Santeri Junttila
Universität Greifswald
Ernst-Lohmeyer-Platz 3, DE-17487 Greifswald
santeri.junttila@iki.fi

Sampsa Holopainen
University of Helsinki
PL 24, FI-00014 Helsingin yliopisto
sampsa.holopainen@helsinki.fi

Juho Pystynen
University of Helsinki
PL 24, FI-00014 Helsingin yliopisto
juho.pystynen@helsinki.fi

DIGITAL ETYMOLOGICAL DICTIONARY OF THE OLDEST VOCABULARY OF FINNISH

This paper presents the new project *Digital etymological dictionary of the oldest vocabulary of Finnish* (University of Helsinki, funded by the Kone Foundation) and discusses the present state and challenges of the (especially digital) etymological resources of the Finnic languages and Uralic languages in general. It is also shown how crowdsourcing of etymology can work, and how the present platform could be used in the etymological lexicography of other languages and language families.

1. Introduction

In this paper, we will describe the aims and the current situation of the project *Digital etymological dictionary of the oldest vocabulary of Finnish*, funded by the Kone Foundation and carried out in the department of Finnish, Finno-Ugrian and Scandinavian studies of the University of Helsinki from 2017 to 2020. The project is led by Santeri Junttila, and the project team consists of Petri Kallio, Sampsa Holopainen, Juho Pystynen, and Juha Kuokkala.

In the paper, we will show how our new dictionary will advance the etymological research of Finnic and other Uralic languages.

2. Background

2.1. Finnish and Finnic

Finnic is a well-defined branch of the Uralic (Finno-Ugric) language family. The Finnic group of languages consists of Finnish (including Meänkieli and Kven, which have separate literary standards in Sweden and Norway), the closely related Karelian (including Olonetsian), Ludian, Veps, Izhorian (Ingrian), Votic, Estonian, South Estonian (which is a separate language and not very closely related to standard Estonian, despite the misleading name; the name *Ugala* is also used sometimes for South Estonian) and the nearly extinct Livonian.

The common ancestor of the Finnic languages is (Late) Proto-Finnic, which started to split-up during the first half of the first millennium (Kallio 2014: 163–165). Proto-Finnic was in intensive contact with Baltic (see Junttila 2012, 2015) and Germanic (LäGLoS; Kallio 2012, 2015), and according to Kallio (2014: 164) had recently come in contact with early Slavic, with which many Finnic languages have later been in intensive contact.

2.2. Proto-Finnic etymology: The current situation

Despite the long history of Finnic etymological research, the existing standard sources of Finnic etymology are not well up to date. All the existing dictionaries, covering the Finnish and Estonian lexicon, also include various methodological shortcomings and misleading views. The oldest, SKES (*Suomen kielen etymologinen sanakirja*) is outdated already due to its age (published 1955 to 1981). It is also natural that its successor SSA (*Suomen sanojen alkuperä*; published 1992 to 2000) cannot include the most recent developments of Finnic etymology and it has also been criticized already at the time of its publication (see Nikkilä 2002) for the extensive use of onomatopoetic and affective explanations. SKES and SSA also lack proto-language reconstructions (except for very few entries

in SSA). NES (*Nykysuomen etymologinen sanakirja*) gives a good overview of Finnish etymology to the general public, but it covers only a part of the oldest Finnish vocabulary, and new developments in the field of Uralic etymology after the early 2000s are not included. The only complete etymological dictionary of Estonian, EES (*Eesti etümoloogia sõnaraamat*) includes up-to-date information about individual etymologies to varying degrees, but it completely lacks references about the individual etymologies.

Various problems concern most etymological dictionaries of other Uralic languages as well. Bereczki's (2013) etymological dictionary of Mari mostly follows the obsolete reconstructions of the UEW (*Uralisches etymologisches Wörterbuch*) – for general criticism of the UEW see below. The Hungarian etymological dictionaries such as EWUNg (*Etymologisches Wörterbuch des Ungarischen*) have a rather conservative approach and generally avoid referring to more recent works on Uralic phonology and etymology, such as Sammallahti (1988). KESK (*Краткий этимологический словарь коми языка*; Lytkin and Gulyaev 1999), the only existing etymological dictionary of Komi (and the Permic languages in general) does not include the most recent views about the complicated phonological developments of the Permic languages. DEWOS (*Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*), the only existing etymological dictionary of Khanty or the Ob-Ugric languages in general mostly reflects the etymological knowledge of the 1960s–1970s. For many languages or language groups, such as Mansi or Samoyed, no etymological dictionaries exist at all.

