

Faculty of Arts  
University of Helsinki

# **INDO-IRANIAN BORROWINGS IN URALIC**

CRITICAL OVERVIEW OF THE SOUND SUBSTITUTIONS AND  
DISTRIBUTION CRITERION

**Sampsa Holopainen**

DOCTORAL DISSERTATION

To be presented for public discussion with the permission of the Faculty of Arts of the University of Helsinki, in Auditorium PII, Porthania, on the 14th of December, 2019 at 12 o'clock.

Helsinki 2019

ISBN 978-951-51-5728-7 (pbk.)  
ISBN 978-951-51-5729-4 (PDF)

Unigrafia  
Helsinki 2019

# TIIVISTELMÄ

Väitöskirjan aiheena ovat uralilaisten (suomalais-ugrialaisten) kielten indoiranilaisperäiset lainasanat. Työssä arvioidaan kriittisesti aiemmissa lähteissä esitetyt etymologiat ja kommentoidaan niiden pohjalta lainoissa esiintyviä äännesubstituutioita ja levikkiä. Aineiston muodostavat eri uralilaisissa kielissä esiintyvät lainasanat, joille on aiemmassa tutkimuksessa esitetty kantaindoiranilaista tai kantairanilaista, tai muuta varhaista iranilaista laina-alkuperää.

Menetelmänä sovelletaan lainasanatutkimuksen yleisesti hyväksytyjä kriteereitä. Etymologioita arvioidaan äänneasun, levikin ja merkityksen perusteella. Työssä on otettu huomioon sekä uralilaisen että indoiranilaisen kielikunnan äännehistorian uusin tutkimus, ja monet työn tuloksista poikkeavat aiemman tutkimuksen käsityksistä, jotka perustuvat jo osittain vanhentuneisiin näkemyksiin äännehistoriasta.

Tärkeimpinä tuloksina esitetään lainaetymologioiden jaottelu todennäköisiin, epätodennäköisiin ja epävarmoinhin etymologioihin sekä esitetään lainoissa esiintyvät äännesubstituutiot ja niiden pohjalta määriteltävä kronologinen jako eri lainasanakerrostumiin. Työn tulosten perusteella voidaan todeta, että merkittävä osa aiemmassa tutkimuksessa esitetyistä etymologioista on epävarmoja tai vääriä. Lisäksi on huomattavaa, että melko useat työssä käsitellyt etymologiat on äänteellisin kriteerein varsin vaikea luokitella kantaindoiranilaisiksi, kantairanilaisiksi tai myöhemmistä iranilaisista kielistä saaduiksi lainasanoiksi. Usean etymologian kohdalla voidaan myös todeta, että indoiranilaisen laina-alkuperän lisäksi myös lainautuminen jostakin muusta arkaaisesta indoeurooppalaisesta kielestä on mahdollista. Lisäksi työssä esitetään muutamia uusia indoiranilaisia etymologioita.

Huomattava tulos on myös, että lainojen äänneasun on esihistoriallisia lainasanoja tutkittaessa levikkiä tärkeämpi kriteeri, sillä vuosituhansien saatossa lainasanoja on muun sanaston mukana hävinnyt ja korvautunut. Lainasanojen levikin yksinomaisen merkityksen uralilaisen kielikunnan taksonomian evidenssinä on siis kyseenalaista. Levikkiä on vääristänyt myös lainojen rinnakkainen lainautuminen, sillä osa varhemmassa tutkimuksessa jo uralilaiseen kantakieleen saaduiksi oletetuista lainasanoista on todellisuudessa varhaisen, rinnakkaisen lainautumisen tulosta.

## ACKNOWLEDGEMENTS

This work has benefited hugely from the assistance of various colleagues and teachers, and it is no exaggeration to mention that without their precious and kind help this work would have never materialized.

I am in eternal debt of gratefulness to my three supervisors who have guided me through the research and the writing process. Professor Janne Saarikivi was my first real guide into mysteries of etymology, and it is largely thanks to him that I got into the study of loanwords. His healthy scepticism towards earlier research and open mind to new ideas has always been an inspiration to me, and the spirit of research that he gave in his lectures about etymology and historical phonology during my student years has guided me through the years. Ulla-Maija Forsberg has been my main supervisor, and thanks to her kind but staunch support I have been able to keep up in some kind of tempo and finish things more or less on time. Professor Ulla-Maija's knowledge of Ob-Ugric languages and helpful and friendly attitude have been essential in the birth of this dissertation.

