Contextualizing historical lexicology

The state of the art of etymological research within linguistics

University of Helsinki, May 15–17, 2017

Abstracts

Organized by the project
“Inherited and borrowed in the history of the Uralic languages”
(funded by Kone Foundation)
Contents

I. Keynote lectures ........................................................................................................ 5

  Martin Kümmel
  Etymological problems between Indo-Iranian and Uralic .................. 6

  Johanna Nichols
  The interaction of word structure and lexical semantics ............... 9

  Martine Vanhove
  Lexical typology and polysemy patterns in African languages ...... 11

II. Section papers ....................................................................................................... 12

  Mari Aigro
  A diachronic study of the homophony between polar question
  particles and coordinators ................................................................. 13

  Tommi Alho & Aleksi Mäkilähde
  Dating Latin loanwords in Old English: Some methodological
  problems .......................................................................................... 14

  Gergely Antal
  Remarks on the shared vocabulary of Hungarian, Udmurt and Komi
  ........................................................................................................ 15

  Sofia Björklöf
  Areal distribution as a criterion for new internal borrowing ........ 16

  Stefan Engelberg
  Etymology and Pidgin languages: Words of German origin in Tok
  Pisin ............................................................................................. 17

  László Fejes
  Sound change from the point of view of etymology and grammar. 18

  Roman Gaidamaško
  On reconstructing the history of Komi, Mansi, and Russian language
  contact in Northern Kama region .................................................. 19

  Alexandra Grigorieva
  Finnish food words and European culinary vocabulary: Diachronic
  mapping of etymological contexts ................................................. 20

  Riho Grünthal & Lotta Jalava
  Spatial postpositions and lexical change in Uralic ...................... 21
Mikko Heikkilä
(Seeming) irregularity as a clue to the origin of a word................. 23

Sampsa Holopainen
Stem vowels in Proto-Uralic words of Indo-Iranian origin......... 23

Adam Hyllested
A stratigraphy of Finnic and Saami loanwords in Germanic........ 25

Tjaša Jakop
Loan words in Slovenian dialects ............................................... 27

Santeri Junttila
Latvialaisia lainoja mordvalaiskielissä?................................. 28

Anders Richardt Jørgensen
Dialectal features in French loan words in Middle Breton........... 29

Vladislav Knoll
On the history of Low German influence in Slavonic languages ... 30

Guus Kroonen
Did Pitted Ware groups leave a linguistic trace in Germanic and
Finnic?......................................................................................... 31

Juha Kuokkala
Statistics and stratistics of the Germanic loanwords in Finnic ...... 31

Denis Kuzmin
Karjalankielisen esikristillisen naisnimistön rekonstruoointi ......... 32

Elena Lastochkina
Концепт вӱд «вода» в языковой картине мира мари................. 33

Jyri Lehtinen
Assessing computational linguistic phylogeny with Uralic historical
phonology..................................................................................... 34

Anna Makarova
Landscape terms in the substrate toponymy of Belozerye: Semantic-
etymological research................................................................. 36

Julen Manterola & Joseba Lakarra
Basque etymological studies: Loanwords, roots and a new
etymological dictionary.............................................................. 37

Niklas Metsäranta
Loans between branches in reconstruction............................... 38
Sergey Myznikov
Balto-Finnic influence on a wide language background .................. 39

Simonas Noreikis
Lithuanian hydroyms supposed to be Finno-Ugric ......................... 40

Minerva Piha
Combining historical lexicology with archaeology: An example from the South Saami prehistory ...................................................... 41

Juho Pystynen
The state of the Finno-Permic lexicon .......................................... 42

Ilona Rauhala
The layers of loanwords in the adjectives of North Saami .............. 43

Michael Rießler, Mervi de Heer, Terhi Honkola, Unni-Päivä Leino,
Kaj Syrjänen, Outi Vesakoski
Basic vocabulary and the phylogenetic approach to the study of Uralic language history .......................................................... 44

Jack Rueter
Parallels in noun and verb stem types in the Mari and Mordvin languages ............................................................................... 45

Janne Saarikivi
Defining Baltic and Slavic in Finnic vocabulary ............................. 47

Johan Schalin
The notorious cruxes of Common Scandinavian umlaut and breaking: A metaphonic feature-based unified solution .................. 48

Revisiting the quantitative phylogeny of the Uralic languages ...... 49

Yoseph Tezazu
Semantic development of Amharic words ........................................ 50

Outi Vesakoski, Mervi de Heer, Terhi Honkola, Unni-Päivä Leino, Jyri Lehtinen, Jenni Leppänen, Timo Rantanen, Kaj Syrjänen & Niklas Wahlberg
Historical linguistics and evolutionary framework (Poster) ........... 51
I. Keynote lectures
Martin Kümmel  
Friedrich-Schiller-Universität Jena

Etymological problems between Indo-Iranian and Uralic

At least since Jacobsohn 1922, the high relevance of pehistic contacts between Indo-European, especially Indo-Iranian, and Uralic has been well known. However, while Uralicists have worked extensively on this topic until today (cf. especially Jacobsohn 1934; Joki 1973; Rédei 1983; Katz 2003; Kallio 2004), it has got much less attention in Indo-European and Indo-Iranian studies, where the evidence of Uralic contacts is often neglected. Although Uralic evidence is regularly mentioned in the most recent etymological dictionary of Old Indo-Aryan (Mayrhofer 1992–2001), this is not the rule in other works dealing with the prehistory of Indo-Iranian. And indeed, one must admit that it may have been wise not to rely too much on such evidence. In recent years, the field of comparative Uralic linguistics has developed very much, and the reconstructions have changed considerably. Even the traditional internal classification of Uralic has been challenged (cf. aikio 2006; Häkkinen 2009), and the result of all this is that older works can only be used with caution. E.g., Katz 2003 (cf. also Aikio & Kallio 2005) still builds very much on dating loanwords by alleged inner-Uralic nodes like Proto-Fenno-Ugric, Proto-Fenno-Permic etc. Since these nodes are now in doubt, arguments based on their existence have become problematic. All this leads to new interpretations that may have major repercussions for etymological research. Katz 2003 argued that most IIr. loanwords reflect “primary palatals“ as PU/PFU *ś́ so that we should reconstruct a palatal sibilant for Proto-Indo-Iranian (PII), since an affricate would have been taken over as such. This clashes with the current reconstruction of these PII and even Proto-Iranian sounds as affricates (cf. Mayrhofer 1989; Kümmel 2007; Lipp 2009; Lubotsky 2017). This reconstruction has been corroborated by the insight of Koivulehto (1999; 2001; 2007) that younger loans from Iranian in Uralic still reflect affricates that had already become depalatalized. If younger loans have an affricate, it is hardly possible that the older loans without depalatalization already had a sibilant. Thus it follows that the Proto-Uralic (PU) sound traditionally reconstructed as *ś must have been *ć instead, although mostly continued as a sibilant (already Sammallahti 1988 had shown that a contrast between *ś and *ć cannot be shown to be
old). Zhivlov (2014: 114 n. 3) has now also argued for this reconstruction on purely internal grounds. Perhaps even more important is the revision of PU vocalism in recent studies (Aikio 2012; 2015; Kallio 2012; Zhivlov 2014). The new reconstruction of *ćarwi ‘horn’ (Zhivlov 2014; Aikio 2015) instead of *śorwa (Rédei 1988) makes it much easier to derive it from a plausible PII preform, probably *ćr̥w-. Far greater would be the impact of adopting the alternative reconstruction of vocalism as per Tálos (1987; 2015, cf. Abondolo 1998: 16f.), which would make PU vocalism (i a u; ĭ ā ŭ ĭ ĭ vs. Aikio’s and Zhivlov’s i e a o u ŭ ĭ) much more similar to PII (i ø a u; ĭ ā ŭ): e.g., PU *ora ‘awl’ would then be *āra with the same accented vowel as its source *ārā (Vedic ārā-), and there would be no need for either supposing a pre-PII loan from *őrā- or an unclear raising of the vowel of *ārā- (the first would be excluded if the preform was *ēlā- = Germanic *ēlō- and not *ṓlā-, for which there would be no exact comparandum in IE). The aim of my presentation is to look at these and other problems from the Indo-Iranian and Indo-Europeanist perspective and discuss their impact on etymology.

