Juha Janhunen (Helsinki)

Proto-Uralic—what, where, and when?

1. Introduction

The traditional picture of Proto-Uralic we have today was created by several generations of comparative linguists, starting with M. A. Castrén (see Janhunen, forthcoming) and ending with the synthetic surveys of Pekka Sammallahti (1988) and Daniel Abondolo (1998). In this framework, Uralic is understood as a regular language family whose members represent the divergent, and basically binary, branches and sub-branches of an originally uniform protolanguage. Proto-Uralic was a fully developed natural language that was chronologically far beyond the glottogonic stage. Its structural properties, as far as they can be reconstructed, may therefore be assumed to have been similar to those attested in modern natural languages. The deepest dividing line within the family is traditionally assumed to exist between Finno-Ugric in the west and Samoyedic in the east. For various reasons, subsequent (Post-Proto-Uralic) diversification has been more profound, or is better preserved, within the Finno-Ugric branch, which is today represented by as many as seven major sub-branches, including (from west to east:) Saamic, Finnic, Mordvinic, Mariic, Permic, Mansic (incl. Hungarian), and Khantic.

In practice, all adherents of the traditional framework have always been conscious of certain problems and limitations that call for minor modifications to the approach. For instance, it is generally acknowledged that the protolanguage was not strictly uniform but dialectally diversified, like any natural language. Also, the branching of the language family need not always have taken place in a binary way, and, in any case, there are isoglosses that cross branch boundaries, including even the boundary between Samoyedic and its immediate western neighbours (Khantic, Mansic, and Permic). Even so, many Uralists agree that the classic family-tree model still remains the best for describing internal family relationships. As a possible modification of the binary family-tree, the fuzzy 'bush' model has been proposed by Kaisa Häkkinen (1984), later followed by the linear 'comb' or 'rake' model of Tapani Salminen (2002). However, even these modified models recognise the validity of Uralic as a language family, as well as the relevance of the comparative method as a diachronic tool.

The Quasquicentennial of the Finno-Ugrian Society.
Suomalais-Ugrilaisen Seuran Toimituksia = Mémoires de la Société Finno-Ougrienne 258.
Helsinki 2009. 57–78.

More recently, the conventional framework of Uralic studies has been challenged from two points of view. On the one hand, the so-called Roots Group, led by Kalevi Wiik (e.g. 2004) and anticipated by János Pusztay (1996), has proposed that the Uralic comparative corpus, or at least a considerable part of it, should be explained as the result of areal convergence, rather than genetic divergence. If this were the case, there would have been no single coherent Proto-Uralic language, but, rather, two or more regional protolanguages and centres of expansion. In this context, Proto-Uralic has also been described as having been formed as a regional *lingua franca* (for a critical review of the issue, cf., e.g., Jaakko Häkkinen 2006). On the other hand, it has been claimed, notably by Angela Marcantonio (2002), that the entire Uralic comparative corpus is simply not valid and thus requires neither a divergence nor a convergence explanation. According to this view, the conventional Uralic comparisons and reconstructions are statistically unlikely to be true. This would be especially so since the comparative corpus shared by Finno-Ugric and Samoyedic is very small, comprising hardly more than 200 lexical items.

One might think that such new points of view have 'endangered' the existence of Proto-Uralic as a valid diachronic entity. This is not the case, however, for the principles and methods of comparative linguistics, created during the 19th century, are solid enough to make any major 'revolution' in the discipline impossible. As is well known, the comparative method has effective tools for distinguishing between divergence and convergence, as well as between cognates and accidental lookalikes. Moreover, although the 'revolutionaries' have raised doubts about the comparative method in the field of Uralic studies, this method continues to be actively used by specialists on virtually all of the other language families in the world. There is, consequently, no serious reason to question the existence of the Uralic language family, nor the validity of Proto-Uralic as the reconstructed proto-form of the Uralic languages. A more relevant question is how much effort should be devoted to arguing against paradigms that are based on an insufficient understanding of the discipline. The situation is analogous to that in the natural sciences, where the theory of evolution is being challenged by religious fundamentalists propagating unscientific 'alternative' 'models', such as 'creationism' and 'intelligent design'.

2. Proto-Uralic in a global context

Although, consequently, Proto-Uralic remains a valid entity, there are many details about it that still call for explanation. Most of the unsolved problems revolve around the classic issues of dating and locating the Proto-Uralic speech community. These issues are usually approached by studying both internal and external evidence, that is, the Uralic comparative corpus, on the one hand, and the traces of contacts with other language families, on the other. For a more general understanding of Uralic, we may, however, also take a look at its position in a global context. There are between 6,000 and 7,000 languages spoken today, and they represent between 400 and 500 separate genetic lineages, or language families. Uralic, with some 30–40 separate languages, is slightly larger than an average-sized family. Judged by the number of speakers it is a relatively small entity, but in view of its territorial extension it is, in fact, one of the larger families in the world. Both the wide territorial extension of the Uralic languages and the small volume of the Uralic comparative corpus suggest that it is an 'old' family, that is, that it was formed a long time ago.

One of the great unsolved questions about human language is how the linguistic map of the world developed in pre-historical times. The best points of reference for understanding the situation would seem to be offered by those parts of the world that are still dominated, or until recently were dominated, by hunter-gatherer societies leading a relatively 'primitive' way of life. Such parts of the world include Australia, New Guinea, Siberia, and much of the Americas. A case in point is the island of New Guinea (including the western part under Indonesian administration), which today has a population of just about nine million people speaking more than a thousand separate languages. Although the size varies considerably, the average number of people speaking a single language varies between 8,000 and 9,000. This may be compared with the situation in Siberia (including the Russian Far East), where in the initial period of Russian rule (16th to 19th cc.) around 50 languages were spoken by speech communities that mainly ranged between 50 and 5,000 people in size (Dolgix 1960). It is reasonable to assume that pre-historical speech communities typically comprised of only a few hundred individuals each.

It also seems that the formation and expansion of many of the large language families of the world took place under circumstances specifically connected with technological and social innovations, as well as population growth. The most decisive factor behind, at least, many 'old' language families seems to have been the so-called Neolithic Revolution, which involved a rapid development of the methods of food production, especially agriculture and cattle breeding, and a subsequent population expansion and social stratification. The Neo-

lithic Revolution took place at different times in different parts of the world, but the important conclusion is that before this there may not have been sufficient grounds for the differentiation of individual languages into large language families. Although linguistic evolution itself must have taken place in the same way as in historical times, the Neolithic opened up a new line which has been shaping the language map of the world ever since.