Lastly, without downplaying the importance of anyone else, I own my greatest debt to docent Petri Kallio, the uncrowned king of Uralic and Indo-European loanword studies. Without his mastery of the two language families my work would be nothing. Petri's quick reactions to my questions and friendly comments to even the poorest drafts is something I cannot thank enough

It goes without saying that the reviewers of my dissertation deserve my deepest thanks. Professor Martin Kümmel and Professor Sirkka-Liisa Hahmo gave me many good, critical but encouraging comments that have been essential help in finishing this study. I am honoured to have such important scholars to read my dissertation. I would also like to thank Professor Kümmel for acting as the opponent in my defence. I would also like to express my gratitude to the representatives of the Faculty of Arts, Professors Jouko Lindstedt and Mika Lavento.

I would also like to thank my other teachers for valuable help during all of my years as a student and researcher. Professor Riho Grünthal was one of my first teachers in historical-comparative Uralic linguistics, and he has also offered valuable help during all my years as a PhD student, for which I am very grateful. Docent Tapani Salminen encouraged me to study Siberian languages and their lexicon, and he has always helped with his immense knowledge of Samoyedic. I am also grateful to my other teachers in Helsinki, Rigina Ajanki, Arja Hamari, Magdolna Kovács, Ildikó Vecsernyés and others. Petri Pohjanlehto's courses in various Indo-European languages have been not only extremely useful but very enjoyable, and I offer him my heartfelt thanks. During my various periods in Budapest, I have benefited greatly from the teachings of Dr. Márta Csepregi and the various Indo-European courses that Dr. Máté Ittész taught; I would like to thank him for welcoming me to his lectures and for being very helpful in many Indo-European matters. In the University of Vienna, Melanie Malzahn was very kind to act as my host as a visiting PhD student during the winter semester 2018–2019 and also to give valuable comments on many etymologies. The teachings by her and Dr. Hannes Fellner have been very useful and influential, and I am very thankful to them.

Among other teachers and mentors whom I would especially like to thank are Professor Ante Aikio (Luobbal Sámmol Sámmol Ánte), whose etymological and historical works have been a great inspiration to me, and whose comments on my early papers and presentations had great influence on my work. I would also like to thank Professor Aikio for the chance to read his manuscript of Uralic etymological dictionary. To Dr. Mikhail Zhivlov I am grateful for the possibility to consult his unpublished dissertation about the development of Ob-Ugric vocalism. I would like to thank also Professor Johanna Nichols for being always supportive of my work and also for giving me a chance to work in an interesting project with her during the final stages of my PhD. To Professor Martin Schwartz I give thanks for commenting many Iranian etymologies and for showing interest in my work.

I am also grateful to Dr. Natalia Kuznetsova, Professor Juha Janhunen and researchers of the Institute of Linguistics, Russian Academy of Sciences, for helping me to get a copy of Alla Lushnikova's dissertation from the archives of the Institute.

My work would not have been possible without the support of various funders over the years. I am grateful to Wihuri foundation for providing the funding in the initial phase of my PhD work, and to the Faculty of Arts of the University of Helsinki for funding most of my time as a PhD student. Final stages of the dissertation work have also benefited from the funding of the Kone foundation.

My many kind and friendly colleagues also deserve my gratitude for the immerse help and comradeship they have provided during the years. My fellow etymology scholars and friends Juha Kuokkala and Niklas Metsäranta have helped me more than anyone else, and without the lengthy discussions with them and the useful comments they provided my work would have much less value than it has now. Juha's encyclopedic knowledge of especially Saami and Finnic, his technical skills and his calm but firm attitude to research have been essential help to me. Niklas has always provided answers to my questions about the tricky problems of Mari and Permic, and his extremely creative mind has produced etymological remarks that have enriched my dissertation vastly. I am also grateful to my friend, Dr. Santeri Junttila for precious help in questions related to Finnic and Baltic, and for pleasant company during various conference trips. Juho Pystynen, one of the best specialist of Uralic historical phonology, has also commented various parts of my dissertation and many of my earlier etymologies, for which I offer my deepest thanks.

The list of dear colleagues is whom I would like to thank is too long to be listed here, but to mention just the few, I thank Chams Bernard, Jeremy Bradley, András Czentnár, Nikolett Gulyás Farkasné, Jaakko Helke, Jeongdo Kim, Veronika Milanova, Oliver Plötz, Zsolt Simon and many others (you know who you are) for help during various phases of my career. Christopher Culver has been of special assistance in proofreading the English of my dissertation and also provided many enlightening comments to many etymologies. For Zoltán Molnár I am forever grateful for the immense help he has given me in the library of the Finno-Ugric department of ELTE, Budapest during the times I have spent in Hungary. A special thanks goes also to my dear friend and colleague Niko Partanen, with whom I have shared some of the greatest moments of my career, all the way from introductory courses of Uralistics to field trips and conferences, and who has always been there when I have been in need of technical assistance, and what is more important, of a friend.

