittite	6,000 – 7,000	± (cf. Hit. *me*(*i*)*u*- ‘4’ [6])
Semitic	6,000	+
Chukchi-Kamchadal	(much?) less than 5,000	+
Dravidian	5,000	+
Elamite-Dravidian	6,000 or (much?) more	– ? (at least ‘3’ is not common[7])
Yeniseic	2,500	+
West Caucasian	relatively young	+
East Caucasian	probably older than IE	± (partly very problematic [8])
North Caucasian	probably much older than IE	– (no reliable counterparts [9])

The dependence between the time depth and the presence of common numerals probably indicates that some major change in counting practices — perhaps the introduction of decimal counting — spread throughout Eurasia some 5000–7000 years ago and led to establishing the new, and since then very stable, systems of numerals. There are therefore little (I would perhaps dare to say: no) chances of reconstructing numerals

[5] There is regretfully no possibility to argue here for individual dates and estimates used in this table, but (a) they are mutually calibrated insofar as such calibration is possible; (b) it can be hoped that most specialists on individual families would find them reasonable rather than voluntary.

[6] The phonetic shapes of most IE numerals are not attested in Hittite texts and other Anatolian sources, or the etymological identifications are problematic [Гамкрелидзе, Иванов 1984: 844–848; Mallory, Adams 1997: 397–405]. At least Hit. *tēri* ‘3’ fits the Late IE system of numerals, and at least *me*(*i*)*u*- ‘4’ does not.

[7] See [McAlpin 1981: 118]. Most Elamite numerals are not attested.

[8] The North Caucasian Etymological Dictionary [Nikolayev, Starostin 1994] suggests the common origin of most East Caucasian numerals. In some cases, however, the comparison runs across serious and only tentatively solved problems (‘4’, ‘6’), and — in view of the fact that usually numerals are preserved as an entire system — the lack of Nakh correspondences for ‘2’ and ‘3’ looks very suspicious. On the other hand, the word for ‘20’ is apparently common!

[9] S. Nikolayev and S. Starostin [1994] attempted to identify the following West Caucasian numerals with their East Caucasian semantic counterparts: ‘1’, ‘2’, ‘3’, ‘5’, ‘6’, ‘10’, ‘100’. In most cases, however, this identification requires comments like ‘Phonetically irregular’ or ‘Complicated case’ — which, taking the predominantly monosyllabic stem structure of West Caucasian into account, signals the Procrustean character of the comparison. I would suggest abandoning the attempts to reconstruct the Proto-North-Caucasian numerals (with the exception, perhaps, of the word for ‘1’) and assuming that the branches of North Caucasian diverged too early to have common numerals.

for the proto-languages which disintegrated much earlier, like Altaic, Nostratic, or Sino-Caucasian.

This cultural innovation seems to originate, like most other innovations of the Neolithic, from the Near East and North-Eastern Africa: the common vocabulary of Semito-Hamitic, which was probably sufficiently older than Semitic, includes several numerals (preserved in Semitic, Berber, Egyptian, partly also in Chadic).

I am naturally aware of the attempts to reconstruct the Altaic numerals or to find Nostratic cognates for the Uralic ones (most recent: [Blažek 1996/1997, 1997]). For the most part, the highly questionable etymological comparisons which result from such wishful attempts (on a line with some very interesting marginal comments)[10] confirm, in my opinion, the above statement rather than refute it. On the other hand, accepting this statement means that the criterion of having common numerals[11] simply loses any relevance when dealing with the oldest proto-languages, so that the effort can be effectively spared.

2.1.3. Due to the fact that the number of common elements in two or several languages of the same origin, as well as the external similarity of these elements, tends to diminish steadily in the course of time, their genetic kinship becomes less and less transparent [12]. This loss of transparency manifests itself directly in the spread of doubts concerning the validity of the corresponding genetic identification and in the loudness of critical voices.

