

Juha JANHUNEN (Helsinki)

ETYMOLOGY AND THE ROLE OF INTUITION IN HISTORICAL LINGUISTICS

I was lucky to start my studies in an environment that still appreciated the traditional paradigm of historical linguistics, that is, the old German tradition of *Sprachgeschichte*, which in the latter half of the 20th century was being undermined by new and “revolutionary” synchronistic theories of language emanating from the Anglo-Saxon world. My professor in Uralic studies was Aulis J. Joki (1913–1989), who had got his PhD with a thesis on the loanwords of Sayan Samoyedic (Joki 1952). His thesis operates with lexical data from many different languages and language families, including not only Uralic, but also Turkic, Mongolic, Tungusic, Yeniseic, Indo-European and Sino-Tibetan.

When Joki appointed me as his assistant in 1973, one of my first tasks was to participate in the indexing of his second large etymological treatise, this time on the early contacts between Uralic and Indo-European (Joki 1973). Joki was an etymologist *par excellence*, and it was natural that he was also involved as one of the authors of the Finnish etymological dictionary (SKES), initiated by Y. H. Toivonen (1890–1956) in 1955 but completed only in 1981. As an etymologist, Joki represented the paradigm of *Wörter und Sachen*, and he was the author of several insightful papers on important trans-Eurasian cultural words, including ‘apple’ (Joki 1964) and ‘salt’ (Joki 1969).

Since I initially also specialized in the Samoyedic languages, I soon realized that the general understanding of their diachrony was still insufficient. Although M. A. Castrén (1813–1852) had already in 1845, on the basis of lexical comparisons, proposed a genetic connection between Finno-Ugric and Samoyedic (Castrén 1846), this idea was not generally accepted during the latter half of the 19th century, and even in the 20th century there were some linguists, notably Gerhard Doerfer (1920–2003), who refused to recognize the Uralic status of Samoyedic. This scepticism was without a doubt due to the fact that the relationship had not been demonstrated in a sufficiently convincing way, even though the work of Heikki Paasonen (1865–1919) was generally considered to have settled the issue (Paasonen 1917).

As a result, I decided to work out a comparative list of internal Samoyedic etymologies (SW), which comprises all those Samoyedic lexical items that are attested in at least one Northern Samoyedic (Nenets, Enets, Nganasan) and one Southern Samoyedic (Selkup, Kamas, Mator) language. It was my intention to continue the work with a second volume, which would have comprised those items that are attested only in two or more either Northern Samoyedic or Southern Samoyedic languages, but this volume never materialized. In fact, the division of Samoyedic into a northern and a southern branch has subsequently been contested, and the internal taxonomy of Samoyedic is indeed more complicated, as was pointed out by Eugene Helimski (1950–2007).

In any case, my work on the reconstruction of Proto-Samoyedic allowed me to approach the question concerning the structure and lexicon of Proto-Uralic, an enterprise which resulted in a list of Uralic etymologies (Janhunen 1981). My list was based on the “conventional” assumption that Uralic is composed of two major branches, Finno-Ugric and Samoyedic, which means that only those lexical items that are attested in both branches can be regarded as verifiably Uralic. I still hold this understanding of Uralic, although some younger Uralists have recently wished to reclassify the internal relations within the family. The fact is that Samoyedic has less than 200 roots in common with any Finno-Ugric branch, while the Finno-Ugric branches all share a larger number of etymons with each other.

Another issue I wished to emphasize in my Proto-Uralic word list was the principle of minimalism: *less is more* also in etymology. A good corpus of etymologies for taxonomic purposes should be as compact as possible. Even when dealing with fully “regular” comparisons, we should always estimate the “value” of an etymology in relation to its phonological complexity, semantic plausibility and geographical coverage. As a Uralist I cannot avoid the impression that Indo-Europeanists are often too loose about these criteria, which means that the Indo-European comparative corpus has grown too large to be true. Needless to say, attempts to “prove” distant relationships with inflated corpora comprising sometimes even thousands of etymologies are counterproductive and serve only to illustrate the hopelessness of such efforts.

I see etymology as a tool, rather than as a goal. While diachronic linguistics, in general, is a tool for etymology, etymology is a tool for understanding the history of language relationships, language contacts, and cultural influences. Etymology is also an important basis for conclusions concerning the dating, both relative and absolute, of linguistic phenomena. For instance, judging by the composition of the reconstructable Uralic lexicon, we have to conclude that the speakers of Proto-Uralic represented a boreal community at the Mesolithic level of cultural evolution. Unlike the situation in Indo-European, we simply have no agricultural or other more developed technological terminology that can be reconstructed for Proto-Uralic.