It is particularly important to note that the formation of the present-day large language families has not necessarily involved massive population migrations, because languages have spread by way of diffusion just as often as by migration. Both migration and diffusion may have been triggered by a variety of internal and external, as well as positive and negative, push and pull factors, including natural calamities, cultural innovation, and population growth. The main process has in most cases been linguistic assimilation, or language shift, in which an original linguistic diversity of languages has been gradually lost in favour of an expansive family. One might think that the expansion of a language family into new areas would lead to a general decrease in the number of local languages, but this is not necessarily the case, because at the same time as a language family expands it also undergoes differentiation into new branches and sub-branches, which function as separate languages. In fact, it often happens that each 'original' local language is replaced by a new language from the expansive family. What is lost is genetic diversity, that is, the number of language families, while the number of separate languages may remain relatively stable.

It may be concluded that the world in pre-historic times may well have had as many languages as there are today, possibly even more. Assuming that the total human population prior to the Neolithic reached, say, some millions, which is a reasonable even if inexact and uncertain estimate, and assuming further that the average size of a speech community was around 500 people, there could well have been around 10,000 separate languages spoken in the world at any given time before the expansion of today's language families. It is more difficult to estimate how many lineages these pre-historic languages represented, but the number must have been larger than today, since the conditions for linguistic expansion were less favourable than later. All pre-historic languages had, in principle, an equal chance of becoming ancestors to lineages surviving up to the present day, but very few of them were successful, the main reason being that they were extinguished by the expansion of the extant language families. We might also say that all pre-historic languages, like most languages today, lived under a constant threat of extinction. In this sense, Proto-Uralic was an 'endangered' language until it started its expansion.

3. The areal position of Uralic

Like all languages today, Proto-Uralic must also have been a member of a language family. This family was not Uralic, however. This situation may be compared to Latin, which was a member of the Italic branch the Indo-European family, but which itself became the ancestor of the modern Romance group of languages. Quite probably, Proto-Uralic had some living relatives which represented the parallel branches of a protolanguage that became extinct in Pre-Proto-Uralic times. The other languages, if they existed, were not Uralic, but they may technically be identified as Para-Uralic, meaning that they represented lineages collaterally related to the lineage of Proto-Uralic. Of course, it is also possible that Proto-Uralic had no contemporary relatives. This would have meant that its lineage had 'never' undergone any branching, or, more plausibly, that any branches that may have existed had either become extinct or were so ancient that traces of an original genetic relationship had been lost before the Proto-Uralic period. In this context, it is good to remember that the linguistically detectable mutual relationship between two or more languages always represents only a section of the time scale. At the same time as differentiation goes on at the shallow end, the traces of the relationship are being erased at the deep end of the time scale.

Every now and then, the possibility of an external relationship between Uralic and other extant language families is raised. If such a relationship could be shown to have been a reality, the other language families would, from the Uralic point of view, represent surviving branches of Para-Uralic. The most prospective Para-Uralic entity would for many reasons appear to be Yukaghir (Yukaghiric), and the arguments in favour of a Uralo-Yukaghiric affinity (on which cf., e.g., Rédei 1999), cannot completely be dismissed. However, the Uralo-Yukaghiric comparative corpus does not correspond to the definition of a language family, in that it would contain a coherent, even if small, corpus of verifiable lexical cognates with more or less regular sound correspondences, complemented by a set of grammatical parallels. In fact, at closer inspection most of the Uralo-Yukaghiric comparisons turn out to be illusory. Even in an areal framework, Yukaghir seems to be an entity more closely linked to its Northeast Asian neighbours (Kamchukotic) than to Uralic. Therefore, if Yukaghir was once related to Uralic, the relationship would be so ancient that it can no longer be reliably detected. The same applies to the other long-range comparisons made between Uralic and other language families.

What is, however, an undeniable fact is that the Uralic languages belong to a single trans-Eurasian belt of agglutinative languages together with the so-called Altaic languages, including Turkic, Mongolic, Tungusic, Korean (Koreanic),

and Japanese (Japonic). In this case, typological parallelism is accompanied by areal adjacency, allowing us to speak of a distinct Ural-Altaic language area and language type. Characteristic features of the Ural-Altaic language type include a modifier-before-headword word order both in the sentence (SOV) and within the nominal phrase (GAN), suffixally marked agglutinative morphology both of the noun and the verb, as well as polysyllabic root structure with simple phonotactic patterns and no suprasegmental distinctions. Deviations from the prototypical Ural-Altaic language type occur in the individual branches and languages, especially in the west (Finnic, Saamic), north (Northern Samoyedic), and east (Koreanic, Japonic), but the basic typological orientation is nevertheless observable throughout the transcontinental belt. The internal uniformity of the Ural-Altaic complex is not annulled by the fact that the reconstructed protolanguages represent a slightly simplified picture due to the potentially distorting effect of the comparative method (Korhonen 1974). It is important to note, however, that the typological similarities are not accompanied by any significant amount of lexical cognates or even lookalikes, except for items documentably transmitted by way of borrowing.

A simplistic explanation of the situation would be to assert that the properties of the Ural-Altaic language type are so trivial and universally so common that their parallel occurrence in several adjacent language families is coincidental. This is certainly not the case, however, for the Ural-Altaic belt has clear areal boundaries which delimit the language type in relation to its neighbours both to the north (Yukaghiric, Kamchukotic) and to the south (Indo-European, Sino-Tibetan), and also to those in the extreme east (Ghilyak, Ainu). Along the margins of the Ural-Altaic belt we may also observe examples of gradual Altaicisation, as in the case of Northern Chinese (Hashimoto 1986), or also de-Altaicisation, as in the case of several Turkic and Mongolic languages of the Amdo Qinghai region (cf., e.g., Janhunen 2007). It has to be concluded, therefore, that the mutual typological parallels of the Ural-Altaic languages are due to actual areal contacts in the past. As far as the so-called Altaic languages are concerned, similarities need not date further back in time than a couple of millennia, when the homelands of the language families concerned were located in a compact area in southern Manchuria (Janhunen 1996: 216, map 6). For Uralic, the issue is more complicated, since this language family seems to be chronologically deeper and the question concerning its homeland has not been solved.