The following table illustrates, with selected examples, the transparency of genetic kinship as dependent on the distance between the nodes of the genealogical tree (= *relative* age of the common proto-language) as well as on the choice between group comparison and binary comparison.

	less than 3000	3,000-6,000	6,000-9,000	9,000-12,000
Cluster of languages	**obvious even to non-specialists** (*Slavic; Germanic; Ob-Ugrian; Samoyedic; Bantu; Turkic; Yeniseic*)	**clear to specialists** (*Indo-European; Uralic; Chadic; Austronesian*)	**traces** (*Semito-Hamitic; Uralo-Yukagir; Altaic; Nostratic* with proto-language reconstructions)	**?** (*Indo-Uralic / Nostratic* without proto-language reconstructions)
Two languages	**clear to specialists** (*Greek* and *Sanskrit; Finnish* and *Lapp*)	**traces** (*English* and *Hindi; Nenets* and *Cheremis; PIE* and *PUr.* on the base of reliable proto-language reconstructions)	**?** (*Estonian* and *South Yukagir; Yakut* and *Manchu; Korean* and *Japanese*)	**???** (*Hungarian* and *English* without proto-language reconstructions)

[10] No more lucky was an attempt to trace the origin of numerals in Samoyedic ([Joki 1975]: perhaps the least successful scholarly contribution of this outstanding etymologist).

[11] This criterion is not infrequently used by the opponents of the Altaic and Nostratic theories, cf. e. g. [Doerfer 1966: 100–101].

[12] However, I would not side with J. Janhunen (s. a.: 2) in claiming that this kinship disappears («language change is really so rapid that genetic relationships inevitably fade away in a relatively short span of time [so that] the maximum lifespan of any language family is some 10,000 years») — just as I would not say that my distant cousin (whom I may have never met and will never recognize when meeting) is not my relative. At best, one can distinguish between a «practically important» and a «practically negligible» genetic kinship.

2.2. The second, much less common type of language kinship is l a t e r a l k i n s h i p or grafting, determined by relexification and observed in the case of «m i x e d» l a n g u a g e s. It seems that all known cases of lateral kinship also fit a single pattern.

The sociolinguistic prerequisite of lateral kinship consists in a situation, in which the representatives of an ethnic or a social group want or need to speak the target (ancestral, titular, prestigious) language but, being adults grown up with another language, are not in a position to cope with the target grammar. Therefore they achieve their aim — partly at least — by relexifying their first language from the target language (case of Anglo-Romani [13], Ma'a or Mbugu [14]), possibly adding also some simpler grammatical elements from this language (case of Copper Island Aleut [15]), or by superimposing the lexical stock of the target language on an appropriate and very simple grammatical framework partly resembling that of their first language (case of most pidgins and creoles, see [Holm 1989]). A linguist should not necessarily evaluate the results of such creative attempts in the same way as the speakers of the corresponding idiom. They naturally tend to view their aim as achieved and to identify their idiom with its target idiom, but there are usually more reasons to treat it on a line with «secret languages» and argots, which also normally preserve the grammatical structure of the starting language intact.

In any case, these processes:

(a) produce only new languages in which grammar and vocabulary are of different origin;

(b) produce in each case one «mixed» (relexified) language rather than a new family;

(c) are due to abnormal and therefore rare sociolinguistic situations;

(d) rarely become independent of these situations and are therefore mostly short-lived.

The recent studies devoted to the above problem do not contain any essential data which do not fit the above pattern — even if they abound in assumptions and guesses aimed at questioning its universal validity. If we leave aside the trivial statements («all languages are mixed») or clumsy attempts to qualify such loan-words-saturated languages as Yiddish or Hungarian as mixed (rather than as belonging to the Germanic, resp. to the Ugric group), the category of mixed languages appears to be in most parts of the world very small and relatively homogenous as far as the sociolinguistic and

[13] See [Thomason, Kaufman 1988: 103–104]: «its grammar is simply that of English». The target language in this case is naturally Romani, the ancestral and titular language of Gipsies.

[14] See [Goodman 1971; Порхомовский 1982: 215–216; Thomason, Kaufman 1988: 223–228]. According to the last mentioned book (and to some previous studies on «the strange case of Mbugu»), the typically Bantu grammar — rather complicated, with word classes, etc. — of this language is borrowed from the Bantu surrounding. Lexical borrowings (to any extent, up to entire relexification, as in the case of secret languages) are natural processes, but I fail to see how an entire grammar can be «borrowed». Obviously the relexification of a Bantu language with a Cushitic lexicon can be the sole explanation of this strange case. The data on the social background of Ma'a («resisters of total cultural assimilation to their Bantu-speaking milieu», who sought isolation «so that they could continue to follow their own customs») confirm the assumption that their language arose as an answer of assimilated Ma'a (who already shifted to Bantu) to the demand to restore their non-Bantu linguistic and cultural identity.