Etymology can also help solve long-disputed taxonomic issues. Róna-Tas – whose classes I had a chance of attending in Szeged back in 1974 – has pointed out that the Chuvash and Common Turkic items for ‘stirrup’ exhibit the phenomenon of rhotacism, that is, the positional change of **s* to **r* in the Bulgharic branch of (Macro-)Turkic (Róna-Tas 1973). Since we know that stirrups were invented in the first millennium BC we can date the phenomenon of rhotacism accordingly and accept the fact that the rhotacist elements in Mongolic and Tungusic (as well as in Samoyedic) can only be loanwords from an early form of Bulgharic, transmitted in the last centuries BC.

Etymology is often considered to be a rather mechanic field of inquiry, in which hard labour automatically yields results. There are linguists who are always able to find an etymology that suits their preconceived idea of what *should* be found. For instance, many Indo-European loanwords in Finno-Ugric, especially in Finnic, can be variously explained as Germanic, Baltic, Aryan, “Pre-Germanic”, or even earlier Proto-Indo-European. Each etymologist tries to support his or her point of view, often with formally elegant but ultimately questionable arguments. This reminds us of the anecdote according to which Zoltán Gombócz (1877–1935), when asked to detect the etymology of a particularly difficult item, answered: “From which language do you wish?”

However, in real etymological work intuition plays a central role. People without etymological intuition are liable to make false conclusions, often based on preconceived stereotypic ideas. In this sense, etymology is like music: some people have an ear for music, while others do not. Distant relationships are typically proposed by people who may or may not be well-trained linguists in other respects, but who lack an inherent ability to understand language evolution. Likewise, there are linguists who are blind to actual relationships supported by true etymological evidence. Too much critique is just as bad as too little critique, but finding the right balance is a delicate thing.

Let me quote a personal example of what I think is intuition in etymological work. In the early 1980s, I was camping in a remote forest area in central Finland with my colleague and friend Asko Parpola, the Indologist. Looking at the flames of the camp fire I thought of the Finnish word *pala-a* ‘to burn’. What was the fire doing? It was *eating up* the wood. I knew that the Khanty expression for ‘to burn’ is etymologically identical with the Uralic verb ‘to eat’ (DEWO 713–716 s.v. *li-essen* = ‘brennen’). Could the Finnish word also be like that? Yes, of course: it makes a perfect cognate for Samoyedic **palä-* ‘to swallow’ (SW 116). Moreover, *pala-* ‘to burn’ is homonymous with the Finnish noun *pala* ‘piece’, as in *suu+pala* ‘mouth-piece’ = ‘bite’.

After these contemplations, the etymology of Finnish *pala-a* was clear to me: we are dealing with a Uralic *nomen-verbum* **pala* ‘bite’ > ‘piece’ : **pala-* ‘to bite,

to swallow, to eat' > 'to burn'. The connection of 'piece' and 'to bite' is, of course, evident from many languages, as in Swedish *bit* 'bite' > 'piece' : *bit-a* 'to bite', while the semantic transition from 'biting' to 'burning' is confirmed by the Khanty parallel. Moreover, although nouns and verbs are morphologically clearly distinct in Uralic, ambivalent *nomina-verba* seem to have been fairly common in Proto-Uralic, a circumstance that may be indicative of a typological change that had taken place rather late in Pre-Proto-Uralic – an interesting topic that I would like to take up in the future.

However, my new etymology of Finnish *pala-a* 'to burn' is also an example on how slowly innovations make their way into standard handbooks. The received explanation, quoted by both Finnish (SKES 471–472) and Hungarian (MSzFE 172–173, TESz 825–826) etymological dictionaries, and also supported by Björn Collinder (FUV 64, 120), separates 'burning' from 'biting', but connects 'burning' with 'freezing', as in the Finnish frequentative *pale-l-la* 'to freeze'. The latter is then linked to Hungarian *fagy* 'frost' : *fagy-* 'to freeze'. Since Hungarian also has *fal-* 'to devour', two separate reconstructions are required: **pala-* 'to bite' vs. **pal'a-* 'to freeze', with a different medial consonantism.