It has to be pointed out that there are also areal and typological parallels that link Uralic and its non-Ural-Altaic neighbours, especially Indo-European. Most importantly, both Uralic and Indo-European, together with Turkic, Mongolic, Tungusic, Yukaghiric, Amuric (Ghilyak), and Kamchukotic, belong to the so-called Mitian languages (on the term, cf. Bengtson 2008: 242, 250), in which

the roots of personal pronouns (in many languages also used as personal markers) contain a labial (*m*/*b*) in the first person and a dental (*t*/*s*) in the second person. In spite of persistent attempts, most recently by Johanna Nichols (2001), these pronominal similarities have never been explained satisfactorily, and they might well have either a genetic or an areal background, possibly even both (cf. also Bancel & de l'Etang 2008). Other cross-family material similarities are less convincing, though some of them, such as the recurrent accusative (**m*) and plural (**t*) markers, present also in Uralic, might be due to non-accidental factors. Irrespective of what the explanation is, both Indo-Uralic and Ural-Altaic are definitely relevant subjects for further study (cf. also Janhunen 2001; Kortlandt 2006, 2008). They do not have any immediate relevance to the debate on Uralic as a language family, however.

4. The structure of the Uralic family-tree

For the absolute dating of Proto-Uralic, the size and type of the comparative corpus are the most important tools. A look at the general picture of the language family also immediately suggests that branching and sub-branching has taken place at several different chronological levels. We may perhaps say that these levels represent three major horizons, which are relatively easy to distinguish. The three horizons may be identified as pre-Iron Age, Iron Age, and post-Iron Age, respectively. The uppermost, or post-Iron Age, horizon may be dated largely to the historical period (starting less than 1000 years ago). Linguistically this corresponds to the dialectal division of the modern Uralic languages. The degree of dialectal differentiation in the individual languages varies considerably, ranging from very shallow, as in the case of Komi Zyryan, to relatively deep, as in the case of the three Western Siberian languages Mansi, Khanty, and Selkup. At the deep end of this horizon there are cases which are on the verge of having become separate languages, such as Tundra Nenets vs. Forest Nenets, and Tundra Enets vs. Forest Enets vs. Yurats.

The medium, or Iron Age, horizon (roughly, between 1000 and 2000 years ago) represents the time period during which most of the individual branches of Uralic underwent differentiation into two or more relatively closely related, but distinct, languages. Again, the degree of internal differentiation varies somewhat, ranging from relatively shallow, as in the case of Permic (3 languages, which started to differentiate perhaps slightly less than 1000 years ago), to relatively deep, as in the case of Samoyedic (6–9 languages, which started to differentiate perhaps slightly more than 2000 years ago). Saamic (10 known languages), Finnic (8–10 languages), and Mordvinic (2 languages) would seem

to represent rather typical intermediate cases (with a differentiation history of perhaps about 1500 years). Typically, the size of the shared vocabulary within each branch differentiated in the Iron Age horizon varies between 800 and 1500 items, as has been shown for Samoyedic (Janhunen 1977) and Saamic (Lehtiranta 1989). The picture is, of course, slightly blurred by secondary contacts within the branches. Also, comparative work is in some cases made difficult by the lack of documentation and early extinction of a number of crucial languages (former Forest Saami, Sayan Samoyedic).

The deepest, or pre-Iron Age, horizon is the most difficult to assess. Those who advocate the 'comb' or 'rake' model would say that Proto-Uralic was more or less immediately divided into the synchronically attested major branches, ranging geographically from Samoyedic in the east to Saamic in the west. The principal problem with this model is that it presupposes an extremely sudden and explosive break-up of Proto-Uralic along a rather narrow east-west trajectory extending from Siberia to the Baltic Sea. This is equal to propagating a very broad homeland for Proto-Uralic, for, technically, the homeland would have comprised the whole area where Proto-Uralic would still have been spoken as a uniform language before the individual branches started differentiating, which would have happened only after the initial explosion. Such a high speed of expansion is, however, unlikely. It is more natural to assume that the protolanguage spread rather slowly, which would have meant that it differentiated at the same rate as it spread to new areas.

The break-up of Proto-Uralic may be compared with that of other 'old' language families, notably Indo-European. For Indo-European it is normally assumed that the protolanguage was dissolved by a simultaneous formation of several parallel descendant branches (cf., e.g., Anthony 1995: 557, fig. 1). In this case, the possibility of a non-binary division is supported by the fact that the break-up seems to have taken place in a radial manner, with the different primary branches advancing in different directions from the original core area. The diffusion of Proto-Uralic, by contrast, seems to have taken place in a linear manner, with a gradual and repetitive advance in one direction. The difference may not be so radical, however, for Indo-Europeanists have always looked for signs of a chronological hierarchy between the branches, and several actual or potential groupings have been discovered, including Balto-Slavonic and Italo-Celtic. Even more substantially, there are serious reasons to assume that the division between Hittite and the rest of Indo-European (proper) is more fundamental than any other branching within the family, resembling the division of Uralic into Samoyedic and Finno-Ugric (proper).

In any case, an unbiased look at the Uralic comparative corpus would seem to reveal a rather systematically westward-branching family-tree, with the divi-

sion between Samoyedic and Finno-Ugric lying at the foot. The basic dichotomy of the language family is particularly difficult to refute (cf. also Michalove 2002), a situation that has not been altered by fresh additions to the corpus of Finno-Ugric-Samoyedic lexical comparisons (Aikio 2002, 2006). There is also motivation to postulate a succession of several lower-level protolanguages, which may be termed Finno-Ugric, Finno-Khantic, Finno-Permic, Finno-Volgaic, Finno-Saamic, and Finno-Mordvinic. The entities that were separated from these protolanguages are Mansic (Mansi and Hungarian), Khantic, Permic, Mariic, Saamic, and Mordvinic, respectively (cf. the table below). It goes without saying that there are many details in this system that may require revision. For instance, the status of Khantic vs. Mansic remains controversial, and it is still too early to completely reject the possibility of a common 'Ugric' protolanguage for all these entities (cf. Honti 1998). Also, the mutual ordering of the three westernmost branches, Finnic, Saamic, and Mordvinic, is open to alternative interpretations. Even so, the basic structure of the family-tree seems to be solid.

Pre-Uralic → Para-Uralic
↘ Uralic → Samoyedic
↘ Finno-Ugric → Mansic
↘ Finno-Khantic → Khantic
↘ Finno-Permic → Permic
↘ Finno-Volgaic → Mariic
↘ Finno-Saamic → Saamic
↘ Finno-Mordvinic → Mordvinic
↘ Finnic & Para-Finnic

Table. The organisation of the Uralic family tree: the binary alternative.