Also, the handbooks of Uralic studies (such as Collinder 1957 and Lytkin et al. 1974) mostly reflect outdated information concerning etymology; an important exception is the handbook *The Uralic languages* (Sinor (ed.) 1988) which includes the ground-breaking article of Sammallahti (1988) that includes a list of Uralic vocabulary based on more strict criteria than that of the UEW.

Although all the existing dictionaries have various problems, many papers from the 1990s onwards have taken the study of Uralic etymology forward: inherited Uralic lexicon has been discussed by such researchers as Helimski (see Helimski 2000 for a compilation of his articles published and various journals and books), and the various loanword layers have been discussed especially by Jorma Koivulehto (see the complication volumes Koivulehto 1999 and 2016): the

bibliographical information of all the literature published after the SSA will be included in our database of research history within the dictionary.

Especially the recent years have seen a resurgence of new etymological works, which deal with the re-evaluation of inherited Uralic vocabulary as well as the various Indo-European loanwords. Major works include Aikio (2009) on Saami loanwords in Finnic and Karelian, *Lexikon der älteren germanischen Lehnwörter der ostseefinnischen Sprachen* (Kylstra et al. 1991–2012) on Germanic loanwords in Finnic, Junttila (2015) on Baltic loanwords in Finnic and Holopainen (2019) on Indo-Iranian loans in the Uralic languages. Smaller studies on Uralic etymology continue to be published as well; a recent compilation volume is Holopainen and Saarikivi (eds.) 2018, and several articles are published in journals such as *Linguistica Uralica* (see for example Aikio 2014a, O'Rourke 2016) and *Journal de la Société-Finno-Ougrienne* (see for example Metsäranta 2017).

2.3. Existing online works

The Uralic etymological database URALONET (*http://uralonet.nytud.hu*) contains the material published in *Uralisches Etymologisches Wörterbuch* (Rédei et al. 1986–1991), the latest general etymological dictionary of the Uralic family, which provides important grounding for Finnic etymology. The database is practical and easy to use (it is also available in three languages, English, Hungarian, and German), but the material has been copied directly from the UEW, and is therefore outdated due to not taking into account research published after the 1980s. This results in two major problems. First, many of the UEW's etymologies have been in later research either refuted or substantially amended (with the addition of new cognates in lesser studies branches, like Samoyed). Second, the reconstructions in the UEW reflect outdated and unsystematic views on Proto-Uralic phonology. Even at the time of publication, the UEW contained views that were not shared by many scholars of Uralic etymology especially in Finland (see Häkkinen 1987; Kulonen 1988). For more general criticism of the UEW, see Ponaryadov (2012); Zhivlov (2014: 113–114). A third problem, though shared by almost all existing works, is the absence of intermediate reconstructions for e.g. Proto-Finnic or Proto-Samoyedic.

Álgu (*http://kaino.kotus.fi/algu*) – the Saami etymological database (hosted by the Institute for the Languages of Finland) is a useful collection of earlier references, but contains no original research. The database lists competing explanations and etymologies, and it can be very difficult to use for non-specialists. Among other sources, *Álgu* contains the material of Lehtiranta's *Yhteissaamelainen sanasto* (*Common Saami vocabulary*), which can be now also accessed online on the website of the Finno-Ugrian Society (*http://www.sgr.fi/fi/items/show/165*).

Eesti etümoloogiasõnaraamat (EES) has been published also as an online edition (*http://www.eki.ee/dict/ety*), containing the material published in the print version. The database is technically easy to use but lacks references to individual etymologies (as does the original printed dictionary), which seriously hampers its use as a research tool.