References


http://www.sgr.fi/susa/92/hakkinen.pdf


The interaction of word structure and lexical semantics

Favored and disfavored directions of lexicosemantic change and lexical replacement are not purely semantically driven; structural types of word morphology are good predictors of change. They identify both favored targets and favored sources in change and renewal. Similarly, greater and lesser diachronic stability of lexemes are determined not purely by semantic factors but also by morphological structure type. In previous work I have raised some hypotheses about these patterns, based mostly on individual etymologies. Now I will attempt to test them more systematically on the larger sets of data collected for various projects in lexical typology. Specifically:

(1) Event structure. In sets consisting of static (e.g. German sitzen; kennen), dynamic (sich setzen; erkennen, kennenlernen), and causative (setzen, bekannt machen; gap for causative of ‘recognize’, perhaps hinweisen), where the static member is the derivational base of the whole paradigm the set is stable; where the base is dynamic or causative the whole set is unstable. Thus, to the historically static-based German sets with their ancient IE roots contrast Spanish static estar de pie ‘stand’, dynamic levantarse, causative levantar with a new root throughout the paradigm, or estar sentado ‘sit’, dyn. asentarse, caus. asentar, with the old IE root replaced by a participial derivative. A favored source of
renewal of both dynamic and causative senses is colorful and/or slangy words for human actions.

(2) Causative alternation. Where intransitives are derivationally basic, verbs are stable; where transitives are basic they are less stable.

(3) In verb-based languages (where verbs are mostly basic and nouns often derived), verbs are stable and nouns less stable or variable; in noun-based languages it is the reverse.

(4) Classification: It is well known that verbs are prone to classify for properties of the S/O (commonly, shape and number). But this describes only intransitives and monotransitives. Ditransitives tend to classify for the T (the more theme-like or patient-like of the two objects), e.g. Russian mazat' ‘apply’ (goo to surface or object), classified for the texture of the T and with no implications about the receiving surface or object; Ingush tuox- ‘strike’ where the T is point-like or a projectile and moves outward and not laterally. Nouns referring to common or prototypical T’s are a common source or target of renewals.

In all of these the derivational base and the central meaning are targets of renewal. Human or animate status of the S/O, and distinctive motion of S/O or T, enhance the likelihood of renewal actually occurring. These various pressures, realized only occasionally but with typological consistency, over time can give a distinctive typological shape to the lexicon of a language. Apart from questions of overall lexical type, they can act as a guide to plausible etymologies.
Lexical typology and polysemy patterns in African languages

The *RefLex* reference comparative database for African languages (Segerer & Flavier 2011–2016) covers over a million entries for over 700 languages. Included are sets of search tools for polysemy patterns that ease identification of areal or genetical patterns of colexification for intrafield and transfield polysemies, as well as a map tool to visualize them.

Partly building on previous works, this presentation will first present some results in several semantic domains: sensory modalities, (some) body parts, temperature (hot/warm; cold), child, want, and work. It will then discuss methodological, theoretical and heuristic issues related to the following questions: How does type of data frame our view of polysemy (dictionaries, grammars, intuition, questionnaires, large databases)? How to identify what is not universal, not borrowed, not inherited? These questions represent real challenges from Sprachbund studies, diachronic and typological studies alike.

References


II. Section papers
A diachronic study of the homophony between polar question particles and coordinators

Several authors have pointed out that in many unrelated languages the polar question particles are homophonous with connectives, such as disjunction and conjunction markers (Bailey 2013; Jayaseelan 2008; Jaggar 2001; Carlson 1994). In some cases the homophony can still be observed while in others the question particle can be etymologically traced to a historical connective.

This research postulates that on the contrary to a widely held view that this is a syntactic phenomenon, this homophony is diachronic by nature. There seem to exist pressures which lead languages to have their polar question particles grammaticalised from this one specific category of function words.

The research will map the connectives and polar question markers in Estonian, Lithuanian, Polish and Hungarian. It will then show that on a closer look the question particles and connectives of several languages, such as the Estonian kas or the Polish czy are not entirely grammaticalised, being either in the process of losing certain properties or yet to gain them. The contexts in which these particles are allowed or restricted are not random, but form a system that supports the pathway hypothesis.

It will also define the possible links in the new grammaticalisation pathway by conducting a case study in Estonian, showing the grammaticalisation pathway of the contemporary initial polar particle kas. The written records going back four centuries will show that the decline of the conjunctive marker kaas and the emergence of the question particle kas fall in the same time period, in accordance with the S-shaped model of historical syntax.

Finally, the research will show that this grammaticalisation pathway is not an isolated case, but the continuation of a pattern of already identified pathways between certain case inflections and connectives. It will explain
the semantic connections between the stages of the process and attempt to give reasons for the force behind the change.

Bibliography


Tommi Alho & Aleksi Mäkilähde
Åbo Akademi University, University of Turku

Dating Latin loanwords in Old English: Some methodological problems

Latin loanwords in Old English are traditionally divided into two (Pogatscher 1888) or three (Serjeantson 1935) periods: continental (before around 400–450 AD), early insular (around 450–600), and late insular (after around 600). However, the division between continental and early insular loanwords is problematic due to the great number of uncertain cases (see e.g. Durkin 2014: 104–105). Furthermore, although the evidence offered by Vulgar Latin is often crucial in dating these loanwords, problems in the interpretation of the Latin evidence are not always acknowledged sufficiently. In this paper, we focus on two types of Vulgar Latin sound changes and the problems they pose for dating Old English loanwords. The first is the loss of contrastive vowel-length in the Latin sound system and the resulting front and back vowel mergers (i, e > [e]; u, o > [o]). The second is the voicing of intervocalic plosives. Using a few well-known problematic cases (e.g. *episcopus* > *bisceop*) as examples, we argue that investigations of Latin loanwords in Old English should re-examine more recent research on Vulgar Latin (e.g. Loporcaro 2015), and re-evaluate the problems posed by the particular nature of data sources such as epigraphic texts and metalinguistic testimonies by Roman grammarians.
Gergely Antal  
Eötvös Loránd University (Budapest)  

Remarks on the shared vocabulary of Hungarian, Udmurt and Komi  

In my study – which is mainly the topic of my thesis for my master degree – I focus on new possible cognates within the Uralic language family, especially looking deeper into the relationship of Hungarian and the Permic languages, namely Udmurt and Komi. It has been suggested and is currently accepted that Old Permic and Old Hungarian had been in contact with each other resulting in a vocabulary found in modern Hungarian which can be traced back to the Permic languages. In the first part of my discussion I would like to measure this contact more precisely by regarding the details and directions of borrowing words between the two. I argue that in some cases the former suggestions of some loanwords should be revised, primarily for culturally linked terms based on geographical and archaeological facts. In the second part I would also like to propose new etymological approach to Hungarian words with no known etymology which in my opinion are linked to Udmurt and Komi cognates, but can be also found in other Uralic or Finno-Ugric languages. I also describe a new possible Finno-Ugric reconstruction of the Hungarian word sügér ‘perch’ linking it to the Udmurt word чорыг ‘fish’ and the Komi word черу ‘fish’.  

References  

Sofia Björklöf  
University of Helsinki  

Areal distribution as a criterion for new internal borrowing

The use of areal distribution as a criterion for internal borrowing between closely related Finnic languages is at the focus of this presentation. Two case studies are presented to describe the use of this criterion. The first contact situation analysed in this presentation is that of the contact between the western subgroup of the Estonian northeastern coastal dialect and Finnish dialects; the second is the contact among Vote, Ingrian, Estonian, and Finnish dialects, which took place in Western Ingria and northeastern Estonia. The Finnic languages are so close to each other linguistically and have such a great amount of shared vocabulary that, traditionally, many of the shared words in these closely related languages have been considered as having been inherited by these languages from their common proto-language. In etymological dictionaries, this question has often been settled through word comparison rather than by presenting information concerning the origin of these words (the numerous Finnish loans in the Estonian Kuusalu dialect are an exception). Even the distribution of a word in a single parish has been sufficient to consider the word in question equally old in all varieties. In dictionaries focusing on a specific variety, internal borrowing in Finnic has been marked to a certain extent (the best example of this is Laakso ed. 1995, Vatjan kielen Joenperän murteen sanasto [Vocabulary of the Joenperä dialect of Vote]). However, defining borrowing relations and the direction of diffusion demands detailed study of vocabulary and the data concerning its areal distribution. When identifying newer borrowings, areal distribution is almost the only criterion with which relationships can be established, as outside of Vote, there are relatively few phonetic correspondences in the Finnic languages with which borrowings can be recognised. When identifying older substrate layers some support can at times be gained from phonetic shape for which the second case study serves as an example.
Stefan Engelberg  
Institut für Deutsche Sprache (Mannheim)

Etymology and Pidgin languages: Words of German origin in Tok Pisin

The Institute of the German Language in Mannheim is currently developing an internet portal for German loanwords in other languages (lwp.ids-mannheim.de). The portal collects bilingual loanword dictionaries and implements complex search options for borrowing processes filtered by grammatical, temporal, semantic and other parameters.