5. The dating of Proto-Uralic

There should, consequently, be no doubt that the age of Proto-Uralic is equal to the time depth of the division between Samoyedic and Finno-Ugric. An exact dating of this division will never be possible, but an approximate dating can be, and has been, attempted on the basis of several circumstances. One way is to proceed by trying to relate Proto-Uralic and the subsequent branchings to archaeological cultures. This is the method chosen by Christian Carpelan and Asko Parpola (2001), and it is potentially very informative, since archaeological cultures can not only be dated but also located. Unfortunately, for both conceptual and definitional reasons, the entire framework of interpreting archaeological cultures in terms of linguistic identities is on a shaky ground. It is

well known from empirical evidence that cultural boundaries do not necessarily correspond to linguistic boundaries. Also, archaeological cultures are typically defined on the basis of a limited number of markers, which means that they do not necessarily correspond to actual cultural spheres.

The archaeological approach to linguistic prehistory involves also territorial and chronological problems. The farther back in time we go, the larger the areas covered by archaeological cultures tend to be. Simplistically thinking, this should mean that linguistic areas were larger in the past than they are today (or in the period with historical documentation). This is why Carpelan and Parpola assume, for instance, that the Neolithic Comb Ceramic culture (5th to 4th millennia BZ[1]) in north-western Eurasia, including Finland, was Proto-Uralic speaking (cf. also Carpelan 2000: 15–16, 19–20). Unfortunately, this contradicts the fact that the same territory has been historically occupied by a multitude of different speech communities, mainly Uralic, but also non-Uralic. Quite obviously, the Comb Ceramic culture comprised at least a comparable variety of languages, and it is impossible to tell whether any of these languages was Uralic and, if so, in what part of the large territorial complex it would have been spoken. Most importantly, however, archaeological cultures tend to be much too early to correspond to what is otherwise known of the chronology of the linguistic map. There is reason to agree with Petri Kallio (2006), who maintains that most datings in Uralic studies are traditionally too deep.

It is, consequently, reasonable to relate linguistic datings to linguistic facts in the first place. Interestingly, however, as far as Proto-Uralic is concerned, linguistic facts would seem to favour a very early dating. It cannot be an accident that the vocabulary shared by the two primary branches of Uralic is not only small in size but also qualitatively indicative of a rather 'primitive' cultural stage (on this, cf., e.g., Kaisa Häkkinen 1998). In fact, the cultural vocabulary we know from Proto-Uralic does not even represent a 'Neolithic' but, rather, a Mesolithic stage of development (as was once pointed out to the present author in an oral comment by Terho Itkonen). Thus, the Proto-Uralic lexicon comprises several words for typical pre-Neolithic cultural innovations, including 'bow' (**yïngsi*) and 'arrow' (**nyïxli*), 'ski/s' (**suksi*), and 'row' (**suxi*-). Other items connected with the subsistence methods of a hunter-gatherer community include the words for 'fish' (**kala*), 'egg' (**muna*), 'nest' (**pesa*), as well as 'hunt' (**nyoxi*-), while, with the single exception of a word for 'tame' (**ïnyi*, implying perhaps the keeping of dogs), there is no indication of any kind of agriculture or cattle breeding, nor of any major social innovations. The often-quoted item for 'metal' (**wäcka*), also discussed by Kallio (2006: 6–8), is not informative in this context due to the

[1] BZ = Before Zero ('Before Common Era'); AZ = After Zero ('Common Era').

possibility of internal and external loan contacts during the post-protolanguage period.

At this point, it is important to reject one mistaken argument that has been presented against the conventional Uralic family-tree. According to this, the small size of the lexical corpus shared by Samoyedic and Finno-Ugric is irrelevant for diachronic conclusions, since, it is claimed, Samoyedic may have undergone a process of 'relexification', in which it would have lost much of its original Uralic vocabulary (cf., e.g., Salminen 2002: 52). This argument is based on the fact that Samoyedic alone, when compared to the other branches of Uralic, lacks many otherwise widespread basic words, including, for instance, the Finno-Ugric words for 'hand' (**käti*) and 'head' (**päxi*). However, actual examples of 'relexified' languages, or 'creoles', suggest that a massive replacement of basic vocabulary is always accompanied by grammatical restructuring, normally simplification. This is not the case with Samoyedic, which, by contrast, is a conspicuously conservative branch of Uralic. As a matter of fact, Samoyedic is in some morphological and phonological respects so similar to the likewise conservative Finnic branch in the west that this has misled Ago Künnap (most recently, 2008) to postulate secondary contacts or 'language shifts' between the two extremities of the family.

The fact that the Finno-Ugric side may also have been innovative is shown by lexical items such as those for 'hare' (Samoyedic **nyoxma* vs. Finno-Ugric **nyoxma-la*, ultimately from **nyoxi-* 'hunt') and 'feather' (Samoyedic **tuxli* vs. Finno-Ugric **tuxl-ka*), in which the derived stem is present only in Finno-Ugric. Nowhere is the Uralic family-tree so obvious, however, as in the numeral system (Janhunen 2000: 60–61). For Proto-Uralic, only the items '2' (**kekta*) and '5' > '10' (**witi*) can be reconstructed. The system was expanded separately in Samoyedic and Finno-Ugric. In the latter branch, the items '3' (**kormi*) and '4' (**nyelya*), as well as '6' (**kuti*), are shared by all sub-branches, while the alternative shape of the item '3' (> **kolmi*) has a Finno-Khantic distribution. The item '2' underwent a restructuring in the Finno-Permic sub-branch (> **kakta*). The remaining items were created separately in the lower-level sub-branches, including Finno-Volgaic ('8', '9') and Finno-Mordvinic ('10'). It is important to stress that the construction of the numeral system may quite well imply a simultaneous evolution of the counting system, which is a culturally conditioned feature. Even the Indo-European numeral system was still in the making at the deepest (Indo-Hittite) level of the protolanguage (Bomhard 2008).

There is, consequently, a lexical basis for postulating a westward-branching hierarchy for Uralic. This hierarchy is also evident in the phonology and morphology (not elaborated on here, but partially summarised in Sammallahti 1988). Although the time between the branchings must have varied, it is not un-

reasonable to assume that variation was not particularly great, especially since the geographical distances between the branches are more or less equal. Let us, therefore, tentatively assume that each branching took an average time of, say, 500 years to be completed. Starting from the west and assuming that Proto-Finnic is located at a depth of 1500 before the present, we then get a succession of increasingly deep datings for the earlier protolanguages: 2000 for Proto-Finno-Mordvinic, 2500 for Proto-Finno-Saamic, 3000 for Proto-Finno-Volgaic, 3500 for Proto-Finno-Permic, 4000 for Proto-Finno-Khantic, 4500 for Proto-Finno-Ugric, and 5000 for Proto-Uralic. It happens that this Proto-Uralic dating (3000 BZ) is surprisingly close to some of the datings established by other methods, irrespective of whether they have been correct or not. It is, however, considerably shallower than the wildest archaeological datings proposed.