The Hungarian etymological dictionary *Etimológiai szótár: Magyar szavak és toldalékok eredete* (Zaicz 2006) is also available as an online edition (*http://www.tankonyvtar.hu/hu/tartalom/tinta/TAMOP-4_2_5-09_Etimologiai_szotar/adatok.html*). However, this is not a database and only consists of a scanned, non-machine-readable version of the printed book, with poor searchability (the book can be downloaded as a machine-readable PDF, however). The work likewise has no individual references.

3. Aims

3.1. Reconstruction of the common Finnic lexicon

Several thousands of etymological comparisons between the Finnic languages have already been compiled in the earlier Finnic etymological dictionaries discussed above. However, as these works have been etymological dictionaries focusing on a single language (either Finnish or Estonian), they have included also words limited only to a small part of the Finnic languages, as well as numerous words exclusive to Finnish or Estonian, whenever anything is known of their etymology. In other words, strictly Proto-Finnic etymology has so far remained only second fiddle to Finnish or Estonian etymology. This task is taken up as the primary aim of this project.

A notable achievement in Proto-Finnic etymology has been recently achieved in *Yhteisitämerensuomalainen sanasto* (*Common Finnic Vocabulary*) by project member Petri Kallio. The work compiles ca. 2000 etymologies covering the oldest vocabulary of Finnish and the other Finnic languages and provides phonological reconstructions for them. At the current stage, seven language varieties are covered independently: Finnish, Karelian, Veps, Votic, Estonian, South Estonian, and Livonian, the language varieties with either a standard literary form or in the case of Karelian and Votic, an interdialectal scholarly standard, employed for lemma forms in the major dialect dictionaries *Karjalan kielen sanakirja* (*The dictionary of the Karelian language*) and *Vadja keele sõnaraamat* (*The dictionary of the Votic language*). The selection of vocabulary is based on regular etymological correspondence, as identified already in earlier research, but for the first time also on distribution. Etymologies are accepted as Proto-Finnic if a word has reflexes both in at least one of the northernmost (Finnish / Karelian / Veps) and at least one of the southernmost Finnic languages (South Estonian / Livonian), or if cognates in a related Uralic language outside of Finnic are known. Following the earlier research tradition, generally, only underived root words are compared, though derivatives are accepted as reflexes where a root word has been lost.

The *Common Finnic Vocabulary* is planned for eventual monograph publication outside of the current project but its first release has taken place online during the project, forming its base. The online publication can be found at *http://sanat.csc.fi/wiki/Luokka:Yhteissuomalainen_sanasto*.

3.2. Loanword coverage

Native Finnic and Uralic etymological research is complemented by loanword research, which has an extensive tradition within Uralic studies. The most important loanword layers in the Finnic languages are those from Germanic, Baltic, and Slavic, but also older Indo-Iranian and Indo-European loans have been

studied in-depth (Koivulehto 1991, 1999, 2016[1]; Holopainen 2019). Loanword research can allow identifying even words of limited distribution as quite old: for example, loanwords from Indo-Iranian already date to Pre-Proto-Finnic. The majority of these loanwords have a widespread distribution in the Finnic languages, and such words are included in the project already through the *Common Finnic Vocabulary*. However, a review of the etymological research literature has revealed that the number of early loanwords that have been lost in several Finnic languages (which could not be reconstructed for Proto-Finnic from the Finnic evidence alone) is also substantial. Altogether an additional 300–400 words can be added to the etymological dictionary because they have known loan etymologies that can be dated to Proto-Finnic or earlier.

3.3. Updated historical phonology

The currently ongoing wave of research in Finnic and Uralic etymology is closely linked to ongoing advances in the study of historical phonology. Traditionally, Finnish had played the role of a "key language" in the phonological reconstruction of Proto-Uralic, which had led to biased reconstructions (see Kallio 2006: 163–64, 172 for criticism), but advances in the study of the related Uralic languages have led to less Finno-centric reconstructions of Proto-Uralic and improved understanding of the phonological history of Finnic. Sound changes separating Proto-Uralic from Proto-Finnic and their chronology have been extensively discussed by Kallio (2007, 2012, 2014) as well as by Pystynen (2018), and the phonological reconstruction of Proto-Uralic (especially of Proto-Uralic vocalism) has seen recent improvements by Aikio (2012, 2014b, 2015) and Zhivlov (2014). These developments are not reflected in earlier dictionaries, and our project is the first etymological dictionary of Finnish/Finnic that includes up-to-