One of the dictionaries that is currently being compiled for the portal is the *Internet Dictionary of Words of German Origin in Tok Pisin* (TP), the main pidgin language of Papua New Guinea. The German influence onto TP has its roots in the German colonial empire, when TP played an important role as a lingua franca in the colony of German New Guinea.

As a pidgin language, TP poses particular problems for etymological research: In its early stages it was unstable, the speakers spoke different native languages and TP as a second language, words varied phonetically according to the origin of the speakers, and there is no inherited vocabulary that forms the background for borrowing processes. Besides Austronesian and Papuan languages, several non-local languages have contributed to the lexicon of Tok Pisin, in particular English (80%), German, and Bazaar Malay, which itself had integrated many Dutch and English loanwords. The existence of Dutch-German-English cognates often makes it difficult to establish the origin of TP words. In some cases, phonetic adaptation strategies of TP point to a specific origin, in other cases the lexicon of older Pacific pidgins provide hints, and often knowledge about particular historical circumstances reveals the origin of a TP word. Nevertheless, in many cases, a multiple origin has to be assumed, or at least a strengthening effect of, e. g., German on the establishment of English loans in TP that had already been in use in pre-TP pidgins.
The talk will briefly present the lexicographic enterprise and then focus on the problems of etymological reasoning specific to the pidgin language under investigation.

László Fejes  
Research Institute for Linguistics, Hungarian Academy of Sciences

Sound change from the point of view of etymology and grammar

Nganasan stems are divided into two harmonic classes, U stems and I stems, according to Helimski (in Abondolo 1998) and Várnai (in Wagner-Nagy 2000). They claim that the harmonic class “in many cases” or “often” cannot be determined on the basis of the vowels the stem contains. However, the descriptions do not discuss cases when the harmonic class can be determined. My recent research shows that Nganasan vowels can be classified into three sets: U class vowels, I class vowels and neutral or ambivalent vowels. If the stem contains exclusively U class vowels, it will be a U stem and if it contains exclusively I class vowels, it will be an I stem in more than 95% of the cases. If the stem contains both U and I class vowels, vowels of all the three classes or exclusively ambivalent vowels, the stem class is unpredictable (however, it remains fairly predictable in some subpatterns).

Helimski (in Abondolo 1998) claims that this kind of Nganasan vowel harmony originates from a frontness/backness harmony. The vowel chain-shift presented by him suggests that U class vowels were back and I class vowels were front about three or four hundred years ago. But how regular this harmony could be? How did it become more irregular owing to sound changes? How can revealed sound changes in stems help us to understand grammatical change? Can understanding grammatical changes help us in the reconstruction of stems? These kind of questions arise when we try to understand the history of Nganasan vowel harmony.
Roman Gaidamaško  
Institute for linguistic studies of RAS  

On reconstructing the history of Komi, Mansi, and Russian language contact in Northern Kama region

The key principles of interpreting the lexical substrate rest upon the methods of areal linguistics. Thus, according to W. Veenker, O. V. Vostrikov, S. A. Myznikov, J. Saarikivi, A. Aikio and other scholars, for the successful etymological interpretation of the lexical substrate, it is necessary to take into account the shape, the semantics, and the area of supposed substrate lexeme, including information about the distribution of its accentological, phonetic, word-formative variations and derivatives.

It should be noted that features of etymological research are directly dependent on the local linguistic realities and historical background. For example, it is well known that prior to Russian colonization of Northern Kama region, local Finno-Ugric peoples (Komi and Mansi) have had a common lexical fund. There could be at least two sources of formation of this fund. Firstly, it can be cognates — lexemes, that are inherited by Komi and Mansi from Finno-Ugric language age. Secondly, it can be a lexicon that is formed as a result of later convergence processes (lexical borrowings from Komi to Mansi, and vice versa). Sometimes these words get into Russian dialects, cf. Russ. арай [araj] (< Komi-Permyak), урай [uraj] (< Mansi) ‘lowland moist or swampy place’, Russ. тулым [tulum], тулымник [tulymnik] ‘a rapid; a bunch of stones in a riverbed’ (~ Komi-Yodzyak ~ Mansi).

The lexical material shows areas of substrate Finno-Ugric dialects and intersections of the lexical isoglosses. Thus, in the Northern Vishera basin, earlier Mansi substrate type boundaries are blurred from the west by a later Komi(-Permyak) substrate type.

This report presents fragments of etymological study of lexicon, that is the result of Komi, Mansi, and Russian language contact in the Northern Kama region. New etymological solutions are proposed for some Russian dialect words.
Finnish food words and European culinary vocabulary: Diachronic mapping of etymological contexts

In my talk I will explore four main loanword layers (Proto-Germanic, Swedish, Russian, and English) that have particularly influenced the development of Finnish culinary vocabulary. Of course many Finnish food words, such as *kala*, *sieni*, *vöi* are originally Finnic. However, other basic words such as *leipä*, *kinkku*, *juusto* etc. are not, they are results of very old contacts with Indo-European languages (Häkkinen 2007). Newer borrowings made in the last few centuries such as *pihvi*, *konvehti*, *hodari* have also been assimilated by Finnish to a point that a native speaker would hardly be aware of their origins. Almost every such borrowing reveals an aspect of common European food culture, and forms part of a general European lexicological pattern even though Finnish comes from a different language family. My research method involves use of naming universals I have formulated while working on diachronic Latin culinary vocabulary (Grigorieva 2005, 2008), as for the etymological approach it comprises both semasiological and onomasiological level (Geeraerts 2015).
Riho Grünthal & Lotta Jalava
University of Helsinki

Spatial postpositions and lexical change in Uralic

The category of adpositions is universally wide-spread (Dryer 2013, Hagège 2010) and covers all Uralic languages. Typologically the language family is quite uniform as postpositions are attested in every languages, whereas prepositions occur only in Saamic and Finnic, the two northwestern branches, and a clearly less frequent than postpositions. Spatial relations are characteristically manifested by both case suffixes, adpositions and relational nouns. From the diachronical viewpoint new adpositions are constantly introduced while others take grammatically more specific functions or may be totally lost. Adverbs and relational nouns, most notably body-part-nouns, constitute the two main categories from which new adpositions are adopted (not only in Uralic, but typologically e.g. Haspelmath 2003: 40, Heine & Kuteva 2002: 46–50).

In this paper, we focus on the maintenance and renewal of the oldest reconstructable layer of Uralic spatial postpositions in terms of etymology, lexical semantics, morphology and language typology. There are five postpositions which have a geographically wide distribution in Uralic and, hence, can be reconstructed in Proto-Uralic: *ala- ‘under, below’ (Sammallahti 1988: 536 *ila-), *eđi- ‘front’, *miŋä- ~ *mǔŋä- ‘behind; after’, *taka- ‘behind’ (cf. Finnic and Samoyedic ‘behind’ in (1a and 2a)), and *(w)üli- ‘above’ (Aikio 2015: 33–34; Sammallahti 1988: 536, 541, 546; UEW: 6, 276–277, 506–507 573–574). The contemporary distribution indicates that in many cases the lexical stem has been replaced by a secondary adposition in various Uralic branches. Furthermore, the degree to which these adpositions are adapted to productive morphological rules and display earlier inflectional patterns varies considerably. This is illustrated in examples (1 and 2).