[15] See [Асиновский, Головко, Вахтин 1983; Thomason, Kaufman 1988: 233–238; Golovko, Vakhtin 1990; Vakhtin 1998].

structural mechanisms of their rise are concerned. It can be added that the data from the technologically backward linguistic areas (like Australia) do not appear to suggest that the phenomenon of lateral kinship was in the early prehistoric past more wide-spread than in the contemporary world.

2.2.1. A mixed language that survives may constitute a problem for genetic linguistics insofar as its lateral kinship with the target language can be mistaken for a direct kinship with it (especially if a study is mostly confined to the vocabulary: that is why some of my colleagues, like Sergei Starostin, insist on the genetic identification of mixed languages with their targets).

2.3. The much-discussed notion of «contact kinship», or «areal affinity»[16], is, from the viewpoint of inductive experience, a mere phantom. Even in the case of the most intensive areal interactions (Romano-Germania, the Balkans, Volga-Kama, Upper Yenisei, India, the Sino-Thai region, aboriginal Australia) the convergency processes result only in far-reaching typological similarities and in abundant borrowings in non-basic vocabulary. A *Sprachbund* producing an outcome similar to the outcome of direct or even of lateral kinship (serial cognates among inflectional and main derivational morphemes, incl. pronouns and other grammatical words; similarities in the domain of irregular and suppletive forms; unity of basic vocabulary or traces of such unity) has never been attested.

The lengthy reflections on how non-cognate languages could have been transformed into a language family due to prolonged and intensive contacts are a poor substitute for examples that are lacking.

Absence of examples does not prove that «contact kinship» due to areal convergencies is impossible. But the above formulated inductive principle of research into linguistic prehistory permits us to dismiss this notion as purely speculative.

2.3.1. Attempts to view the relationships between languages through the prism of convergent rather then divergent patterns have a long history — from Hugo Schuchardt via Giuliano Bonfante and especially Nikolay Trubetzkoy [1939] to contemporary Anti-Altaicists and «rebel» Uralists («root-finders»). Still, the above-mentioned predecessors overtly rejected the classical comparative and historical linguistic methodology of the Junggrammatiker[17]. I'd suggest their contemporary spiritual heirs do the same. A borderline or even a schism between different schools of linguistic thinking is preferable to a misleading pseudo-unity.

3. The admission of direct kinship between PIE and PUr. (K. B. Wiklund, B. Collinder, V. M. Illič-Svityč) is the logical consequence of the inductive principle. The comparison of PIE and PUr. provides us with a picture which corresponds most exactly to the theoretical expectations: the scanty and well known

[16] Other designations or similar concepts: language diffusion, allogenetic relationship, (Hung.) *rokonságszerű kapcsolat*, proto-language as *lingua franca*, etc. (see [Хелимский 1986д: 253–260]).

[17] A consequent adherent of «language convergencies» must inevitably finish with stating that e. g. French and Spanish existed as such on corresponding territories from times immemorial, and acquired their present Romance shape due to mutual convergency under the aegis of Rome.

stock of cognates (see [Collinder 1965; Иллич-Свитыч 1971: 6–37]) is found mainly among the most stable elements of both proto-languages.

3.1. An excellent opportunity to demonstrate the direct relationship between the stability of words (stems) and their occurrence in the common vocabulary of several Eurasian language families is provided by the paper presented at this symposium by Kaisa Häkkinen. Evaluating the distributional certainty of Uralic etymologies as they are reflected in contemporary etymological dictionaries, she singled out a group of 18 items «with 100 % etymological certainty», meaning that their counterparts can be traced throughout the Uralic family, and another group of 23 items «with 90 % certain etymology», meaning that their counterparts are missing or dubious in only one of 10 relevant languages or branches [Häkkinen 1999: 11–12].

The criteria of selection are purely formal (so that some of the most popular and absolutely certain Uralic etymologies, like those for Fi. *kala* 'fish' or *vesi* 'water', remain outside the list only because of the loss of corresponding words in Permian, Saami or Ostyak), but the results are by no means accidental: belonging to the first group means that an item was capable of being preserved in the course of ca. 37,000 years [18], and belonging to the second group means withstanding potential replacements in the course of ca. 33,000 years. It is natural to expect that the items which proved their stability in «Post-Uralic» times possessed this quality also earlier, in the «Pre-Uralic» period.