---

[1]   Regarding the earliest possible loanwords from Indo-European to Proto-Uralic, there are various differing views, some (for example Koivulehto 1991) arguing that the earliest loanwords have been borrowed from Proto-Indo-European already, and also a layer of North-West-Indo-European loanwords has been suggested (Koivulehto 1999: 12, 341–358; Junttila 2015: 242–243). There are also more critical views concerning the existence of a Proto-Indo-European loanword layer (see especially Simon [in press]), and our dictionary contributes to this ongoing discussion. Possible external connections of the Uralic family (such as Nostratic or Indo-Uralic) are not discussed in the dictionary, as these hypotheses are both beyond the scope of Finnic etymology and generally not very convincing (see Koivulehto 1999: 295–306; Campbell and Poser 2008: 243–246).

date reconstructions of Proto-Uralic.

The Proto-Finnic situation has remained more stable. The reconstructions in *Common Finnic Vocabulary* still closely follow earlier results achieved in Finnic historical phonology, a lion's share of them dating back already to the foundational work of Setälä (1891). The most important recent updates to Finnic phonological reconstruction followed in the work are outlined in Kallio (2007, 2014). More important, however, is the fact that no earlier etymological source has presented Proto-Finnic reconstructions systematically. Scholars familiar with the established results in Finnic historical phonology have for long been able to work out the proper reconstruction from a set of cognates if needed, but for all others, this information has remained unavailable. Our project thus remedies a long-standing omission by systematically presenting Proto-Finnic reconstructions for all etymologies included in the dictionary. This also provides material for potential further research on Finnic historical phonology, by revealing a small number of remaining phonological irregularities that have not been systematically treated in earlier research.

3.4. Coverage of research history

Research has not always settled on clearly established etymologies for Finnic etyma. Often numerous different pre-Finnic etymologies have been proposed, and in such cases, no unique loan or native origin beyond Proto-Finnic can be defined. Instead of favoring one proposal over the other, the project aims for analytic coverage of the research literature, summarizing different proposals, their refinements, and their critiques. To aid in dealing with these relatively common cases, an extensive bibliographic database of Finnic etymological research is also being prepared as a part of the dictionary project.

Among earlier etymological dictionaries, SSA already provides a good analytic picture of research history up to the 1990s/2000s, with detailed bibliographical references listed at the end of every etymological entry. Our project will build on SSA's research history entries, filling possible gaps, and providing all the research history from the 1990s onwards.

4. Structure

4.1. Basic divisions

In its basic structure, the project is divided into two sections, providing two distinct ways of accessing etymological information. A database of etymological comparisons covers the full data – etymological and bibliographic – for scholarly use, while a popular dictionary summarizes the current knowledge of the etymologies of the Finnish words covered by the project. The popular dictionary will be based on the database: cognate relationships entered into the database can be automatically retrieved, displayed, and formatted also elsewhere within the project. Currently, the database presents the full etymological corpus and a small part of the bibliographic data. The work on the articles in the popular dictionary has also begun in early 2020.

4.2. Technical description

The project is technically implemented on Semantic Mediawiki, and hosted by the Language Bank of Finland on the open data wiki service *http://sanat.csc.fi*. The environment provides a suitable combination of database and publication features: collaboration and formatting tools are present natively, while also custom semantic templates and lists can be constructed to provide flexible and customizable ways for presenting the etymological data. If required, custom search and edit forms can be also set up to aid navigation for specialists. The popular dictionary and the etymological database comprise their own sections (namespaces) within the wiki platform, allowing formatting standards and contribution permissions to be set up independently for both.