Tundra Nenets
(1a) xarad-ʔ iaxə-na
    house-GEN behind-LOC(PP)
    ‘behind the house’
(1b) xarad-ʔ xew-xəna
    house-GEN side-LOC
    ‘beside the house’
Finnish

(2a) talo-n taka-na (2b) talo-n viere-llä / viere-ssä
house-GEN behind-ESS house-GEN side-ADE / side-INE
‘behind the house’ ‘beside the house’

Our paper sheds light to the renewal and diachronic long-term stability of postpositions as a grammatical and lexical category by means of a detailed morpho-syntactic analysis of the spatial adpositions in Uralic.

References


Mikko Heikkilä  
University of Tampere

(Seeming) irregularity as a clue to the origin of a word

Phonological regularity is a major, although not the only, criterion in the two interdependent fields of historical linguistics, sound history and etymology, which is the methodological reason why proposed etymologies that contain an irregular sound correspondence are a priori inferior to proposed phonologically (fully) regular etymologies as regards the solidity of the etymology. This does not, however, automatically mean that etymologies that seem phonologically irregular at first sight would not be worth a closer scrutiny. Sometimes (seeming) phonological irregularity can, in fact, be an important clue to the origin and history of a word – which can in turn open up for new, previously unknown etymological correspondences and sound historical developments. Such a situation can arise e.g. when a word is borrowed from language A to the mutually related languages B₁ and B₂ via a substrate language S either parallelly or as a borrowing to their common proto-language B₀.

The purpose of my conference talk is to discuss (methodologies of) etymology, loanword layers and general tendencies of phonological development in neighboring related languages in the light of some (new) etymologies in the West-Uralic languages.

Sampsa Holopainen  
University of Helsinki

Stem vowels in Proto-Uralic words of Indo-Iranian origin

In Uralic linguistics, the study of non-initial-syllable vowels (stem-vowels) has always been in the background, and many aspects of their development has been poorly understood until recently (see Kallio 2012 for a survey of recent developments). Because of this, the study of the substitution of non-initial-syllable vowels in various Indo-European loanword layers in the Uralic languages has also been more poorly understood than the substitution of the initial-syllable vowels. However, the study of Germanic and Baltic loans in Finnic has shown that in many
cases the substitution is not based on simple phonetic similarity, but sometimes other factors, such as acquisition of loanwords to a certain Uralic stem-type plays a more significant role (see Junttila 2015 for stem-vowels in Baltic loans of Finnic).

This seems to be true of the Indo-Iranian borrowings also. While some loans show a neat match between PU and PII stem vowels, such as PU *śata ‘hundred’ ← PII *ćata- id., in many of the loans the chosen stem-type is more difficult to understand. This issue is complicated by recent research which has reshaped some of our ideas of Uralic stem-vowels. Aikio’s (2015) new reconstruction influences also the stem-vowels of some well-known PII loans, such as PU *śarwi ‘horn’, instead of traditional reconstruction *śorwa, from PII *ćrwa-.

In this presentation the substitution of the stem-vowels in the earliest Indo-Iranian borrowings in Proto-Uralic will be critically evaluated. The expected results of the research are that it is possible to establish regular substitution patterns and that the choice of stem-type is not arbitrary.

References


A stratigraphy of Finnic and Saami loanwords in Germanic

An important advance in etymological research in recent years is an increased importance of distinguishing different strata within loanword relations. That Finnic and Saami contain many early Germanic and other Indo-European loanwords has long been known, and while it has always been assumed that these borrowings reflect continuous processes of interaction across centuries or even millennia, it has only recently been shown how they can be separated into chronological layers, mainly on the basis of diverging sound substitutions (e.g. Sammallahti 2001; Koivulehto 2003; Aikio 2006, 2012; Kallio 2009, 2012, 2015).

Since I have claimed (e.g. Hyllested 2010, 2012, 2014, 2017) that borrowing at early stages also took place in the reverse direction, from Finnic and Saami to early Indo-European languages in Northern Europe, it seems timely to try to establish a corresponding stratigraphy of this, admittedly much smaller, material. For reasons of time and space, I will restrict myself to Germanic.

A handful of lexemes borrowed before the sound-laws defining Germanic actually operated, such as \( *o > *a \) and \( *k > h \), are nevertheless confined to the Germanic area. I have previously suggested (Hyllested 2010, pace Ruohonen 2010) that PGmc. \( *\text{haljō} \) ‘hell, abode of the dead; female being presiding over it’ is a borrowing from Early Proto-Finnic \( *\text{kolja} \) ‘id.’ (Fi. Koljo ‘a giant connected with death’ ~ Mansi kul’ - in names of the king of the netherworld, but in the rest of Uralic mostly ‘evil spirit’, ‘devil’ or ‘disease’), originally \( *\text{kole-ja} \) ‘the one who lets die’.

Since Saami contains loans from Pre-Proto-Germanic – such as N \( \text{muošmi} \) ‘meat between the thigh and ribs (of a reindeer)’ < Pre-PGmc.\( *\text{moms-mo-} \) ( > PGmc. \( *\text{mamma-} \) > Gothic \( \text{mammo} \) ‘flesh’; Hyllested 2014, pace Aikio 2012) – we may also expect to find Saami loanwords in Pre-Proto-Germanic predating the sound shift and vocalic changes. PGmc. \( *\text{harunda/ō-} \) ‘flesh’ (> e.g. ON \( \text{hörund} \) ‘human flesh; skin’, OSw. \( \text{harund} \) ‘flesh; skin’, Early Mod.Da. \( \text{harend} \) ‘body’) is usually etymologized as a derivative of PIE \( *\text{kṛH-ōn-} \) ‘meat; flesh’ (Lat.
carō), but seems even closer to Proto-Saami *korōntē ‘butchered meat; carrion, carcass’ (> S gārrodh, N gorut, Inari koorood); it is already known to have been borrowed into N Fi.dial. as kurento ‘carcass’, and, remarkably, it belongs to the same narrow semantic field as muošmi which may instigate a discussion about whether such an argument is relevant for two reverse loan directions.

In my paper I sketch out a stratigraphic treatment of my earlier etymological proposals in both directions between Uralic and Indo-European/Germanic, and further present new etymological suggestions.

References

Loan words in Slovenian dialects

The diversity of the Slovenian dialect lexicon is frequently subject to extra-linguistic factors: historical, administrative-political, geographical and so on. Most lexemes are of Slavic origin, while others have been borrowed from neighbouring languages at various periods and are mostly of Germanic origin; and there are a fair number of Romance loan words, in the west in particular (in contact with the Italian and Friulian languages). Lexemes of Hungarian origin are rarest, appearing in the extreme north-east of the country (the smallest part of the Slovenian linguistic area), in contact with the Hungarian language. Along the border with Croatia one can also found “Croatisms”; due to the close relationship between the two south-Slavic languages, Slovenian and Croatian, these are more difficult to distinguish.

This paper presents dialect vocabulary from the semantic fields “human body and family” and “Slovenian (farm)house, its architectural elements, and interior furnishings”. The dialect material from Slovenian linguistic territory was collected for the Slovenian Linguistic Atlas (SLA 1 [2011] and SLA 2 [2016]). The vocabulary for parts of the body is quite uniform and contains mainly non-borrowed words of a Slavic origin, while some other lexemes (e.g. expressions for household fixtures and fittings) show a greater degree of borrowing (predominantly Germanic loan words).

Interpretation of the material in the SLA is based on the interpretation contained in the Slavic Linguistic Atlas (OLA), which examines the distinction between non-borrowed (Slavic) and borrowed lexemes. Inherited (Slovenian) lexemes are phonetically abstracted and are followed by their proto-Slavic transposition and their word-formational
predecessors. Borrowed lexemes are followed by their sources from foreign languages (e.g. Italian, Friulian or German). This paper also highlights the ratio between borrowed and non-borrowed words in Slovenian dialects.

Santeri Junttila
University of Helsinki

Latvialaisia lainoja mordvalaiskielissä?


Kantabaltin g on etuvokaalin edellä muuttunut latiassa ja kantaslaavissa dz:ksi, useimmissa slaavilaisissa kielissä edelleen z:ksi. Liettuan sanalla giminė tai sen kantasanalla gimti, lv dzimt ’syntyä’, ei kuitenkaan ole vastineita slaavissa.