6. The physical type of Uralic speakers

It may be concluded that there is no basis for the assumption that Samoyedic would be any less Uralic in its lexical composition than the Finno-Ugric languages. On neither side is there any sign of massive 'relexification', and even if there were, there would be no way of telling on which side, Samoyedic or Finno-Ugric, the presumed 'relexification' would have taken place. The situation is, incidentally, very similar when the physical features of Uralic speakers are considered. Uralic speakers, in general, represent a continuum in which western, or 'European', features are dominant at the Baltic Sea (Finnic), while eastern, or 'Asian', features are strongest in the east (Samoyedic). Since Samoyedic speakers constitute a numerical minority of all Uralic speakers, it would be easy to argue that they have 'changed' their genes, that is, that they actually represent a physically different population, or a group of populations that once secondarily adopted a Uralic language. There is, however, no way to show that this was the case.

As a matter of fact, the physical continuum among Uralic speakers is relatively smooth, meaning that speakers of the Uralic languages are congruous with the trans-Eurasian continuum of populations occupying the region between Fennoscandia (the Baltic region) and eastern Siberia (the Baikal region). This is so irrespective of whether we are looking at those features described by classical anthropology or at the variation studied by modern molecular genetics. The question as to what the 'original' physical type, or genetic composition, of any given protolanguage-level speech community was cannot be easily answered (for a critical survey, cf., e.g., Häkkinen 2007). Population genetics tells us what the distribution of specific genetic markers is on the map, but for the time being,

at least, it does not give us reliable tools to specify the direction of movement and absolute age of the underlying gene flows. Still less does it tell us what the correlation between gene flows and languages might be. For the absolute dating of actual migrations in the past, palaeoanthropology and archaeology are potentially more informative, but even they remain helpless when it comes to the identification of linguistic correlations.

In the few cases where we have a relatively sharp boundary in the distribution of physical types among Uralic speakers we have to assume recent migrations and/or language shifts. The best known example is offered by the modern Saami, who, without a doubt, represent an originally non-Uralic-speaking population in northern Fennoscandia. The expansion of Proto-Saamic to the physical ancestors of the modern Saami is likely to have taken place very late on the time scale, most probably only during the last millennium. In this process, Proto-Saamic was divided into the modern Saamic languages, perhaps in a rough correlation with the earlier linguistic map of the Saami area. Another example of a sharp physical boundary is that between the western Tundra Nenets and their Uralic-speaking neighbours, most importantly the northern Komi. The western Tundra Nenets are clearly 'Asian' in their physical type, while the Komi are basically 'European'. In this case, also, the contact zone between the two physical types is very recent and is based on migrations which have brought the Asian type (probably together with the Nenets language) to the west and the European type (together with the Komi language) to the north.

As the physical difference between European and Asian population types nevertheless seems to reflect an old dichotomy of human evolution in Eurasia, it is possible that the Uralic language family, at some time, spread across a relatively sharp 'racial' boundary. It is even likely that the original Proto-Uralic population was 'racially' coherent, meaning that its dominant physical features may have been either 'European' or 'Asian'. There is, however, no easy way of determining which of the two physical types was 'originally' connected with the Uralic language family. This is an issue that is more closely connected with the direction of expansion of the language family than with the numerical proportions of the physical types among the modern Uralic-speaking populations. The growth of speech communities depends on a variety of extra-linguistic factors, including cultural and political circumstances. If only the numerical proportions are considered, we would have to conclude that the original Uralic 'type' was close to that of the modern Hungarians. We know, however, that Hungarian speakers represent a local complex of Central European physical types that are also present among the speakers of neighbouring languages, including Romanian, Serbian, and Slovak. There is hardly any Uralic-speaking population that would be farther from the Proto-Uralic physical type than the Hungarians.

To take a similar example from another language family: Turkish is today the 'largest' Turkic language in terms of the number of speakers. This could be mistaken to imply that the modern Turks of Turkey represent the 'original' physical type of Turkic speakers. Nothing could be less true, however. In reality, Turkic was until the latter half of the first millennium (AZ) spoken by a predominantly 'Asian' population, which inhabited the region today known as Mongolia. Due to demographic, cultural, and political circumstances the Turkic languages spread across Central Eurasia as far as Turkey, whereas in Mongolia itself they were marginalised by the expansion of the Mongolic language family. We might say that the modern Mongols of Mongolia are more or less direct descendants of the ancient Turks, while the modern Turks of Turkey represent a more or less direct continuation of the old local population of Anatolia, which historically has spoken a succession of non-Turkic languages, ranging from Hattic and Hittite to Armenian and Byzantine Greek.

7. Locating the Uralic homeland

It follows from the preceding that Proto-Uralic must have been a language spoken by a relatively small and geographically strictly localised speech community whose members very probably represented a coherent physical type. This type may have been either Asian or European, depending on where the speech community was located and when it was dissolved. The cultural stage reflected by the lexicon of Proto-Uralic speakers seems to have been 'Mesolithic', which means that it may have been a question of a relatively 'primitive' and most probably non-settled hunter-gatherer community comprising no more than some thousands of people, at most. However, once Proto-Uralic had started to expand, the process seems to have become cumulative, with ever new branches and subbranches being generated until the modern family-tree had become complete.

Much of the territorial expansion of the Uralic language family must have taken place by way of language shift, in which Uralic speech spread to populations that had earlier spoken other languages. Traces of the original non-Uralic linguistic diversity can be discerned in the contact-induced structural and lexical properties of the individual Uralic branches and languages. In fact, most modern Uralic languages are likely to have been locally preceded by one non-Uralic substrate language or more. In some cases, these substrate languages can be positively identified, often as Indo-European, while in others they remain unknown. In the latter case, we may only generically speak of a diffuse 'Palaeo-European' (Saarikivi 2004; cf. also Aikio 2004) or, on the Asian side, 'Palaeo-Asiatic' substrate influence. Irrespective of this, the secondary properties caused by sub-

strates and other types of language contact should not be confused with the primary properties connected with the Uralic lineage. For purposes of genetic linguistics, it is only the lineage that counts, since it represents the most ancient and, in principle, invariant core of the language.