Last but not least, I would like to thank my family: my parents, my brother and sisters and my grandparents for always supporting me in everything. My sister Reeta also helped me with various technical details of my dissertation, for which she deserves special thanks.

The greatest thanks goes to my love Laura, for always being there for me and also for providing a great deal of help with my dissertation, especially during the final phases where her help was essential. Köszönöm szépen, Szerelmem!

Budapest, 25.11.2019  
Sampsa Holopainen

# CONTENTS

Tiivistelmä  
Acknowledgements  
Contents

1.1. Aims .....	1
1.2. Research history .....	4
1.3. The research material and sources.....	13
1.3.1. The loanword material .....	13
1.3.2. Sources on Uralic and Indo-Iranian etymology and phonology.....	14
1.4. The Uralic languages.....	15
1.4.1. The taxonomy of the Uralic languages and the earliest split .....	15
1.4.2. Proto-Uralic phonology and its historical developments in the daughter languages .....	19
1.4.2.1. *ɨ .....	20
1.4.2.2. Non-initial syllables .....	21
1.4.2.3. Alternative views on Uralic vocalism .....	22
1.4.3. Consonants .....	23
1.4.3.1. The Uralic velar spirant *ɣ .....	23
1.4.3.2. The reconstruction and development of the PU affricates and sibilants .....	23
1.4.3.3. PU *ć and *ś .....	24
1.4.3.4. Reflexes of PU *č and *š .....	25
1.5. The Indo-Iranian languages .....	25
1.5.1. The taxonomy of Indo-Iranian languages .....	25
1.5.2. Indo-Iranian phoneme system .....	26
1.5.3.3. From PI to Old Iranian.....	34
1.5.3.4. Middle Iranian .....	34
1.6. Geographical and chronological context of the contacts.....	37
1.6. Methods .....	37
1.6.1. General notes on loanword research .....	37
1.6.2. On sound substitution.....	40
1.6.3 Distribution as a criterion in the dating of loanwords .....	43
1.6.4. Semantics.....	46
1.7. Sound substitutions.....	48
1.7.1. Vowels.....	49
1.7.2. Consonants.....	51

1.7.3. Word-initial consonant clusters .....	53
2. The etymologies: a critical analysis .....	53
2.1. Introductory notes .....	53
2.2. Indo-Iranian loanwords in Uralic .....	54
2.3. New etymologies .....	323
3.1. Analysis of the sound substitutions .....	327
3.2. Vowels .....	327
3.2.1. PII/PI *a and *ā .....	327
3.2.1. PRE-II *e .....	329
3.2.2. Pre-II *o .....	329
3.2.3. *r̥ .....	329
3.2.4. *r̥H .....	329
3.2.5. The substitutions of PII and Pre-II vowels and the vowel reconstruction by Táros (1984; 1987) and Abondolo (1996; 1998).....	330
3.2.6. Non-initial syllables .....	330
3.3. Consonants .....	331
3.3.1. Laryngeals.....	331
3.3.3.1. VHV .....	331
3.3.3.2. #HV .....	332
3.3.3.3. CH .....	332
3.3.4. Stops .....	332
3.3.4.1. Unvoiced stops .....	332
3.3.4.2. Voiced and aspirated stops .....	332
3.3.4.3. Iranian *x, *θ, *p .....	333
3.3.5. Sibilants .....	333
3.3.5.1. PII, PI *s .....	333
3.3.5.2. PII, PI *š .....	333
3.3.6. Affricates.....	334
3.3.6.1. PII *č, *č', *č' h .....	334
3.3.6.2. PI *ts, *dz .....	334
3.3.6.3. PII & PI *č, *č̣ .....	335
3.3.6.4. Later Iranian *s, *z (from *ts, *dz) .....	335
3.3.6.5. Glides .....	335
3.3.6.6. *l, *r .....	335