Grünthalin selitystä olisi helppo pitää vääränä, sillä oletus latvian ja esimordovan välisistä kontakteista on sekä maantieteellisesti että kronologisesti ongelmallinen. Oletettava äänenteellinen kehitys ei sinänsä ole epätavallinen, sillä palataalisessa asemassa velaariklusiilit ovat alttiita affrikaatiolle, mutta esimerkiksi itäisistä balttilaisperäisistä hydronyymeistä ei sille ole etsitty tukea. Olen kuitenkin löytänyt mainitusta äännesuhteesta neljä rinnakkaistapausta mordvan tähän saakka etymologiomattomasta sanastosta.

1. ers сиземс, mkš șizėms ’väsyä, uupua’ ~ lv dzist: (prees. 3) dziest: (pret. 3) dzisa ’sammua; jäähtyä, hiipua, himmetä’.

2. ers šijādīms ’шить, вышивать / nähen, sticken’ ~ lv dzija ’Garn’, lt gija ’Faden im Aufzug eines Gewebes’
3. ers šiē, mkš šiā į paral. slovo k napo в проклятиях / Par. - Wort zu paro 'Gutes' in Verfluchungen ~ lv dzirt: dzir: dzīra 'rühmen', dzirties 'sich brüsten, prahlen, sich etw. vornehmen, wollen, versprechen', lt girti: giri: gyrė 'rühmen, loben', gyrius, gyriai, gyra 'Großtun, Prahlerei', geras 'gut, tüchtig, angenehm, trefflich, freundlich'

4. ers šivē, mkš šivā 'награда, плата / Lohn, Belohnung; Bezahlung'; ers šivedems mkš šivādems 'нанимать, нанять; сдавать внаём; наниматься; брать взаймы (за плату) / mieten, in Dienst nehmen, dingen; vermieten; sich verdingen; (gegen Entgelt) leihen' ~ lv dzīve, dziviba 'Leben', dzīvot, -āt 'leben, arbeiten, beschäftigt sein', dzīvotājs 'Lebender, Lebenskräftiger, Arbeiter', vrt. lt gyti 'gesund, heil werden', (vanh.) 'leben'

Anders Richardt Jørgensen
Uppsala University

Dialectal features in French loan words in Middle Breton

In this talk I will present some aspects of an ongoing study of French (and earlier Romance) loan words appearing in Middle Breton (1450–1650). I will first present a basic outline of the different strata of loan words, dating from the early Middle Ages until the end of the Middle Breton period around 1650 AD. After this, I will focus in more detail on the issue of “vanished” donor language features, in this particular case the presence of dialectal innovations appearing in the loan words in Middle Breton, innovations which are either completely unknown in the presumed NW French donor dialect or are no longer present but known from older attestations.
Vladislav Knoll  
Slavonic Institute of the Czech Academy of Sciences  

On the history of Low German influence in Slavonic languages

The Germanic influence in the Slavonic languages is a broadly studied theme both from lexical as well as from the morphosyntactic point of view. The interest of studying the history of the Low German contact consists in the sociolinguistic aspect of such study and the possibility to compare the results within the context of the Low German impact in the Northern Europe in general.

We can distinguish several types of Low German-Slavonic contact situations: there are areas with intensive and durable Low German contact (Pomerania, Wendland in Low Saxonia), areas with the time-limited contact (Lower Lusatia, North-western Russian areas) and the areas with no or mediated contact. The common feature of the areas with higher Low German influence is a complicated sociolinguistic development, characterized by graduated withdrawal of both Slavonic and Low German dialects and the rising impact of other idioms and/or high language varieties. In simple words, one can deduce the sociolinguistic history of Low German studying certain Slavonic idioms.

The richest material of this mutual evolution of Low German and Slavonic can be taken from the diachronic study of the Kashubian idiom (in Pomerania). Here we found different layers of the Low German loanwords that date minimally from the Middle Low German period up to the modern local Low German dialects. The specificity of the language situation in Pomerania also consists in the mediated influence of the East Central German and the rising influence of Standard German and Polish competing with the Low German loanwords, and in the role of the Low German heritage in the codification process of Kashubian. The main aim of the paper would be an attempt to sum up the main aspects and specifics of Low German-Slavonic contact from diachronic and areal perspective as attested in the Low German loan lexicon in the Slavonic languages.
Guus Kroonen  
Copenhagen University

Did Pitted Ware groups leave a linguistic trace in Germanic and Finnic?

The Pitted Ware was a marine oriented culture primarily based along the South Swedish and Northeast Danish coasts in the third millennium B.C.E. Though parts of the population had adopted aspects of animal husbandry and agricultural practices from the neighboring Funnel Beaker culture, isotopic and aDNA studies have shown a marked predominance of seal in the diet in eastern Sweden, which has given the Pitted Ware people the nickname “Inuit of the Baltic”. Though the Pitted Ware groups disappeared in the second half of the third millennium, we investigate the possibility that they left a linguistic trace in Germanic and Finnic. We then evaluate the ramifications of this linguistic contact for dating of the arrival of Germanic and Finnic speaking populations in the Baltic Sea area.

Juha Kuokkala  
University of Helsinki

Statistics and stratistics of the Germanic loanwords in Finnic

There is a considerable bulk of lexical items borrowed into Common Finnic from Proto-Germanic or language stages close to that (Pre-Germanic to North Germanic). The by far most comprehensive compilation of these old Germanic loanwords has been published in the Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen I–III (LÄGLOS, 1991–2012), which lists ca. 500 confirmed and another 500 possible loan items. The current presentation draws from a digitized corpus of LÄGLOS’s lexical material, showing some general statistics as well as more specific figures on sound correspondences between the Germanic and Finnic items. In addition, an experiment towards age stratification of the data based on the sound correspondences is demonstrated. Different correspondences, i.e. loan substitutions, of one sound (combination) can often be interpreted to reflect different borrowing times. Tagging the loan items with each alternating sound
correspondence value, we can sort the items to an approximate age order and further see how consistently each sound correspondence set follows an age pattern and find possible deviating items. The analysis is accompanied by visualizations of the patterns on time scale.

Denis Kuzmin
University of Helsinki / Karelian Research Centre

Karjalankielisen esikristillisen naisnimistön rekonstruointi

Концепт вûд «вода» в языковой картине мира мари

Концепт вûд «вода» в рамках этого исследования направлена на выявление отличительных черт национального характера. В словаре марийского языка это понятие толкуется так: 1) вода, 2) река, 3) ручей, 4) сок, 5) жидкость, рассло, настой, раствор, 6) перен. вода (речь без смысла) [СМЯ 1990, 318]. Данный концепт очень важен для понимания марийской культуры. Марийцев отличает бережное отношение к природе, так как она для них храм и место, где можно восстановить силы. Вода очень часто встречается в фольклоре народа мари: в загадках – Вûд оҕыл – вишкыде, лум оҕыл – ошо (шöр). Не вода, но жидкая, не снег, но белая (молоко); в пословицах – Вûдшү пүрэн ончыде, келгытшым палаш пеш йöсö. Чтобы узнать глубину воды, нужно в нее залезть. Вûдыш пүрыде ияш от тунем. Чтобы научиться плавать, надо залезть в воду; в песнях –

Изи вûдет йогалеш, Небольшая вода течёт
Тыгыде колет мобылдалеш Маленькие рыбки играют.

в запретах – Вûдыш шûведаш ойöр – сулык. Нельзя плеваться в воду – грех.

Для выявления языковой картины мира мари был проведен ассоциативный эксперимент. В нашем эксперименте вода ассоциируется у большинства респондентов с чистотой, рекой, родником. По логическим связям преобладает центральная реакция, т.е. связь между словами существует (18), периферическая (5), с точки зрения грамматической связи парадигматических реакций больше (15), чем синтагматической (8). Под парадигматическими реакциями испытуемых понимается связь в пределах одной грамматической категории. Например: вûд – памаш (родник), тазалык (здоровье), ошма (песок), отдых (каныш) и др. Судя по результатам вода – это источник чистоты, т.к. в древности и в настоящее время только после бани, т.е. после водных процедур, человек становился чистым и телом, и духом.
Assessing computational linguistic phylogeny with Uralic historical phonology

One of the biggest issues in using computational phylogenetic methods in historical linguistics is the role of linguistic data that are used in the analyses. As Pereltsvaig & Lewis (2015: 2) put it, especially in regard to Indo-European phylogeny as studied by Bouckaert et al. (2012), “incorrect and in some cases incoherent linguistic information is fed into complex equations, systematically corrupting the results; as the common adage puts it, ‘garbage in/garbage out’.”