It is immediately clear that the connection between 'burning' and 'freezing' is real and reflects the similar sensory and physiological effects of burning and freezing. In Mordvin the two meanings are expressed by the same basic verbal root **pala-* > *palo-ms* 'to burn; to freeze' (MWb 1516–1517). Most probably, the semantic transition was from 'biting' or 'eating' to 'burning' to 'freezing', but there is also the possibility of a direct connection between 'biting' and 'freezing', as is evident from expressions such as *frostbite*. In any case, Finnish *pala* 'bite' : *pala-* 'to burn' : *pale-l-* 'to freeze' can be easily derived from one single source, the Proto-Uralic *nomen-verbum* **pala(-)*, while Hungarian *fagy(-)* is formally different and must represent a separate etymon.

Even so, the most recent – though no longer up-to-date – Uralic etymological dictionary still repeats the dual etymology of **pala-* 'bissen, fressen' vs. **pal'a-* 'frieren' (UEW 350, 352). The same opinion is held by Róna-Tas in his treatment of Hungarian historical phonology (WOT 1275, 1302–1303). On the other hand, the new Finnish etymological dictionary already mentions the semantic connection between 'biting' and 'burning' as a possibility (SSA 2: 298–299), though it does not take a firm stance on the issue. At the same time, it notes the phonetic difficulty of connecting Hungarian *fagy(-)* with the Western Uralic data.

Looking for a more plausible etymology for Hungarian *fagy(-)*, one is tempted to link it to Hungarian *fáz-ik* 'to freeze', as was proposed in older etymological sources (e.g. MSzSz 71). However, this comparison is rejected by later etymologists, who prefer to find separate comparative evidence for both *fagy(-)* and *fáz-ik* in the Ugric languages. In reality, this evidence is rather thin, for *fagy(-)* has a possible cognate only in Mansi (with more problematic counterparts in Khanty), while *fáz-ik* has a possible cognate only in Khanty (with an areally transmitted counterpart in

Northern Mansi). Due to their limited distribution, neither of these etymologies may be regarded as certain.

We might nevertheless opt for the possibility that Hungarian *fagy(-)* is a cognate of Mansi *pāl'-i-* ‘to freeze’, as is unanimously maintained by the standard sources. If so, the most probable reconstruction would be **palya-* ‘to freeze’, in which the cluster **l+y* would have yielded Hungarian <gy> (voiced palatal stop), as also in *négy* : *negy-* ‘four’ < **nelya* and possibly *figy-el* ‘to listen’ < **pelya-* (WOT 1034). As an item peculiar to Hungarian and Mansi, **palya-* ‘to freeze’ could be one of the innovations shared by these two languages, which also otherwise seem to form a distinct branch of Uralic, either within the context of “Ugric” (with Khanty) or perhaps separately (without Khanty).

Whatever the background and mutual relationship of Hungarian *fagy(-)* and *fáz-ik* may be, it has to be concluded that my intuition at the campfire was correct: Finnish *pala* ‘bite’ : *pala-* ‘to burn’ : *pale-l-* ‘to freeze’ are all reflexes of Proto-Uralic **pala(-)* ‘bite; to bite’, an etymon that is also present in Hungarian *fal-* ‘to devour’ (: *fal-at* ‘bite’). There is, however, also another issue that this example teaches us: one should never think that an etymology is “new” before one has checked the extant literature, which in the case of comparative Uralic studies is massive and covers at least 200 years of etymological work.

In the actual case, I thought I was the first one to have noticed the connection between *pala-* ‘to burn’ and **pala-* ‘to swallow’. Only much later did I realize that this connection had already been proposed by Toivo Lehtisalo (1887–1962), who also correctly separated Hungarian *fagy(-)* from the Uralic *nomen-verbum* **pala(-)* (Lehtisalo 1933: 236–237). So, the priority of the correct conclusion belongs to Lehtisalo, not me. However, Lehtisalo’s observation was totally ignored by the standard handbooks. Only after I had independently arrived at the same conclusion half a century later has the idea become more widely known. It may take another half a century for the correct explanation to win its place in the pool of generally accepted etymologies.

Juha Janhunen
Department of World Cultures
00014 Helsinki, Finland
[asiemajeure@yahoo.com]

R e f e r e n c e s

- Castrén, M. A. (1846). Anteckningar om Samojediskans förvandtskap med de Finska språken. *Suomi*, vol. 5, pp. 177–186. Helsingfors.
- Joki, Aulis J. (1952). *Die Lehnwörter des Sajansamojedischen*. Mémoires de la Société Finno-Ougrienne 103. Helsinki.