Database structure allows us to arrange comparative Finnic etymologies in a more balanced form. Previous monolingual etymological dictionaries have gathered cognates from elsewhere in Finnic only under a lemma form from either Finnish or Estonian. If one does not already know the Finnish or Estonian cognate, this presents difficulties for finding cognates for words in smaller Finnic languages. Words not found in the focus language have also been excluded from previous etymological dictionaries. Within the database, both problems are solved by cross-searchability by any language variety. Ultimately, every word

covered by the dictionary will have its own entry in the database, linked to its cognates elsewhere in Finnic and to its Proto-Finnic origin both by human-readable hyperlinks and machine-readable ancestor/descendant relations.

Cognate relationships between the seven modern language varieties are however not encoded individually. Instead, Proto-Finnic serves as a "hub" variety. It is these entries, currently already available online, that could be thought of as the database basic lemmas. Cognate relationships can be automatically rederived or displayed by querying the Proto-Finnic reconstruction attached to an attested word, following by querying its descendant in a wanted language. Clearly established loan and native Uralic etymologies can be eventually treated similarly, through defining further etymological relationships.

The detailed technical implementation of the etymological database has been already mostly designed but remains under coding work, and will not be covered in detail in this paper.

4.3. Popular dictionary

The popular etymological dictionary follows a more traditional model, close to NES, SSA, and SKES. The lemmas of this section are the modern Finnish words treated by the project. Each word will be given, on the basis of information gathered in the database, a streamlined explanation in Finnish of what is known of their etymology, and a list of their cognates in the other Finnic languages; each entry also provides links into the database for further information. A list of the currently written article drafts (subject to further editing and formatting) can be found at *http://sanat.csc.fi/wiki/Luokka:Yleistajuiset_etymologiasivut.*

5. Future goals

5.1. Enabling collaboration

As shown above, etymology is a constantly progressing discipline, which creates a periodic need to update printed reference works. Our project, however, does not only present an up-to-date picture but also aims to remove altogether the need for repeated re-edition work, as the online database can be easily updated

to reflect new research results. We hope for the project to remain sufficiently active in the future to "close the loop", allowing general and particular etymological research to build on one another without unnecessary delay.

This goal will still require the contribution of work. To help with this, we aim for the project membership to be open to collaboration and expansion. Moreover, as the project is hosted in a wiki environment, also **crowdsourcing** is already possible: even scholars unaffiliated with the project are welcomed to comment on and improve the etymologies presented within the project, as much or as little as they see fit. An initial workshop introducing scholars of Finnic etymology to the project and ways to contribute to it has been held in the spring of 2019, and future workshops are planned as well.

5.2. Expansion possibilities

Like all Uralic languages, the Finnic languages feature complex **word derivation**, in many layers of different age (see e.g. Hakulinen 1979 for an overview of the historical background of word derivation in Finnish). Derivatives remain mostly out of the scope of the current project. Our data model will, however, easily allow the presentation of derivational relationships if desired, both within a single language and with the further etymological comparison between languages.

The widest possibilities of expansion, however, are entirely outside the field of Finnic etymology. The flexible wiki-based etymological database can be in principle extended to cover **any other etymological data**. Etymological research on other Uralic languages will be a natural direction of collaboration in the future, as they still share a common root in Proto-Uralic and common research literature. Etymological data of any important contact language groups would also have natural benefits to integration on a common platform. These, in turn, will have their own relatives and contact languages. With luck, our project could in time end up pioneering even a much wider trend of online digital infrastructure for etymological research.

References

AIKIO, ANTE. 2009. *The Saami Loanwords in Finnish and Karelian*. PhD thesis. University of Oulu. Oulu. 386 pages. http://www.academia.edu/4813912/The_Saami_Loanwords_in_Finnish_and_Karelian (accessed 29 June 2020).

AIKIO, ANTE. 2012. On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic *x. *Per Urales ad Orientem: Iter polyphonicum multilingue. Festskrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012*. Eds. Hyytiäinen, Tiina; Saarikivi, Janne; Sandman, Erika. Suomalais-Ugrilainen Seura. Helsinki. 227–250.

AIKIO, ANTE. 2014a. Studies in Uralic etymology I: Saami etymologies. *Linguistica Uralica* XLIX. 161–174. dx.doi.org/10.3176/lu.2013.3.01.