In order to strengthen the use of linguistic data in computational phylogeny and to aid in developing the methodology for use in historical linguistics, in this study phonological innovations (sound changes) in shared vocabulary are used as data in phylogenetic analyses. This means that the phylogenetic algorithms are made to employ the same kind of information as in qualitative historical linguistics employing the comparative method. This allows for the optimal assessment of the reasons for differences between qualitative study and results from computational phylogeny better than using results from e.g. basic vocabulary phylogeny (for the Uralic languages, Syrjänen et al. 2013; Honkola et al. 2013; Lehtinen et al. 2014), which uses data that are not as central in established qualitative research.

The sound change data used in this study are based on the attested forms in 28 present-day Uralic languages as well as reconstructed forms in 8 intermediate protolanguages and in Proto-Uralic, corresponding to 58 reconstructed Proto-Uralic lexical roots. The amount of roots was restricted using criteria of phonological and semantic correspondences, distribution in different Uralic subgroups and the quantitative assessment
of their etymological stability based on reconstructed meaning using data from the World Loanword Database (Haspelmath & Tadmor 2009). This was done to obtain a dataset of manageable size and to maximally weed out possible effects of borrowing and chance correspondences from the data.

References


Anna Makarova  
Ural Federal University

Landscape terms in the substrate toponymy of Belozerye:  
Semantic-etymological research

The report examines landscape terms, evidenced in substrate toponymy of Belozerye in a proper toponymic use (mow Podša, swamp Čupa), and as part of attributive combinations (swamp Savinskaja Pohta). The research material is the data of field card indexes of Toponymic expedition of the Ural Federal University (Yekaterinburg) and Institute of Language, Literature and History, Karelian Research Centre RAS (Petrozavodsk), collected during field trips in 1961–2015. We also used the data of dialect dictionaries of the given area.

The analysis involved only place names, etymology of which caused a problem: the discrepancy between semantics in the source language and the object, denoted by the place name, and/or appellative semantics in Northern Russian dialects; the necessity to choose from several homonymous etymons of the different origin etc.

The study outlines some of the semantic patterns associated with the transition of words of certain semantic groups to the landscape toponymic vocabulary. In particular, there is a transition of weather vocabulary to the landscape terms: snow, frost → swampy place (kuftýř ‘shaky’, pohta ‘swamp’); of vocabulary related to the naming of shape: angle → bay (poča ‘the swamp, the old bed of the river’, čupa ‘swampy place, bay’). The report discusses alleged continuations of the Fennic and Saami etymons in modern Finnic languages and Northern Russian dialects, which enables better explicate semantic history of the studied lexemes.

This research was supported by the Russian Scientific Fund, project 17-18-01351 “Contact and Genetic Connections of North Russian Vocabulary and Onomastics”.
Basque etymological studies: Loanwords, roots and a new etymological dictionary

A new era in Basque diachronic linguistics started with Mitxelena. Based on Martinet’s work (1950), he extends the reconstruction of occlusives to the whole phonological system (1961–1977); this reconstruction is shaped by 1200 etymologies. He proposes a full reconstruction of the protobasque phonological system, basing his proposal mainly on loanwords from the latin-romance continuum.

This reconstruction has also served as a basis for the phonological reconstruction of the inherited lexicon: the intervocalic change $VlV > VrV$ is well documented in pairs such as lat. *angelus, gula* > basq. *aingeru* ‘angel’, *gura* ‘desire’. This offers us a better understanding of the inherited pairs *euskara* ‘Basque language’ / *euskaldun* ‘Basque speaker’, *sari* ‘payment’ / *saldu* ‘to sell’, for which original laterals are reconstructed: *euskala*, *sali*. But the information pulled from the adaptation of loanwords is nearly exhausted.

A new model for the reconstruction of Basque lexicon has been proposed by Lakarra (1995), based on the definition of the canonical root: proto-Basque roots are claimed to have a basic CVC structure. Subsequently, many new analyses of old known data have been made, and we have identified new roots and suffixes from the inherited lexicon: cf. the *bel* ‘black, dark’ root (cf. *beltz* ‘black’, *bele* ‘crow’, *arbel* ‘slate’, *orbel* ‘dead leaves’), the *din* ‘become’ root (cf. *jin* ‘come’ < *edin*), the *hor* ‘dog’ root (cf. *hortz* ‘tooth’), etc. At the same time, typologically well described grammaticalization paths allow us to identify the etymological source of suffixes such as -di (adject. suff.), which has its source on *din* ‘become’ (Dimmendaal 2011: 129).

The Royal Academy of the Basque language has promoted the elaboration of a new etymological dictionary of Basque, which puts together the latest advances made in the reconstruction of Basque lexicon.
Bibliography


*Niklas Metsäranta*
*University of Helsinki*

**Loans between branches in reconstruction**

Loan substitutions and loan words in general can be a helpful tool in trying to determine the sound value of a given reconstruction. It can also give us clues about the chronology of the changes. Researchers working on Proto-Permic vowel reconstructions for example have used evidence provided by loan words in their reconstructions (Lytkin 1964; Zhivlov 2014). Zhivlov for example posits that Udm *už* ‘stallion’, Komi *už* should be reconstructed as Proto-Permic *ož* partly because the word in question has been borrowed from Proto-Permic to Proto-Mari as *ožə* (Mari E *ožo*, W *ož*). Phonologically the proposal seems perfectly sensible, but does invite questions regarding its chronology. As Zhivlov himself points out, the shift *o > u* has occurred in the Volga-Kama linguistic area. Why don’t we then find *u* in Mari as well? We could probably assume that the word was borrowed into Proto-Mari after the shift *o > u* had already occurred, but this kind of reasoning is far from unproblematic. There are no constants in reconstruction, however, and all sounds are at all times liable to change. Several other examples from literature discussing Uralic reconstructions are considered in an attempt to explore the common problems associated with using loan words as evidence for reconstruction.
Balto-Finnic influence on a wide language background

The paper attempts to analyze etymologically the northern Russian vocabulary which was interpreted earlier as Baltic-Finnic origin. Particular emphasis in the work being done is on a wide coverage of the material, including East Slavic language data, given the importance of taking into account the various areas of foreign language exposure. We consider the probability of various etymologies of words found in dialects as well as nation-wide vocabulary. In some cases the Russian dialectal word has a wide areal distribution extending marginally to Ukrainian, probably as a result of Russian influence.

One of the names for a shoe is found in the Ukrainian, Belarusian and Russian languages, despite the fact that it is a Baltic-Finnic borrowing. It should be noted that in the Russian language the corresponding word in the XVIII century has been presented in a standard lexicon, cf. rus. *kenga*, commonly used in the plural *kengi* ‘winter calfskin or deerskin shoes with wool, which in underfur fur or bikes, similar to cats and worn for warmth over shoes and boots’. In the XIX century the word is also indicated as widespread: *kenga* ‘winter shoes, hemmed with fur inside or bike and worn over shoes or boots’. The tradition of quoting the word continues, but in the XX century the word already has a dialectal tone.

In etymological terms, the North Russian *kenji* has a long tradition of research which relates the material to Karelian-Vepsian sources: karel. *kengä* ‘shoes, boots’, karel. Tver. *kengä* ‘shoes, footwear (often worn on the shoe)’, Veps. *keng'ad* ‘shoes (any)’, Fin. *kenkä* ‘shoes’, Votic. *tšentša* ‘shoes’. The Livonian *käănga* ‘shoes’ was borrowed into Latvian *k'eng'e* ‘kind of shoes’, and the Finnish *kenkä* into modern Swedish *käng* ‘shoes’.

The Komi language data can hardly be interpreted as a Nenets borrowing, as it is done in «Brief etymological dictionary»: Komi *tenga* ‘felt boots for children’, with the Nenets. *t'enkdei p'iββa* ‘id.’. In our view, the direction of borrowing may be as follows: Balto-Finnic > dialectal Russian > Komi > Nenets. Moreover, the variability of [k] / [t] before front vowels is very characteristic for the Russian dialects.
Simonas Noreikis  
*University of Helsinki*

**Lithuanian hydronyms supposed to be Finno-Ugric**

In 1913 Jan Rozwadowski claimed that there might be hydronyms of Finno-Ugric origin in Lithuania. Later they were presented also by Kazimieras Būga and Tadeusz Lehr-Spławiński. Many etymologies by the latter one were rejected. Most such etymologies were suggested by Aleksandras Vanagas, and nowadays they are believed to be credible. The number of such hydronyms is 30–40.