- Janhunen, Juha (1981). Uralilaisen kantakielen sanastosta. *Journal de la Société Finno-Ougrienne*, vol. 77, pp. 219–274. Helsinki.
- Joki, Aulis J. (1963). Omenan vaellus. *Virittäjä*, vol. 67, pp. 134–142. Helsinki. [German version:] Der wandernde Apfel. *Studia Orientalia*, vol. 28, pp. 1–17. Helsinki, 1964.
- Joki, Aulis J. (1969). Suolan suku. *Acta Universitatis Tamperensis*, vol. A 26 (Juhlakirja Paavo Siron täyttäessä 60 vuotta), pp. 47–54. Tampere.
- Joki, Aulis J. (1973). *Uralier und Indogermanen: Die älteren Berührungen zwischen den uralischen und indogermanischen Sprachen*. Mémoires de la Société Finno-Ougrienne, vol. 151. Helsinki.
- Lehtisalo, T[oivo] (1933). Uralische Etymologien. *Mémoires de la Société Finno-Ougrienne*, vol. 67 (Liber semisaecularis Societatis Fenno-Ugricae), pp. 233–237. Helsinki.
- Paasonen, Heikki (1917). *Beiträge zur finnischugrisch-samojedischen Lautgeschichte*. Budapest: Druck des Franklin-Vereins.
- Róna-Tas, András (1973). Did the Proto-Altaiic people know the stirrup? *Studia Mongolica*, vol. 1 (9), fasc. 13, pp. 169–171. Ulan-Bator: Permanent Committee, International Congress of Mongolists.

Lexicographical sources

- DEWO = Wolfgang Steinitz (ed. in-chief). *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. Abhandlungen der Deutschen Akademie der Wissenschaften zu Berlin. Berlin: Akademie Verlag, 1966–1993.
- FUV = Björn Collinder. *Fenno-Ugric Vocabulary: An Etymological Dictionary of the Uralic Languages*. A Handbook of the Uralic Languages, part 1. Second revised edition. Hamburg: Helmut Buske Verlag, 1977.
- MSzFE = Lakó György (ed. in-chief). *A magyar szókészlet finnugor elemei: Etimológiai szótár*, vols. 1–3. Budapest: Akadémiai Kiadó, 1967–1978.
- MSzSz = Bárczi Géza. *Magyar szófejtő szótár*. Budapest: Királyi Magyar Egyetemi Nyomda, 1941.
- MWb = H. Paasonens *Mordwinisches Wörterbuch*, vols. 1–6. Zusammengestellt von Kaino Heikkilä, bearbeitet und herausgegeben von Martti Kahla. Lexica Societatis Fenno-Ugricae, XXIII. Helsinki: Suomalais-Ugrilainen Seura, 1990–1999.
- SKES = Y. H. Toivonen & Erkki Itkonen & Aulis J. Joki & Reino Peltola. *Suomen kielen etymologinen sanakirja*, vols. 1–7. Lexica Societatis Fenno-Ugricae 12. Helsinki, 1955–1981.
- SSA = Erkki Itkonen & Ulla-Maija Kulonen (eds. in-chief). *Suomen sanojen alkuperä*, vols. 1–3. Suomalaisen Kirjallisuuden Seuran Toimituksia, vol. 556. Helsinki: 1992–2000.

- SW = Juha Janhunen. *Samojedischer Wortschatz: Gemeinsamojedische Etymologien*. Castrenianumin toimitteita, vol. 17. Helsinki: Suomalais-Ugrilainen Seura, 1977.
- TESz = Benkő Loránd (ed. in-chief). *A magyar nyelv történeti-etimológiai szótára*, vols. 1–4. Budapest: Akadémiai Kiadó, 1967–1984.
- UEW = Károly Rédei (ed. in-chief). *Uralisches etymologisches Wörterbuch*, vols. 1–3. Wiesbaden: Otto Harrassowitz, 1988–1991.
- WOT = András Róna-Tas & Árpád Berta. *West Old Turkic: Turkic Loanwords in Hungarian*, parts 1–2. *Turcologica*, vol. 84. Wiesbaden: Harrassowitz Verlag, 2011.