AIKIO, ANTE. 2014b. On the reconstruction of Proto-Mari vocalism. *Journal of Language Relationship* 11. 125–157.

AIKIO, ANTE. 2015. The Finnic 'secondary e-stems' and Proto-Uralic vocalism. *Suomalais-Ugrilaisen Seuran Aikakauskirja / Journal de la Société Finno-Ougrienne* 95. 25–66. doi.org/10.33340/susa.82642.

Álgu – sámegielaid etymologalaš diehtovuođđu [The Etymological database of the Saami languages]. 2006. Institute for the Languages of Finland. http://kaino.kotus.fi/algu (accessed 29 June 2020).

BERECZKI, GÁBOR. 2013. *Etymologisches Wörterbuch des Tscheremissischen* (*Mari*)*: Der einheimische Wortschatz (Nach dem Tode des Verfassers herausgegeben von Klára Agyagási und Eberhart Winkler)*. Harrassowitz Verlag. Wiesbaden.

CAMPBELL, LYLE; POSER, WILLIAM J. 2008. *Language Classification: History and Method*. Cambridge University Press. Cambridge.

COLLINDER, BJÖRN. 1957. *Survey of the Uralic languages*. Almqvist & Wiksell. Stockholm.

Eesti etümoloogiasõnaraamat [Estonian etymological dictionary]. 2012. Eds. Metsmägi, Iris; Sedrik, Meeli; Soosaar, Sven-Erik. Eesti Keele Sihtasutus. Tallinn. http://www.eki.ee/dict/ety (accessed 29 June 2020).

Etymologisches Wörterbuch des Ungarischen Sprache. 1992–1997. Ed. Benkő, Loránd. Akadémiai Kiadó. Budapest.

HAKULINEN, LAURI. 1979. *Suomen kielen rakenne ja kehitys [The structure and development of Finnish]*. Otava. Helsinki.

HÄKKINEN, KAISA. 1987. Uralilaisen kielikunnan vanhimmat sanat [The oldest words of the Uralic language family (Review of UEW)]. *Sananjalka* 29. 232–240.

HÄKKINEN, KAISA. 2004. *Nykysuomen etymologinen sanakirja [Etymological dictionary of modern Finnish]*. WSOY. Juva.

HOLOPAINEN, SAMPSA; SAARIKIVI, JANNE (eds.). 2018. *Περὶ ὀρθότητος ἐτύμον. Uusiutuva uralilainen etymologia*. Suomalais-Ugrilainen Seura. Helsinki.

HOLOPAINEN, SAMPSA. 2019. *Indo-Iranian borrowings in Uralic. Critical overview of sound-substitutions and distribution criterion*. PhD dissertation. University of Helsinki. Helsinki. 403 pages. http://urn.fi/URN:ISBN:978-951-51-5729-4 (accessed 29 June 2020).

JUNTTILA, SANTERI. 2012. The prehistoric context of the oldest contacts between Baltic and Finnic languages. *A linguistic map of prehistoric Northern Europe*. Eds. Grünthal, Riho; Kallio, Petri. Suomalais-Ugrilainen Seura. Helsinki. 261–296.

JUNTTILA, SANTERI. 2015. *Tiedon kumuloituminen ja trendit lainasanatutkimuksessa: Kantasuomen balttilaislainojen tutkimushistoria [The cumulation of knowledge and the trends in loanword research: Research history of the Baltic loanwords in Proto-Finnic.* PhD dissertation. University of Helsinki. Helsinki. 291 pages. http://urn.fi/URN:ISBN:978-951-51-1842-4 (accessed 29 June 2020).

KALLIO, PETRI. 2006. Suomen kantakielten absoluuttista kronologiaa [Absolute chronology of the protolanguages of Finnish]. *Virittäjä* 110. 2–25.

KALLIO, PETRI. 2007. Kantasuomen konsonanttihistoriaa [Proto-Finnic consonant history]. *Sámit, sánit, sátnehámit: Riepmočala Pekka Sammallahyii miessemanu 21 beaive 2007*. Eds. Ylikoski, Jussi; Aikio, Ante. Suomalais-Ugrilainen Seura. Helsinki. 229–249.