The belief that in Lithuania there are Finno-Ugric substrate hydronyms is based on the fact that there are traces of the Comb Ceramics culture, which has been connected to Finno-Ugric speakers. Hydronyms which were difficult to etymologize were compared by Vanagas to proper names and appellatives in modern Finno-Ugric languages (mostly Finnic).

The aim of my research is to review the etymologies of the above presented hydronyms using modern methods. Firstly, the comparison has to be not to the modern but to the reconstructed forms in the supposed source languages. Secondly, the etymology of the suggested source word must be checked, as it might be more recent than the supposed contact. Thirdly, the types of toponym construction as well as semantics should be compared in the compared languages. Finally, all the hydronyms which are supposed to be Finno-Ugric have to be compared using an areal-typological method and checked if they can belong to one type which can be substrate of a certain Finno-Ugric group.

For example, river name *Robata* is equated to an Estonian word *raba* ‘swamp’. The claim is wrong, as *raba* is a Germanic loan from 3–8 century AD, while substrate contacts in Lithuania according to the hypotheses took place ca. 2000 BC.

Due to time restrictions I will present only 4–5 hydronyms. Using concrete examples I will show how I apply the methods described above.
Combining historical lexicology with archaeology: An example from the South Saami prehistory

In this presentation, I will consider how historical lexicology, and loanword studies in particular, can be combined with archaeological research. Historical lexicology is the area of linguistics with which we gain knowledge from the past. Another field of research that studies the prehistorical times is archaeology. Words, were they loans or inherited words, as well as archaeological remains tell us about the human environments, cultures and language contacts of the past. In Finland, linguists have searched confirmation to their hypothesis from the results of archaeology and vice versa, but the usage of theory and methodology of both sciences in the same research has been rare.

The purpose of my research is to find out how to fruitfully work with loanword material and archaeological material together. Advances and problems that arise with the combination of these two sciences will be carefully considered.

As an example material I will present the early Iron Age (CE 1–600) Proto-Scandinavian language contacts of the Saami. My point of view is in South Saami language and the geographical area of the South Saami people in Central Scandinavia. I will explain the method of cultural inventory of the Proto-Scandinavian loanwords. I will also take a look at the archaeological remains and artifacts in which contacts between the Saami and the Scandinavians can be seen. I will examine how the results of the analysis of these two materials contribute to our knowledge of the South Saami prehistory. Finally, I will ponder the pros and cons of combining archaeology and historical lexicology in my case study.
The state of the Finno-Permic lexicon

Most traditional family tree models of the Uralic language family have posited a Finno-Permic (FP) node, comprising the five westernmost unambiguous subgroups (Finnic, Samic, Mordvinic, Mari, Permic). Starting from the late 20th century, this view has been subject to extended critique: multiple studies have drawn attention to a lack of clear grammatical or phonological innovations that would define a distinct FP grouping, by now seemingly leaving the concept obsolete.

Major reference works on the lexical reconstruction of Proto-Uralic, assembled under the earlier assumption of a FP grouping, have however collected and isolated a substantial body of lexical material absent from the more eastern subgroups of Uralic. The present paper aims to firstly review this data in light of new developments in Uralic etymological research. Secondly, the data will be taken as a basis to discuss the following theoretical issues in lexical reconstruction:

1) the problem of distinguishing shared innovations from shared retentions;

2) the problem of distinguishing common innovations from parallel innovations;

3) the importance of control hypotheses in evaluating models of language subgrouping.
Ilona Rauhala  
University of Helsinki

The layers of loanwords in the adjectives of North Saami

This presentation concentrates on the loanwords in the adjective system in North Saami. The central question is, how the loanwords have been adapted to North Saami, and how they have affected the adjective attribute system.

The Saami languages have a complex adjective attribute system. It dates back to Proto-Saami, but it is assumed that it has been more regular than in the modern Saami languages. The adjective system in North Saami has been described e.g. by Sammallahti (1998), but the description is synchronic. It does not try to describe the development of the system. The examination of the oldest common Saami adjectives suggests that the choice of the attribute form has depended on the last syllable of the adjective (examples 1 and 2).

1) a. Biila lea ođas.  
   car.NOM is new.NOM  
   b. ođđa biila  
   new.ATTR car.NOM

2) a. Bárdni lea njoahci.  
   boy.NOM is slow.NOM  
   b. njoazes bárdni  
   slow.ATTR boy.NOM

Adjectives with suffixes have different attributive forms, such as adjectives in -at, which traditionally have an attributive form in -es (example 3).

3) a. Biila lea ruoksat.  
   car.NOM is red.NOM  
   b. rukses biila  
   red.ATTR car.NOM

In this presentation, I will try to describe how the loanwords have been adapted to the adjective system and whether the loanwords have been adapted differently in different stages of the language. I will divide the loanwords in North Saami adjectives according to the age of the loan and observe, how the adaptation mechanisms have changed.

References

Basic vocabulary and the phylogenetic approach to the study of Uralic language history

In this presentation we offer an overview of the basic vocabulary data with cognate (correlate) assessments we have collected for Uralic languages, and the Bayesian phylogenetic methods we are using to analyse these data for studying the history of the Uralic language family.

We will start by introducing our data, which consists of cognate sets taken from basic vocabulary in 26 Uralic languages. Cognate assessments are taken from earlier literature on Uralic historical linguistics and show which meanings share a common origin. We also discuss – and illustrate with examples – how the more and less stable subsets in the vocabulary affect the analyses and how the benefits of the phylogenetic approach are lost if the amount of data is too small.

Since the type of data we are using resembles binary data used in evolutionary biology it can be analysed with the well-established phylogenetic computational tools. We will show, how the computational approach offers new tools to study historical linguistics. This is especially interesting for the objective handling of large data sets. However, it is important to understand that using of these methods as well inferring the outcomes are not completely transparent tasks. We will discuss these issues to be taken into account in planning the analyses and in inferring the language family trees using Bayesian algorithms.

Our presentation aims specifically at clarifying some of the common misunderstandings regarding our approach to historical linguistics, which were recently also raised in a monograph on “The Indo-European Controversy” by Pereltsvaig & Lewis (2015). Although their critics are not entirely unjustified, we will describe how our approach, called language phylogenetics”, takes the advantage of the vast knowledge resulting from earlier basic research in qualitatively oriented historical-comparative linguistics, and converts it into quantitatively-oriented “applied historical linguistics”.
Parallels in noun and verb stem types in the Mari and Mordvin languages

There are parallels between noun and verb stem types in individual languages. These parallels might help for finding correlations between word stems in different subbranches of the Uralic language family.

A concatenational perspective indicates the Mari and Mordvin language groups differ in the number of stem types they have. Mari nouns and verbs come in two stem types, whereas Mordvin nouns and verbs can attest to even three. The Hill Mari (mrj) and Meadow or Eastern Mari (mhr) languages have two noun stem types: consonant-final, and vowel-final. Likewise the two conjugation types attest to a salient break between consonant and vowel stem types demonstrated in paradigmatic and derivational patterns.

(1.1) кол ‘fish’
(1.2) (mrj) понгы, (mhr) понго ‘mushroom’
(2.1) (mrj) возаш¹, возам: ваз; (mhr) возаш¹, возам: воз Imp.Sg2 ‘to fall’
(2.2) (mrj) возаш², возем: вазы; (mhr) возаш², возем: возы Imp.Sg2 ‘to write’

The division of Erzya (myv) and Moksha (mdf) noun stems into three groups is observed in a small and almost predictable group of vowel-stem
nouns that can lose their stem-final mid vowel in some instances: the indefinite nominative plural, for example. Analogously the verb stem types are also divided into vowel and consonant-final stem types as observed in combination with the indicative, first preterite Sg3 -ś. A third stem type is observed in the Erzya verbs in -OtOms, where the same -ś formant is preceded by an extra -t- element.