As has been pointed out above, it is a virtually hopeless task to try to locate the Uralic homeland with the help of non-linguistic disciplines, including archaeology, palaeoanthropology, and population genetics. The available linguistic tools, on the other hand, are also problematic, since linguistic material is, in principle, independent of the geographical context. Potentially the most informative method for locating the homeland would seem to be offered by linguistic palaeontology. It has long been argued that, especially, dendronyms require the Uralic homeland to be placed rather far in the east, possibly on the Siberian side of the Urals (Hajdú 1969: 257–258), the crucial argument being provided by the Proto-Uralic item for 'cedar' (**siksi*). It may be recalled that the typological orientation of the Uralic languages in the Ural-Altaic areal context also favours the assumption of an 'eastern' homeland. Linguistic arguments in favour of a 'western' homeland, located possibly as far west as the Baltic region, are mainly based on alleged protolanguage-level lexical parallels between Uralic and Indo-European (Koivulehto 2001 and elsewhere). Unfortunately, the parallels in question are highly controversial (cf., e.g., Helimski 2001) and can hardly serve as a basis for further conclusions, especially as the question concerning the Indo-European homeland also remains unsettled.

The most uncontroversial information on the pre-historical location and movements of Uralic on the map is, however, provided by the internal taxonomy of the language family. The very fact that the branchings of Uralic seem to become chronologically shallower the farther west we proceed suggests that the main direction of expansion has been systematically from east to west. In other words, the Uralic language family seems to have been formed as a more or less binarily organised hierarchical chain, in which a new branch has always been formed on the western side of the previous ancestral branch. By the classic principle of linguistic geography this also has to mean that the deepest boundary within the language family must correspond to the original location of the first break-up, that is, the linguistic homeland. This criterion places the break-up of Proto-Uralic in the region which historically forms the boundary between Samoyedic and its immediate Finno-Ugric neighbours (Khantic and Mansic). The region in question is the borderline between the Ob and Yenisei drainage areas in Siberia, and until the contrary is shown, it qualifies as the most likely candidate for the Uralic homeland.

To be exact, we do not know whether the first break-up of Proto-Uralic also involved a westward-branching division, for it is also possible that Proto-

Samoyedic moved eastwards from the homeland, while Proto-Finno-Ugric remained in the original location until it entered into its westward-branching history. It is, however, noteworthy that the subsequent geographical centre of the Samoyedic languages is formed by the Minusinsk basin on the Upper Yenisei, a compact region with an exceptionally well-documented sequence of archaeological cultures, extending from the Eneolithic Afanasievo culture (3500–2500 BZ) through the Bronze and Iron Age Okunevo (2500–2000 BZ), Andronovo (2000–1500 BZ), Karasuk (1500–800 BZ), Tagar (800–100 BZ), and Tashtyk (BZ 100–400 AZ) cultures up to the historical Yenisei Kirghiz (from 400 AZ), Mongols (from 1300 AZ), and Russians (from 1700 AZ). This is a much more specific record than anything established so far in the sparsely inhabited forest zone between the Volga and the Baltic Sea, the traditional candidate for a 'western' homeland of Uralic.

Without going into the question concerning the possibility of a Proto-Uralic presence in the Minusinsk basin, it is relatively safe to follow the ethnolinguistic history of the region backwards to the arrival of Turkic (later Yenisei Turkic), which ended the Tashtyk (or Hunnish) period in the region. The historical distribution of the local ethnolinguistic groups strongly suggests that the dominant language in the Minusinsk basin before Turkic, that is, the language of the Tashtyk Culture, was Yeniseic (Proto-Yeniseic), while the dominant language before Yeniseic, that is, the language of the Tagar Culture, must have been Samoyedic (Proto-Samoyedic). Much speculation has been presented concerning the possible linguistic identities of the Karasuk, Andronovo, Okunevo, and Afanasievo Cultures, but nothing certain can be said. Even so, the Indo-European elements in Samoyedic suggest that some early eastern form of Indo-European (Proto-Tocharian?) may have been present in the region either before Samoyedic or in parallel with it (Janhunen 1983).

8. Uralic in time and place

Uralic is one of the relatively few (probably less than 100) reliably established 'old' language families of the world. The small size of the Uralic comparative corpus, especially as far as the lexicon is concerned, suggests a relatively early dating for the protolanguage. The cultural stage reflected by the shared lexicon is even more indicative of a very early first break-up, possibly datable to the Mesolithic level. The structure of the language family, as well as the available palaeolinguistic evidence, suggests that its original homeland was located relatively far to the north, probably within the boreal zone or, at least adjacent to it, and relatively far to the east, probably on the Asiatic side of the Urals. The

protolanguage-level speech community may or may not have been dominated by Asian physical features. The subsequent expansion of the language family took place mainly by way of linguistic assimilation, in which process a number of local populations with different cultural backgrounds and physical heritage gradually became Uralic speaking.

Although the original number of Proto-Uralic speakers must have been very small (hardly more than a few thousand, perhaps even less), the fact that the language started to expand and became the source of cumulative branchings suggests that there was an initial and recurrent edge that favoured linguistic expansion. The crucial question is what this edge could have been. It was certainly not demographic (bigger population), nor can it have been military (stronger striking power). Very probably, it was not one of material culture (more advanced technology), social structure (more effective organisation), or spiritual heritage (more attractive traditions), either. As one possibility, Carpelan and Parpola (2001: 109–110) have pointed out the significance of trading, especially in the context of the so-called Bronze Age Seima-Turbino 'transcultural phenomenon' (1800–1500 BZ), though in their model this becomes relevant only in the Post-Proto-Uralic period. However, the edge may also simply have been a strategic position at the boundary between the forest and steppe belts, or also in the vicinity of the southern end of the Urals, a region which became one of the first sources (perhaps the very first source) of metal age cultures in Eurasia.

Although the development of metal age technologies, as well as the rise of agriculture and cattle breeding in Eurasia took place in linguistic environments other than Uralic, Uralic speakers were never too far away from the centres of cultural innovation, and their successful linguistic expansion in the northern forest belt may well have been been related to their role as satellites of their southern neighbours, many of whom spoke Indo-European languages. Typically, most of the interaction between the two language families involved the influx of Indo-European elements into Uralic, rather than vice versa. The material suggests that contacts were initiated only in the Post-Proto-Uralic period and grew stronger with time. It is no accident that the westernmost branches of Uralic, that is, Finnic and Saamic, exhibit lexical traces of an almost complete succession of Indo-European donor languages, ranging from Pre-Iranian through Iranian to Baltic, Germanic, and Slavonic. Certainly, in spite of claims to the contrary (Koivulehto 1983), none of the earlier layers of loanwords was received in the current location of the Finnic and Saamic languages. Rather, the distribution and diachronic properties of the borrowings reflect the geographical movement of the ancestral forms of Finnic and Saamic across the forest belt between the Urals and the Baltic Sea.