KALLIO, PETRI. 2012. The non-initial-syllable vowel reductions from Proto-Uralic to Proto-Finnic. *Per Urales ad Orientem: Iter polyphonicum multilingue. Festskrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012*. Eds. Hyytiäinen, Tiina; Saarikivi, Janne; Sandman, Erika. Suomalais-Ugrilainen Seura. Helsinki. 163–175.

KALLIO, PETRI. 2014. The Diversification of Proto-Finnic. *Fibula, Fabula, Fact: The Viking Age in Finland*. Eds. Ahola, Joonas; Clive Tolley, Frog. Suomalaisen Kirjallisuuden Seura. Helsinki. 155–168. dx.doi.org/10.21435/sfh.18.

KALLIO, PETRI. 2015. The Stratigraphy of the Germanic Loanwords in Finnic. *Early Germanic Languages in Contact*. Eds. Askedal, John Ole; Frede Nielsen, Hans. Benjamins. Amsterdam – Philadelphia. 23–38.

Karjalan kielen sanakirja [The dictionary of the Karelian language]. 2009 Eds. Virtaranta, Pertti et al. Suomalais-Ugrilainen Seura. Helsinki. http://kaino.kotus.fi/cgi-bin/kks/karjala.cgi (accessed 29 June 2020).

KOIVULEHTO, JORMA. 1991. *Uralische Evidenz für die Laryngaltheorie*. Verlag der Österreichischen Akademie der Wissenschaften. Wien.

KOIVULEHTO, JORMA. 1999. *Verba Mutuata: Quae vestigia antiquissimi cum Germanis aliisque Indo-Europaeis contactus in linguis Fennicis reliquerint.* Ed. Ruppel, Klaas Ph. Suomalais-Ugrilainen Seura. Helsinki. 1–16.

KOIVULEHTO, JORMA. 2016. *Verba vagantur. Jorma Koivulehto in memoriam*. Eds. Holopainen, Sampsa; Kallio, Petri; Saarikivi, Janne. Suomalais-Ugrilainen Seura. Helsinki.

Kulonen, Ulla-Maija. 1988. Uralilaiset etymologiat tulevat [The Uralic etymologies are coming (Review of UEW)]. *Virittäjä* 92. 283–92.

Lexikon der älteren germanischen Lehnwörter in den Ostseefinnischen Sprachen I–III. 1991–2012. Eds. Kylstra, A.D. et al. Rodopi. Amsterdam.

lytkin et al. 1974. = Лыткин, В. И. и др. 1974. *Основы финно-угорского языкознания I (вопросы происхождения и развития финно-угорских языков)*. Наука. Москва. [Lytkin, V. I. i dr. 1974. Osnovy finno-ugorskogo jazykoznanija I (voprosy proishoždenija i razvitija finno-ugorskih jazykov). Nauka. Moskva.]

Lytkin and Gulyayev 1999. = Лыткин, В. И.; Гуляев, Е. С. 1999. *Краткий этимологический словарь коми языка [Concise etymological dictionary of the Komi language]*. Коми книжное издательство. Сыктывкар. [Lytkin, V. I.; Guljajev, E. S. 1999. *Kratkij ètimologičeskij slovarʹ komi jazyka [Concise etymological dictionary of the Komi language]*. Komi knižnoe izdatelʹstvo. Syktyvkar.]

Metsäranta, Niklas. 2017. Päivitettyjä permiläisiä etymologioita ja rinnastuksia [Updated Permic etymologies and cognates]. *Suomalais-Ugrilaisen Seuran Aikakauskirja / Journal de la Société Finno-Ougrienne* 96. 211–241. doi.org/10.33340/susa.70228.

Nikkilä, Osmo. 2002. Suomen sanojen alkuperät päätökseen [Origins of the Finnish words finished (Review of SSA)]. *Virittäjä* 106. 124–36.

O'rourke, Patrick. 2016. Comments on Proto-Uralic Etymology: Derivations and Lexemes. *Linguistica Uralica* LII. 241–246. doi.org/10.3176/lu.2016.4.01.