It appears that among the 18 items of the first group 12 have IE and other Nostratic parallels according to [Иллич-Свитыч 1967, 1971, 1984] (in 9 cases in IE [19], in 3 cases only in non-IE Nostratic [20]; in 10 cases these parallels are accepted and mentioned also in the UEW [21]), and 2 more are supplied with indications of non-Uralic parallels in UEW [22].

The second group includes, among 23 items, 6 numerals, none of which (with the exception of *kaksi* '2' [23]) has Samoyedic counterparts and none of which has Nostratic parallels (this confirms that the subgroup of numerals appeared and acquired absolute stability only in Finno-Ugric times, cf. above, 2.1.2). The situation with the remaining 17 items resembles that of the first group, though the figures are a bit lower: 8 Nos-

[18] This approximate figure reflects the sum of lengths (in years) of all branches of the family tree connecting the nodes for the 10 relevant Uralic languages/groups.

[19] The Uralic proto-forms of Fi. *ala-* 'under-' (IE **Hel-* 'deep', Alt. **ala* 'under', Yuk. *al*), *ku*(*ka*) 'who' (IE **kʷo*, Alt. **qa-*/**qo-*, Yuk. *xadi*, etc.), *me* 'we' (IE **me-s*, Alt. **bä*/**män-*, Yuk. *mit*, etc.), *mi*(*kä*) 'which' (IE **mo-*, Turk. **mi-*, etc.), *minä* 'I' (IE **me*/**mene-*, Alt. **bi*/**min-*, Yuk. *met*, etc.), *nimi* 'name' (IE **nem-*, **nōm-n̥-*, Yuk. *niu*), *punoa* 'to weave' (IE **spen-*, Drav. **puɳʌ-*, etc.), *suoni* 'sinew' (IE **sneĤ*(*u̯*)-, etc.), *tuo* 'that' (IE **to-*, Yuk. *tuŋ*, etc.)

[20] *kadota* 'to disappear' (Alt. **qala-*, etc.), *sydän* 'heart' (? Alt.), *uida* 'to swim' (Tung. **uju-*).

[21] Exceptions: *mi*(*kä*) and *sydän*.

[22] *maksa* 'liver', *nuoli* 'arrow' (but the Alt. parallels are very problematic and can hardly be viewed as pieces of Nostratic evidence).

[23] This word must actually belong to the first group. The reason for lowering its certainty are the two question marks put by K. Rédei in the UEW before the Samoyedic counterparts of *kaksi*. However, none of the statements which substantiate his doubts («Die Zugehörigkeit der sam. Wörter ist unsicher, weil ihr anlautender Konsonant sowohl einem **k* als auch einem **c* entsprechen kann. Die Vertretung der ursprünglichen inlautenden Konsonantenverbindung ist unregelmäßig» [UEW: 139]) can be accepted.

tratic parallels (6 also in IE [24], 2 only in non-IE Nostratic [25]) plus 5 indications of external parallels (which mostly cannot be viewed as Nostratic cognates) in the UEW.

The extraordinary high concentration of words with evident (and semantically exact) external parallels in the most stable strata of the Uralic vocabulary [26] fully corresponds to the theoretical expectations outlined in 2.1.1.

Needless to say, treating the above items as «borrowings» (from IE or Altaic into Uralic, or vice versa, or from Uralic into Yukagir, etc.; see also below, 4.1) creates an absolutely improbable picture of a language which borrows all crucial and most stable elements of its lexical (and morphological) structure from a foreign source. «Borrowing on such a scale is practically equivalent to borrowing the entire structure of the language, to switching to another language, and this leads us again to admitting the genetic relationship of the proto-languages in question» [Иллич-Свитыч 1971: 5].

3.2. A lateral kinship between PIE and PUr. is hardly imaginable, since:

(a) the similarities are found both in grammar and in vocabulary;

(b) nothing in the linguistic structures of PIE or PUr. implies their «mixed» past;

(c) the scenario of lateral kinship is statistically rare. (However, I would not completely exclude a version of this scenario in the case of Uralic-Yukagir kinship.)