3.3.6.7. Evidence for Andronovo Aryan and Indo-Aryan loanwords?	335
3.4. List of convincing, uncertain and unconvincing Indo-Iranian etymologies	336
3.4.3. Pre-Indo-Iranian	336
3.4.4. Proto-Indo-Iranian	336
3.4.2.1. Uncertain cases	336
3.4.3. Proto-Iranian	337
3.4.4.1. Uncertain cases	337
3.4.5. Old Iranian and ambiguous early Iranian (no PI criteria, no Osseticisms)	337
3.4.6. Alanic/“Pre-Ossetic”	337
3.4.7. Ambiguous early loans (can be either from PII or PI)	337
3.4.7.1. Uncertain cases	337
3.4.8. Other Middle Iranian/late Iranian	338
3.4.9. List of uncertain Indo-Iranian etymologies	338
3.4.10. Etymologies that were probably borrowed from another Indo-European source (PIE, PBSI, Germanic, Baltic)	338
3.4.11. Unconvincing etymologies	338
3.5. Early loans into already differentiated branches	339
3.6. Discussion about the distribution and its impact on Uralic taxonomy	343
3.7. Impacts on dating and the location of the contact zones	344
3.8. Lexical evidence of cultural influence	345
3.9. Directions for further studies	345
Abbreviations of languages	375
APPENDIX I: Later Iranian loans in Permic languages	377
APPENDIX II: Possible Uralic loans in Indo-Iranian	385



# 1. INTRODUCTION

## 1.1. AIMS

The objective of this dissertation is to offer an up-to-date synthesis of the lexical influence of Indo-Iranian on the Uralic languages. The work continues the tradition, already over a century old, of loanword studies dealing with Indo-Iranian loanwords in the Uralic (Finnic-Ugric) languages (here these terms are used as synonyms; the reason for this will be explained later). The starting points of this tradition are Munkácsi's major work on Indo-Iranian–Uralic contacts in 1901, which was the first ample synthesis of Indo-Iranian–Uralic contacts, and Vilhelm Thomsen's pioneering research on the contacts of Finnic with Baltic and Germanic (Thomsen 1869, 1890), which paved the way for modern loanword studies in Uralic linguistics and in historical linguistics in general.

In this dissertation, the loan etymologies that have been suggested in earlier studies have been gathered together, presented and commented on critically. The loanwords are reviewed based on phonological, distributional and semantic arguments. Some new etymological proposals will also be made in Chapter 2.2. The main task will be to present a stratigraphy for the borrowings, which can only be based on critical scrutiny of sound substitutions.

The central research questions of my dissertation are the following:

1. How many of the proposed Indo-Iranian loan etymologies for Uralic words are correct?
2. How many different layers do these loanwords represent, and based on what criteria can they be stratified?
3. How can Indo-Iranian loanwords be distinguished from other early Indo-European (such as Proto-Indo-European or Proto-Balto-Slavic) loanwords into Uralic?
4. Which loanwords have been borrowed independently into different branches of Uralic, and how can these be recognized?
5. At which phase of Uralic linguistic prehistory were the earliest loanwords from Indo-Iranian acquired, and what can the loanwords tell us about the taxonomy and early divergence of the Uralic language family?

While it is also an important task for science to seek an answer to the questions relating to the geographical area of the prehistorical contacts between Uralic and Indo-Iranian, in this work this matter is of secondary significance. The same must be said about the absolute datings of the loanwords. Although the relative chronology of the loanwords and their distribution in the Uralic language family are among the main objectives of this work, the extralinguistic conclusions must be left to other specialists to deal with. That means that even if I am able to show that a given word has been acquired from, let us say, Proto-Iranian to Pre-Mordvin, I still do not have knowledge of when and where this particular borrowing took place.

Without a strict analysis of the linguistic material, it is impossible to draw conclusions on extralinguistic issues, and when dealing with contacts that have taken place thousands of years ago, without adequate results of linguistic research one cannot hope to achieve

success in extralinguistic matters. The other major field of study which can assist in locating the early contact zones is of course archaeology, but it should go without saying that without a proper linguistic background, there is no possibility to convincingly connect the results of archaeology with those of linguistics. I deliberately leave out the references to various archaeological cultures here, for the main idea is to concentrate on linguistic evidence, which is the central, key evidence when one is studying loanwords. The extralinguistic conclusions can then be drawn from the results of etymological research with the help of archaeologists, archaeogeneticists and other specialists, but it cannot be done in this work.

This dissertation consists mainly of a critical discussion about etymologies that earlier researchers have presented. While one could question the need for a study that offers little in terms of new loan etymologies, it is by no means futile to focus one's research on previously suggested etymologies. Although large amounts of data have been assembled in earlier research, it has not been worked out or properly analysed in the framework of modern Uralic and Indo-Iranian studies. This means that for researchers who are interested in this topic, but not familiar with all of the relevant languages and their historical developments, it is an enormously difficult task to judge which of the etymologies are convincing. Furthermore, the etymologies are split across a number of different sources (some of which are even unpublished), making it very difficult to approach the topic.