Considering the, presumably, very small size of many local populations and speech communities, it is not unlikely that there were also cases of language shift from Indo-European to Uralic. As the north-western branches of Indo-European, especially Baltic and Germanic, continued their expansion towards the west, their last remnants in the east may well have been absorbed by their Uralic partners and satellites, especially Finnic and Saamic. Such a development is especially likely to have taken place in the Volga-Ilmen-Ladoga region, which must have lain on the trajectory of Indo-European expansion, but which ultimately came to form the homeland of Finnic and Saamic. Much later, and under somewhat different circumstances, Hungarian (of the Mansic branch) was absorbed into the steppe under the impact of Turkic and transplanted into Pannonia, where it replaced a number of earlier Indo-European languages. By this time, Turkic itself had already replaced Indo-European in the Central Eurasian steppe zone.

With the exception of Hungarian, the east-to-west geographical sequence of the branches of Uralic, extending from the Baikal region to the Baltic Sea was complete by the Iron Age (c. 2000 years ago). From this time on, the principal direction of expansion of the Uralic languages has been from south to north. In this process, most of the Uralic branches, notably Finnic, Saamic, Permic, Khantic, Mansic, and Samoyedic, spread from their individual homeland regions northwards towards the Arctic coast, which they reached perhaps a millennium ago, or later. In the case of Saamic and Samoyedic, the expansion continued horizontally along the tundra belt, again mainly in an east-to-west direction. The northern expansion of Uralic caused the extinction of an unknown number of earlier languages, a process which may have ended only a few centuries ago (Helimski 2000). Development has been rapid, however, and some of the expansive Uralic languages have themselves been extinguished by Russian, which forged its way to the Arctic coast in the immediate footsteps of Uralic (especially Finnic).

Chronologically, Uralic remains ambiguous. On the one hand, it is obvious that the modern locations of many present-day Uralic languages are very recent. The south-to-north dimension of the Uralic language belt has a chronological depth of less than two millennia, which is also the maximum age of the internal differentiation of most of the individual branches of Uralic. On the other hand, the geographical length of the east-to-west chain and its systematically westward-branching structure, as well as the Mesolithic cultural level reflected by the Proto-Uralic lexical corpus, suggest a very early dating for the language family as a whole. The external evidence provided by the earliest layers of Indo-European loanwords (considering only uncontroversial data) also suggests that the first split in Proto-Uralic took place very early, and in any case before contacts with Indo-European were initiated. However, the Mesolithic, like the Neo-

lithic, can have widely different absolute datings in different parts of the world. Assuming that the Proto-Uralic speakers were hunter-gatherers of the boreal zone somewhere in Central Eurasia, who, due to their strategic position were drawn into a process of linguistic expansion, it is possible to moderate the datings and place Proto-Uralic at a chronological level perhaps not so much earlier than the earliest stage of Proto-Indo-European (Indo-Hittite).

Acknowledgements: Among the many colleagues who have contributed, directly or indirectly, to the formation of the ideas contained in the present report, the author wishes to recognise his gratitude, in particular, to the late Eugene Helimski (1950–2007). At the same time, the author wishes to stress that many of the interpretations presented above are no longer as controversial or 'counter-revolutionary' as they would have been only a few years ago. After a couple of decades of intensive, and certainly useful, disputes and revisions of the paradigm, several crucial issues concerning time and place in Uralic studies are now approaching a new level of consensus, especially among younger scholars, as summarised by Jaakko Häkkinen (2009), with whose conclusions the present author largely agrees.

References

Abondolo, Daniel 1998: Introduction. – Daniel Abondolo (ed.), *The Uralic Languages.* Routledge Language Family Descriptions. London – New York: Routledge. 1–42.

Aikio, Ante 2002: New and old Samoyed etymologies. – *Finnisch-Ugrische Forschungen* 57: 9–57.

Aikio, Ante 2004: An essay on substrate studies and the origin of Saami. – Irma Hyvärinen, Petri Kallio & Jarmo Korhonen (eds.), *Etymologie, Entlehnungen und Entwicklungen. Festschrift für Jorma Koivulehto zum 70. Geburtstag.* Mémoires de la Société Néophilologique de Helsinki 63. Helsinki. 5–34.

Aikio, Ante 2006: New and old Samoyed etymologies (Part 2). – *Finnisch-Ugrische Forschungen* 59: 9–34.

Anthony, David W. 1995: Horse, wagon and chariot: Indo-European languages and archaeology. – *Antiquity* 69: 554–565.

Bancel, J. Pierre & l'Etang, Alain Matthey de 2008: The millenial persistence of Indo-European and Eurasiatic pronouns and the origin of nominals. – John D. Bengtson (ed.), 439–464.

Bengtson, John D. 2008: The languages of northern Eurasia: Inference to the best explanation. – John D. Bengtson (ed.), 241–262.

Bengtson, John D. (ed.) 2008: *In Hot Pursuit of Language in Prehistory: Essays in the four fields of anthropology. In honor of Harold Crane Fleming.* Amsterdam – Philadelphia: John Benjamins Publishing Company.

Bomhard, Allan 2008: Some thoughts on the Proto-Indo-European cardinal numerals. – John D. Bengtson (ed.), 213–221.

Carpelan, Christian 2000: Essay on archaeology and languages in the Western end of the Uralic zone. – *Congressus Nonus Internationalis Fenno-Ugristarum. Pars I.* Tartu. 7–38.

Carpelan, Christian & Parpola, Asko 2001: Emergence, contacts and dispersal of Proto-Indo-European, Proto-Uralic and Proto-Aryan in archaeological perspective. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), 55–150.

Carpelan, Christian & Asko Parpola & Petteri Koskikallio (eds.) 2001: *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological considerations.* Mémoires de la Société Finno-Ougrienne 242. Helsinki.

Dolgix 1960 = Долгих, Б. О.: *Родовой и племенной состав народов Сибири в XVII в.* Труды Института этнографии 55. Москва.

Grünthal, Riho & Johanna Laakso (eds.) 1998: *Oekeeta asijoo: Commentationes Fenno-Ugricae in honorem Seppo Suhonen sexagenarii.* Mémoires de la Société Finno-Ougrienne 228. Helsinki.

Hajdú, Péter 1969: Finnougrische Urheimatforschung. – *Ural-Altaische Jahrbücher* 41: 252–264.

