

**Tapani Salminen** (University of Helsinki)

## **Uralic**

*Loanwords:* At least in the Uralic wordlists, the number of loanwords appears larger than anticipated. For instance, the Finnish corpus contains no fewer than 17 Indo-European loanwords, thus one sixth of the total. It would be an interesting exercise to check what would happen to the ASJP tree if these were disregarded. As it now stands, the tree basically classifies lexicons instead of languages, and since languages unlike organisms follow Lamarckian rather than Darwinian principles of evolution in that acquired characteristics are inherited, this is indeed a drawback as long as phylogenies based on real developments of languages represent the desired goal. Could it be expected that two consecutive stages of the same language, say, Early Proto-Uralic before the influx of Indo-European loanwords and Late Proto-Uralic thereafter, would be classified quite differently through ASJP?

*Binary classification:* The traditional classification of Uralic supported by conservative specialists recognizes a large number of intermediate proto-languages and yields a more or less binary scheme for the family, not unlike typical ASJP tree structures. The more modern view of Uralic taxonomy, however, is based on the virtual absence of shared innovations except among the well-known groups of closely related languages, i.e. Saami, Finnic, Mordvin, Mari, Permian, Hungarian, Mansi, Khanty and Samoyed. Furthermore, whatever scanty support there may be for higher-level units, similar evidence could be found for establishing other nodes, contradictory to the traditional classification, and alternative ASJP trees attained via different algorithms interestingly reveal different results.

*Language splits:* It is true that language splits in general, insofar as they arise through the extinction of transitional dialects, must in reality have been binary, but it is another matter to detect the correct order of those splits when several of them have by all accounts followed each other in rapid succession, which would not be unusual for historically expansive language families such as Indo-European and Uralic. Many details in the ASJP tree seem to reflect such scenarios, and presumably slight changes in the data would often reverse the order of splits, which suggests the desirability of non-binary nodes instead. As for Uralic, we simply cannot tell where the first split occurred.

*Genetic versus areal features:* Within Uralic there have demonstrably been numerous contacts between neighbouring branches, and they seem to have the potential to skew the ASJP tree. For instance, if the distances in the ASJP tree are to be taken at face value, the relationships of Khanty and Mansi as opposed to Nenets and Selkup appear incoherent, as Nenets and Selkup are closely related members of the same branch while Khanty and Mansi are certainly not closely related, and many of their similarities are due to mutual contacts and shared substrate influences. Again, the different layers of primary and secondary similarities among languages merit serious consideration within the ASJP framework.

*Comments on the data and general conclusion:* Questions may be raised concerning, for instance, the role of function words, as they are typically very short elements that are often poorly translatable or tend to be expressed morphologically. Near-synonyms could also be easily dealt with by providing expressed contexts for the items. Furthermore, the ASJP code might be revised so that it would respect the phonological properties of languages to a larger extent. It is, however, safe to say that the ASJP project clearly has the potential to invigorate the subfield of language classification within historical linguistics, including Uralic studies, and if more critical attitudes towards classification methods are thereby successfully established, the project will indeed be highly commendable.