(3.1) кал ‘fish’
(3.2) (myv) пуло, пулонь: пулот Pl.Nom.Indef ; (mdf) пула, пулонь: пулот Pl.Nom.Indef ‘tail’
(3.3) (myv) панго, пангонь: пангт Pl.Nom.Indef ; (mdf) панга, пангонь: панкт Pl.Nom.Indef ‘mushroom’
(4.1) молемс: мольсь ‘to go’
(4.2) самс: сась ‘to arrive’
(4.3) (myv) сатомс: сатотсь ‘to be enough’

This paper also contemplates Mari-Mordvin correlations (mrj) пӈчк- ‘to cut’, (myv) печк- ‘to cut’ and (mhr) мұры- ‘to sing’; (mdf) мора- ‘to sing’.
Defining Baltic and Slavic in Finnic vocabulary

In my paper, I investigate the characteristics of Slavic and Baltic borrowings in Finnic languages. The point of outcome is that the so-called Baltic borrowings in Finnic probably derive from historical dialects that subsequently developed into Slavic. This is reflected in the fact that some toponymic dublets in Ingria (cf. *Laukā[njoki] ~ Luga*) pointing to a phonematically archaic Slavic in the presumed oldest contact zone of Finnic and Slavic.

The study of Slavic borrowings in Finnic has not utilized the Russian dialectological material that has been published actively beginning from the 1970s. Further, the new findings regarding the phonology of Novgorod birch barks, an early source for vernacular Slavic, for instance, the absence of the 2nd palatalization, has not been taken into account in the etymological analysis of the Finnic languages.

If it is assumed that many of the Slavicisms of the Finnic languages originally derive from Old Novgorod and Pskov vernaculars, one should probably reconsider the characteristics of some of the words traditionally considered as Baltic borrowings in Finnic. Some of them may derive from archaic Slavic dialects, or from the Baltic dialects that subsequently evolved into Slavic. In addition, some words may be proposed new Slavic etymologies.

In my paper, I will discuss the limits and characteristics of Slavic and Baltic layers of borrowings in the light of cases with the 2nd palatalization (*käärmė ‘snake’, *käämi ‘Weberspule’, *keltainen ‘yellow’) and diphthongs (*hauki ‘pike’, *rauta ‘iron’). Especially, I will stress the importance to consider the Russian dialectal material and the archaisms that can be found in Novgorod and Pskov vernaculars.
The notorious cruxes of Common Scandinavian umlaut and breaking: A metaphonic feature-based unified solution

In the transition from Proto- to Old Scandinavian target vowels in an initial main stressed syllable were assimilated into triggering vowels in a following syllable by means of front umlaut, rounding umlaut and by certain forms of breaking, and this in a manner that they frequently ended up as new vowel phonemes. Some 200 years ago the rules of Germanic umlaut were framed and outlined by Rasmus Rask and Jakob Grimm. To date, some very fundamental problems remain unsolved, including the relative chronologies and the mechanism of transmission, as well as the question of how, or indeed whether, the genesis of new phonemes depended on the weakening of umlaut triggers or their deletion. Particularly in Scandinavian, the distribution of umlaut and breaking in the lexicon persistently defies a good explanation.

In the presentation it is shown that by reconstructing one further pre-Germanic archaism and one chain shift in the Proto-Scandinavian vowel system, as well as a prominence system based on mora count, the most recalcitrant distribution of umlaut, namely when a trigger in a light syllable had followed a target in another light syllable, may be economically explained.

Finnic loanwords do not by and large reflect sub-phonemic umlaut together with reflexes of an unreduced syllable structure. In the presentation three etymons which appear to originate from the umlaut era are discussed: *olut, *rohkeda and *kari.
Revisiting the quantitative phylogeny of the Uralic languages

Phylogenetic tools have gained a foothold in recent historical linguistics over the last years, with applications that already reach further than creating quantitative topologies for languages. In addition, effort is put into fine-tuning existing methods and data, and re-evaluating the existing phylogenies. Here we present approaches with which we optimize and re-evaluate the quantitative Uralic phylogeny by using cognate coded basic and less basic vocabulary data. First, we take into account the heterogeneous stability of meanings in lexical data. We experiment this with both manual partitioning strategies, where the characters are divided into clusters with different rate of change using a priori knowledge (e.g. semantic category), and algorithmic partitioning strategies, where the characters are divided using heuristic techniques that assign stability ranks for each character. The second approach is to improve the divergence time estimates by analyzing the partitioned data and using improved evolutionary models designed for languages. Finally, we also examine the divergence structure of the Uralic family using a new tool, which allows for assessing whether the divergence event resembles polytomous or binary branching. When comparing different ways to partition the data, we found that algorithmic techniques were more successful than manually created partitioning approaches. The timed tree based on algorithmic partitioning and language-specific evolutionary models improved our earlier timing analyses. According to these analyses, Proto-Uralic started disintegrating ca. 5000 years ago (with error variation 6600–3700 YBP in the preliminary analyses), with a binary split separating Samoyedic from Finno-Ugrian. The Finno-Ugrian divergence, in turn, occurred either rapidly or polytomously some 4200 years ago (5400–3100), including the divergence of the Finno-Saami entity. Saami languages diverged from Finnic languages some 2800 (3800–2000) years ago.
The major purpose of this study is to investigate the semantic development of Amharic (an Ethio-semitic language) words. Attention is given to investigating factors that have triggered the changes, analyzing and classifying the findings based on the taxonomies of semantic change, and exploring possible ‘laws’ that govern the changes.

The semantic behavior of the words under study is compared in the older and the newer dictionaries to achieve the foregoing objectives. The results of the study show that a multiplicity of factors has led to change in the semantics of Amharic words. Various kinds of meaning changes are identified on the basis of the different factors and association mechanisms that underlie the words.

Words have changed their meanings pejoratively, aškar ‘boy or girl’ > ‘servant’; amelioratively, zega ‘beggar’ > ‘citizen’; metaphorically, ṏmmärtä ‘leaping’ > ’progress’; metonymically, c ’ènk ’ōllat ‘brain’ > ‘skull’. Meanings have also undergone specialization, ḍs’ ‘plant’ > ‘drug’; generalization, wäyzäro ‘queen’ > ‘a title for any married woman’; and semantic bleaching fit ‘face’ > temporal (past or future). In addition to these general tendencies, some specific tendencies are also seen in “person body-parts” andäbät ‘palate’ > ‘speech-producing organs’.

Metaphor is found to be by far the most pervasive force in many of these changes. This discussion of semantic change suggests that most of the meaning changes come under the overriding principle: concrete meaning becomes abstract.

The study had to confront two basic problems: the acute lack of an etymological dictionary of Amharic, and the absence of a systematic classification of the historical stages of the language. Scholars should give serious attention to both of these problems.
Outi Vesakoski, Mervi de Heer, Terhi Honkola, Unni-Päivä Leino, Jyri Lehtinen, Jenni Leppänen, Timo Rantanen, Kaj Syrjänen & Niklas Wahlberg
University of Turku etc.

Historical linguistics and evolutionary framework (Poster)

BEDLAN is a multidisciplinary research group focusing on the evolution of linguistic variation, including mechanisms of languages divergence and linguistic dispersion. This is done by applying frameworks and methods from evolutionary biology to different types of data sets: One subproject compares cross-linguistic data from the whole Uralic family, another focuses on historical Finnish dialects and third on current sociolinguistic change in speech. The applied data sets and methods allow for co-operating with international projects conducting similar studies on other language families. By comparing the results of our studies to 250 years of research in historical Uralistics, we can offer a comprehensive synthesis of the prehistory of the Uralic language family to international researchers of phylogenetic linguistics as well as to a transdisciplinary audience interested in human prehistory. Specifically, we aim at coupling our project with studies in Uralic archaeology and genetics carried out by other groups. Ultimately, we aim at drawing a transdisciplinary picture of the prehistory of Uralic speaking populations. One specific idea of BEDLAN, which we will present on our poster, is the study of the mechanisms behind linguistic divergence. This has been done mostly on the “microevolutionary level”: studying linguistic diversification with Finnish dialects. Until now, the team has allocated the resources in order to building solid data foundations to our studies as well as by testing various new methods to be adopted to traditional linguistic studies. Recently funded follow-up projects will expand the study to patterns and mechanisms of concurrent dialectal change (i.e. diachronic data on the urban Helsinki dialect) and to study how much old dialectal areas co-occur with cultural, administrative and biogeographical areas within Finland. In the latter approach, advantage is taken of methods provided by geographical information systems (GIS).