Häkkinen, Jaakko 2006: Uralilaisen kantakielen tutkiminen. – *Tieteessä tapahtuu* 2006: 1: 52–58.

Häkkinen, Jaakko 2007: Jälleen suomalaisten juurilla. – *Tieteessä tapahtuu* 2007: 6: 53–64.

Häkkinen, Jaakko 2009: Kantauralin ajoitus ja paikannus: perustelut puntarissa. – *Journal de la Société Finno-Ougrienne* 92: 9–56.

Häkkinen, Kaisa 1984: Wäre es schon an der Zeit, den Stammbaum zu fällen? Theorien über die gegenseitigen Verwandtschaftsbeziehungen der finnisch-ugrischen Sprachen. – *Ural-Altaische Jahrbücher* 32: 1–4.

Häkkinen, Kaisa 1998: Uralilainen muinaiskulttuuri sanahistorian valossa. – Riho Grünthal & Johanna Laakso (eds.), 188–194.

Hashimoto Mantaro J. 1986: The Altaicization of Northern Chinese. – John McCoy & Timothy Light (eds.), *Contributions to Sino-Tibetan Studies.* Cornell Linguistic Contributions 5. Leiden: Brill. 76–97.

Helimski, Eugene 2000: Before the Uralians came: White spots on the historical language map of Northern Eurasia and the Uralic languages. – *Congressus Nonus Internationalis Fenno-Ugristarum. Pars II.* Tartu. 72–73.

Helimski, Eugene 2001: Early Indo-Uralic linguistic relationships. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), 187–205.

Honti, László 1998: Ugrilainen kantakieli – erheellinen vai reaalinen hypoteesi? – Riho Grünthal & Johanna Laakso (eds.), 176–187.

Janhunen, Juha 1977: *Samojedischer Wortschatz: Gemeinsamojedische Etymologien.* Castrenianumin toimitteita 17. Helsinki.

Janhunen, Juha 1983: On early Indo-European-Samoyed contacts. – Juha Janhunen, Anneli Peräniitty & Seppo Suhonen (eds.), 115–127.

Janhunen, Juha 1996: *Manchuria: An Ethnic History.* Mémoires de la Société Finno-Ougrienne 222. Helsinki.

Janhunen, Juha 2000: Reconstructing Pre-Proto-Uralic: Spanning the millennia of linguistic evolution. – *Congressus Nonus Internationalis Fenno-Ugristarum. Pars I.* Tartu. 59–76.

Janhunen, Juha 2001: Indo-Uralic and Ural-Altaic: On the diachronic implications of areal typology. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), 207–220.

Janhunen, Juha 2007: Typological interaction in the Qinghai Linguistic Complex. – *Studia Orientalia* 101. Helsinki. 85–103.

Janhunen, Juha (forthcoming): Language and the search for identity: M. A. Castrén and the rise of the Uralic concept, 1800–1880. – Michael Branch (ed.), [Proceedings of the conference:] *Research and Identity: Non-Russian Peoples in the Russian Empire, 1800–1855* [Kouvola, 2006]. [To appear in the series:] Studia Fennica. Helsinki.

Janhunen, Juha & Anneli Peräniitty & Seppo Suhonen (eds.) 1983: *Symposium Saeculare Societatis Fenno-Ugricae.* Mémoires de la Société Finno-Ougrienne 185. Helsinki.

Kallio, Petri 2006: Suomen kantakielten absoluuttista kronologiaa. – *Virittäjä* 110: 2–25.

Koivulehto, Jorma 1983: Seit wann leben die Urfinnen im Ostseeraum? Zur relativen und absoluten Chronologie der alten idg. Lehnwortschichten im Ostseefinnischen. – Juha Janhunen & Anneli Peräniitty & Seppo Suhonen (eds.), 135–157.

Koivulehto, Jorma 2001: The earliest contacts between Indo-European and Uralic speakers in the light of lexical loans. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), 235–263.

Korhonen, Mikko 1974: Oliko suomalais-ugrilainen kantakieli agglutinoiva? – *Virittäjä* 78: 243–257.

Kortlandt, Fredrik 2006: Indo-Uralic and Altaic. – http://www.kortlandt.nl/publications/art216e.pdf.

Kortlandt, Fredrik 2008: Indo-Uralic and Altaic revisited. – http://www.orientalistik.uni-mainz.de/robbeets/verbalmorphv14/_Media/abs_kortland.pdf.

Künnap, Ago 2008: On the Finnic and Samoyedic accusative plural. – *Linguistica Uralica* 40: 34–40. [http://www.ceeol.com/aspx/getdocument.aspx?logid=5&id=e218d4f4-a7e6-4b32-8554-7f1adb7ba3ee]

Lehtiranta, Juhani 1989: *Yhteissaamelainen sanasto.* Mémoires de la Société Finno-Ougrienne 200. Helsinki.

Marcantonio, Angela 2002: *The Uralic Language Family: Facts, Myths and Statistics.* Publications of the Philological Society 35. Oxford: Blackwell Publishers.

Michalove, Peter A. 2002: The classification of the Uralic languages: Lexical evidence from Finno-Ugric. – *Finnisch-Ugrische Forschungen* 57: 58–67.

Nichols, Johanna 2001: Why "me" and "thee"? – Laurel J. Brinton (ed.), *Historical Linguistics 1999.* Amsterdam: John Benjamins. 253–276.

Pusztay, János 1995: *Diskussionsbeiträge zur Grundsprachenforschung (Beispiel: das Protouralische).* Veröffentlichungen der Societas Uralo-Altaica 43. Wiesbaden: Harrassowitz Verlag.

Rédei, Károly 1999: Zu den uralisch-jukagirischen Sprachkontakten. – *Finnisch-Ugrischen Forschungen* 55: 1–58.

Saarikivi, Janne 2004: Is there Palaeo-European substratum interference in western branches of Uralic? – *Journal de la Société Finno-Ougrienne* 90: 187–214.

Salminen, T[apani] 2002: Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. – *Лингвистический беспредел. Сборник статей к 70-летию А. И. Кузнецовой.* Москва: Издательство Московского университета. 44–55.

Sammallahti, Pekka 1988: Historical phonology of the Uralic languages with special reference to Samoyed, Ugric, and Permic. – Denis Sinor (ed.), *The Uralic Languages: Description, History and Foreign Influences.* Handbuch der Orientalistik VIII, 1. Leiden: E. J. Brill. 478–554.

Wiik, Kalevi 2004: Suomalaisten juuret. – *Tieteessä tapahtuu* 2004: 6: 28–34.