3.3. Postulating the Indo-Uralic (and wider) genetic kinship is referred to here as the Nostratic theory. I feel obliged to remark, however, that accepting this theory is not the same as accepting uncritically the entire etymological material contained in Illič-Svityč's works and especially in recent Nostratic research [27] (many critics of Nostratism — cf. [Doerfer 1973a, 1993; Reinhart 1988] — prefer to disregard this distinction). The core of the problem lies in the contradiction between the subjective wish to find new proofs of the prehistoric kinship and the objective scarcity of comparable data (see 2.1.1).

4. The reasons for rejecting some alternative treatments of the set of PIE-PUr. cognates are provided by the same inductive principle.

4.1. The similarity between some lexical items in PIE and PUr. is so obvious that this material cannot be simply discarded by the opponents or doubters of the genetic kinship. The idea of borrowing from PIE into PUr. therefore suggests itself.

K. Rédei, who advocates this approach, restricts the lexical evidence to a group of 7 words which he treats as PIE loans in PUr. The PUr. words in question are (the reconstructed forms and meanings are quoted after Rédei): **miγe-* 'geben; verkaufen', **muśke-* (**mośke*) 'waschen', **nime* 'Name', **sẹne* (**sōne*) 'Ader; Sehne', **toγe-*

[24] *kuolla* 'to die' (? IE **g^{u}e(H)l-*, Drav. **kol-* 'to kill', Kartv. **q̇wil-* id.), *kyynär* 'elbow' (IE **ĝenu/ĝneu* 'knee', etc.), *käsi* 'hand' (IE **ĝhes-*, Drav. **kac-*, etc.; cf. [Хелимский 1990a]), *mennä* 'to go' (? IE **men-* 'to step upon'), *pelätä* 'to fear' (IE **pelH-*, Alt. **pēli-*, etc.), *tämä* 'this' (IE **te-*, Alt. **ṭä-*, Yuk. *tiŋ*, etc.).

[25] *e-* negative verb (Alt. **e-*, etc.), *veri* 'blood' (? SH **br-*).

[26] On the whole, the share of PUr. and PFU stems with external parallels (in [Иллич-Свитыч 1971–1984] and [UEW]) hardly exceeds 10–15 %, and very often they are extremely problematic, especially in comparison with the majority of cases from the preceding footnotes.

[27] I suspect that many «long-ranger» studies will not withstand the procedure of imitative absurd etymologizing as suggested and applied in [Helimski 1987c; Хелимский 1989г].

'bringen, holen, geben', **waśke* 'irgendein Metall, Kupfer', **wete* 'Wasser' [Rédei 1986: 40–43]. J. Koivulehto, who essentially shares this viewpoint [28], adds the following PUr. items to this list: **pele-* 'fürchten, sich fürchten', **puna-* 'spinnen, flechten', **pura* 'Bohrer; bohren' [Koivulehto 1994: 137–139, 1999: 1–2].

The semantic spectrum of presumed loan-words is in itself the strongest, most scathing argument against the version of borrowing.

The inductive basis of studying loan-words is rich and uniform: in all situations of language contact, the words to be borrowed first and foremost are «useful borrowings» — technical and other cultural terms as well as the terminology of local natural milieu. If the contacts are long and especially intensive, the penetration of some elements of the basic vocabulary – of «useless borrowings», because they serve only as replacements for the words which already existed – can be observed and expected; but these always remain quantitatively overshadowed by the «useful» ones. No known cases of «bilingualism in adjoining regions» or «motivation through the economic and cultural prestige» [Rédei 1986: 21] produce a different effect.

Here we are faced, on the contrary, with a list including four verbs with the most basic semantics – «give» (the meaning «sell» in Uralic is obviously derived, s. [UEW: 275]), «bring», «wash», «fear», with three nouns of a no less basic nature – «sinew», «name» (in no human language can the idea of naming be a cultural innovation, contrary to [Rédei 1986: 41]), «water», and two notions which ceased to be technical novelties long before the PIE and PUr. epoch – «spin» and «drill». No «useful» borrowings, which abound e. g. in the well-known lists of Indo-European (Pre- and Early Aryan) loan-words in Finno-Ugric and Finno-Permic [Rédei 1986: 43–64], are present in the list [29].

Therefore, in accordance with the inductive principle, this list (in its present form, at least) can be safely discarded from consideration as evidence of language contacts between PIE and PUr. The parallels in question can testify either to a genetic relationship or (if the adversaries of the Nostratic theory prefer this) to erroneous etymologizing.