Ponaryadov 2012. = Понарядов, В. В. 2012. Насколько достоверный современные реконструкции прафинно-пермского вокализма [How reliable are the modern reconstructions of Proto-Finno-Permic vocalism]?. *Вопросы финно-угорской филологии* 3/8. 78–91. [Ponarjadov, V. V. 2012. Naskolʹko dostovernyj sovremennye rekonstrukcii prafinno-permskogo vokalizma [How reliable are the modern reconstructions of Proto-Finno-Permic vocalism]?. *Voprosy finno-ugorskoj filologii* 3/8. 78–91.]

Pystynen, Juho. 2018. *Itämerensuomen pitkien vokaalien alkuperä [The origin of the long vowels of Finnic]*. Master's thesis. University of Helsinki. Helsinki. 116 pages. http://urn.fi/URN:NBN:fi:hulib-201806132569 (29 June 2020).

Rédei, Károly et al. 1988. *Uralisches Etymologisches Wörterbuch*. Harrassowitz. Wiesbaden.

Sammallahti, Pekka. 1988. Historical phonology of the Uralic languages with special reference to Samoyed, Ugric and Permic. *The Uralic languages: description, history and foreign influences*. Ed. Sinor, Denis. E. J. Brill. Leiden. 478–554.

Setälä, Emil Nestor. 1891. *Yhteissuomalainen äännehistoria [Common Finnic historical phonology]*. Suomalaisen Kirjallisuuden Seura. Helsinki.

Simon, Zsolt. [in press]. Urindogermanische Lehnwörter in den uralischen und finno-ugrischen Grundsprachen. Eine Fata Morgana?. *Indogermanische Forschungen* 125.

STEINITZ, WOLFGANG. 1966–1993. *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. Akademie Verlag. Berlin.

Suomen kielen etymologinen sanakirja [Etymological dictionary of Finnish] (1–7). 1955–1981. Eds. Toivonen, Y. H.; Itkonen, Erkki. Suomalais-Ugrilainen Seura. Helsinki.

Suomen sanojen alkuperä: Etymologinen sanakirja [The origin of Finnish words. Etymological dictionary] (I–III). 1992–2000. Eds. Itkonen, Erkki; Kulonen, Ulla-Maja. Kotimaisten kielten tutkimuskeskus & Suomalaisen Kirjallisuuden Seura. Helsinki.

Uralonet database. 2010. Research Institute for Linguistics, Hungarian Academy of Sciences. http://uralonet.nytud.hu (accessed 29 June 2020).

Vadja keele sõnaraamat [The dictionary of the Votic language]. 2013. Eds. Grünberg, Silja et al. Eesti Keele Instituut. Tallinn. http://www.eki.ee/dict/vadja (accessed 29 June 2020).

ZAICZ, GÁBOR. 2006. *Etimológiai szótár. Magyar szavak és toldalékok eredete [Etymological dictionary. The origin of Hungarian words and affixes]*. Tinta Könyvkiadó. Budapest. http://www.tankonyvtar.hu/hu/tartalom/tinta/TAMOP-4_2_5-09_Etimologiai_szotar/adatok.html (accessed 29 June 2020).

ZHIVLOV, MIKHAIL. 2014. Studies in Uralic vocalism III. *Journal of Language Relationship* 12. 113–148.

Digitalni etimološki rječnik najstarijega finskog leksika

Sažetak

U radu se predstavlja novi projekt *Digitalni etimološki rječnik najstarijega finskog leksika* (Sveučilište u Helsinkiju, financira ga zaklada Kone) i prikazuje trenutačno stanje i izazovi (posebno u vezi s digitalizacijom) koji stoje pred etimološkim izvorima za finske i uralske jezike općenito. Također se pokazuje kako se masovna podrška može primijeniti u etimologiji i kako prikazana platforma može biti upotrijebljena u etimološkoj leksikografiji drugih jezika i jezičnih porodica.

Keywords: etymology, Finnish, Finnic languages, Uralic languages, digital humanities
Ključne riječi: etimologija, finski, finski jezici, uralski jezici, digitalna humanistika