4.2. Other attempts to prove such contacts deal with more or less the same lexical set, occasionally enlarging it by adding e. g. personal, demonstrative, and interrogative pronouns (the corresponding items are partly presented in the footnotes to 3.1). K. Rédei, on the other hand, excluded this material from consideration entirely, since for him «die lautliche Ähnlichkeit der uralischen und indogermanischen Pronomina beruht aller Wahrscheinlichkeit nach auf Lautsymbolismus» [1986: 19]. This view is

[28] He adds, however, a criterion according to which close phonetic similarity and/or regular phonetic correspondences between PIE and PU forms serve as evidence of borrowing. This criterion is not fulfilled by «Gleichungen, bei denen (bisher) keine genaue lautliche Übereinstimmung erreicht werden konnte; unter diesen letzgenanntren Fällen könnten sich theoretisch urverwandte, indo-uralische Wörter finden» [Koivulehto 1994: 137]. The list of these provisionally unexplained and therefore provisionally related words includes PU **nime* 'name' (the laryngealistic reconstruction of IE **h*$_1$*neh*$_3$*men-* does not fit the author's ides of IE laryngeals as being reflected and preserved in Uralic as **k* or **š*), s. [Ibid.: 140].

[29] The only potential Kulturwort of the set, «metal», must have originally denoted «ore (as a special type of stone)» and have belonged to the same group of the basic vocabulary as «stone», «sand», «pebble», «clay», etc. On the other hand, this parallel can be viewed only in the context of contacts between Samoyedic (or some other later branch of Uralic) and Tocharian, and not as a piece of earliest Indo-Uralic evidence [Szemerényi 1988: 172].

expressed also by G. Doerfer [1973a: 84, 1993: 25]. The morphological parallels between PIE and PUr. (e. g. *-*n* of the genitive and *-*m* of the accusative) receive a similar treatment.

The idea that pronouns are of phonosymbolic origin («theory of elementary kinship») belongs obviously to the domain of glottogony. Even in this case, however, it does not explain the fact that the «primary elements» *m* (1Sg), *t* (2Sg), *k* (interrogation), *t* (deixis) systematically occur in these functions only in the languages for which the Nostratic kinship is postulated, while in other languages they fulfil these functions no more frequently than other consonants. On several occasions I asked the adherents of this idea, whether the fact that in so many branches of Nostratic the elements *m* (1Sg) and *t* (2Sg) are followed by *n*, especially in oblique pronominal forms, should also have a phonosymbolic meaning — but never managed to get an intelligible answer. There is therefore every reason to insist that the theory of elementary kinship as an explanatory model finds no support in the experience of diachronic studies.

4.3. According to the most likely localisation of the proto-homelands, PIE before its disintegration was spoken somewhere to the north of the Black Sea (and/or of the Caspian Sea), while the PUr. habitat was restricted to Western Siberia and the north-eastern corner of Europe, which leaves no possibility for any direct contacts between the speakers of the two idioms (they were only two widely-spaced areas on the variegated linguistic map of the 8.–7. mill. B. P., in no way suggesting the future spread and glory of their descendants) [30]. The search for PIE loan-words in PU, or vice versa, is probably deemed to be continued, but I would qualify this occupation as lacking in perspective.

5. Somewhat similar considerations (stemming from the same inductive principle) also argue for the rejection of the majority of the etymologies suggested by J. Koivulehto and his followers (together with the newly rewritten pages in the history of Indo-Uralic, Germanic-Fennic, etc. contacts).

In [Хелимский 1995] (cf. also [Helimski 1990a: 29–33; Хелимский 1997]) I have attempted to summarise those features that distinguish these newly coined etymologies from more traditional results in the study of IE loan-words:

Lexical scope: the loan etymologies of verbs and adjectives occur almost as frequently as the etymologies of nouns; the designations of qualities and properties, abstract notions and valuations are numerically superior; in too many cases the loan etymologies are proposed for the names of universally known objects and elementary actions.

Semantics: the meanings of the FU words and their presumed IE sources often differ, so that intermediate semantic shifts which are not attested must be postulated (the rich experience of IE linguistics makes the demonstration of the possibility, in principle, of such shifts a too-easy task).

Stem structure: quite often the presumed sources differ from the words which are actually attested in the IE languages in the presence or absence of a suffix, in their *Ab-*

[30] Very different views on the ethnic and linguistic prehistory of Western Eurasia are being developed e. g. in [Pusztay 1995; Sammalahti 1995; Julku 1996, 1998; Carpelan 1998] as well as in some papers of the present symposium.

lautstufen, etc. (here, too, the rich potential of studies on IE stem-formation is exploited).

Phonetics: in addition to long-established rules J. Koivulehto makes use of new ones, which, however, can be illustrated only with other etymologies belonging to the same author.

Each of these features is suspicious in itself; their combination deprives the corpus of etymologies obtained with such methods of any cognitive value (which certainly does not exclude the acceptability of some — relatively few — new etymological findings).

More recent publications by J. Koivulehto (e. g. [1994, 1999; Койвулехто 1997]) demonstrate once again his brilliant inventiveness in coining etymologies and his highest technique in dealing with IE materials, but give no reason to change this position. The following example of his newer etymological findings can be quoted:

> «FM **veneš* ‘boat’ (> Finn. *vene*(*h*) ‘boat’, Lapp. Mord. id.)
> ← IE/Pre-Aryan/Early Indo-Aryan **wen*-(*e/o*-) > OInd. *ván*- (root noun) ‘wood’, *vána*- (n.) ‘wood, tree; timber; wooden vessel’. — **veneš* was an “Einbaum”: a dug-out stem used as a canoe/boat/vessel» [Koivulehto 1999: No. 43].

The reflections on the circumstances under which this borrowing could have occurred inspired me to compose the following sketch:

5.1. A Pre-Finno-Mordvinian comes to a neighbouring Proto-Indo-European village and looks around in bewilderment.

— What are you looking for here? — someone asks him (a Pre-Aryan, as it turns out later — though there are also many Pre-Greeks, Pre-Slavs, Pre-Balts and especially Pre-Germans dwelling in the same village).

— Oh, I’d like to borrow a word for boat… — the guest answers.

— What are you speaking about? Do you want to borrow one of our boats?

— Why should I? We have plenty of boats ourselves. In fact, we are a long way ahead of you in boat-making! How else could we become skilled fishermen? I only need your *word* for boat!

— Don’t you have your own word for it?

— Surely we do! But you know, nowadays it’s all the rage — to use Indo-European loans!

— Well, — the pre-Aryan scrunches up his brow. — Naturally we do have a name for boat. It is **nāus* — everybody, except these stupid and stubborn Pre-Slavs and Pre-Balts, knows and uses it! But I just cannot lend this word to you! I need it for myself, and for my Old Indic offsprings, who will call boats *nāu,* and for my Ossetic descendants, so that they could call them *naw*! No, you won’t get this precious lexical item!

— What shall I do then? I cannot come back empty-handed. Maybe you will find for me something less valuable or little needed, if you have such a thing? And you must have, Pokorny tells us that your language is so rich in stems!

— He is right, we do have some other boat-names, **aldhu*-, **(s)kolmos*, and **plou̯os*, for example. But lending them is out of question, forget it! We Indo-Europeans need these items for ourselves, if we are going to have our languages spread over all continents!

— Have pity, give me something, at least! — moans the poor creature.

— I've had enough of your begging! Here, take the word **wen-(e/o-)* — this is the only one I can give you! At present nobody really uses it here — this word will emerge only in Sanskrit as *vána-* and in Avestan as *vanā-*, without any Indo-European etymology and without any trace of the vowel *e*. So nobody will now notice it is missing. But I must warn you, this word does not really denote a boat! It is a word for tree, or for wood, or for timber. At best you can refer to a chunk of wood or a wooden vessel, like a bucket or a trough, with this name...

— No matter, our boats are after all no less wooden then troughs! You know, sometimes we just use dug-out stems as canoe boats! That will suit me! Thank you very much indeed, now I can head home with this wonderful new loan!

— Hey, wait a moment! You cannot borrow **wen-(e/o-)* just so as it is. What if one of our guys hears and recognises it? He'll take it back, and I'll get into trouble for squandering words! You must disguise the loan. Look, you may add some unusual non-Indo-European suffix to it. For example, *-š* — this will be a proper disguise.

— What a wise idea! I will do so. Many thanks again, it was so kind of you!

And the happy Pre-Finno-Mordvinian leaves the village whispering: «**Veneš*, **veneš*! How sweet these Indo-European words are!»

5.2. I do not think that the etymological proposal by J. Koivulehto is much more realistic than the above dialogue.