The situation with the corpus of the Indo-Iranian loanwords is, in a way, similar to the situation of the Uralic inherited vocabulary in the 1980s, as described by Janhunen (1981: 221): the corpus of the relevant words exists, but the critical analysis remains an urgent task. It is obvious that as the size of our reconstructed Proto-Uralic lexicon has increased since Janhunen's time, especially by Ante Aikio in numerous works over the past two decades (see Aikio 2002, 2006, 2014b, 2015a&b), it is still possible to increase the corpus of Indo-Iranian loanwords with new findings. Nevertheless, the most pressing duty for a researcher is to come up with a detailed stratigraphy of the etymologies based on modern research tools.

The reason why specifically lexical influence from Indo-Iranian to Uralic is the topic of this study is simple: there have been no serious suggestions of other kinds of influence (such as morphological or syntactic) that the Indo-Iranian languages might have had on Uralic. While it is not impossible that one could find evidence for these kinds of influences too, this question is beyond the scope of this study and must be left to future researchers to solve.

It should also be stated that in this work only Indo-Iranian loans in Uralic are dealt with, and not vice versa. This is simply because few, if any, loans in the opposite direction have been found, as is generally the case also with other early Uralic–Indo-European contacts (see Parpola 2010 and Kümmel 2018a for some potential proposals). These tentative loans in the opposite direction will be listed and briefly discussed in Appendix II.

Starting from the late 19th century and the emergence of modern historical-comparative linguistics, there have been multiple studies on the early contacts between the Indo-Iranian and Uralic languages, making this topic one of the most widely studied field of Uralic etymology and contact linguistics. However, the fact is that no up-to-date survey exists, and in 21st-century Uralic and Indo-Iranian studies the views of various scholars differ more widely than perhaps ever before. This is not only due to the fact that the

research results of the field leave possibilities for many different interpretations, but also due to the fact that researchers in different countries (mainly Finland, Germany, Hungary and Russia) have been reluctant to refer to each other's work (possibly also because not everyone has been able to follow the research written in various languages). Furthermore, the recent developments in Uralic historical phonology are not taken into account by all of the researchers to the same degree. It is clear that there is research into this topic of both high quality and lesser quality, and one of the tasks of the present work is to combine the existing information in a way that is easily accessible to those researchers who are in need of a reference on the topic but are not competent in either Uralic or Indo-Iranian linguistics.

As emphasized by Koivulehto (1999a: 208), the early contacts between the two families can offer crucial evidence for scholars of both Indo-European and Uralic, but in practice the reception of loanword studies in the two fields has been very different throughout history. Differences can be found both between the two research fields and also inside them. In Uralic linguistics the existence of Indo-Iranian loanwords is mainstream knowledge and no serious scholar of Uralic would deny their existence. The differences in views concern mainly the stratigraphy of the loanwords. However, even in Uralic studies the importance of such loanwords in historical phonology and etymology is often neglected, and especially references to more recent research are often lacking.

In general, the situation is worse on the Indo-European side. With slight exaggeration one can agree with the remark by Koivulehto (1999a: 209–210) that the Indo-Europeanists often use outdated sources or are simply uninterested in the topic. The problem is further complicated by the various and often obsolete views expressed in even relatively modern Uralicist works, such as those of Rédei (1986c; 1988) or Katz (2003); for a discussion of these works and their deficiencies, see below. For example, Mallory & Adams (2006) adequately refer to the importance of the early loanwords, but they use mostly Rédei's outdated reconstructions and stratigraphy in support of their theories.

What is even worse, in some cases the existence of these loanwords has been simply rejected. Heggarty & Renfrew (2014: 1694) have argued that all the loanwords can simply be later, Iranian loans from Scythian, whose presence in the Pontic steppe is documented in Greek sources in the first millennium BC (Heggarty & Renfrew argue that this is the time when Iranian first appeared in the East European steppe, and that the *Urheimat* of Indo-Iranians was located east of Anatolia, far from the steppes). These arguments neglect all the linguistic evidence. Although the situation is not generally this hopeless, it is obvious that there is a need for a critical scrutiny of Indo-Iranian–Uralic contacts from the Indo-European side, too.

It should also be emphasized from the beginning that in this work only the Indo-Iranian loanwords are studied, and this is not meant to be an overview of all the early contacts between Uralic and Indo-European. Traditionally these two topics have been studied in an intertwined fashion, but this is because of the older view that all the early loanwords into Proto-Uralic or its daughter languages were borrowed specifically from the Indo-Iranian branch. There were possibly contacts already at the proto-language level (between PIE and PU) before entities such as Indo-Iranian emerged. While it is now clear that there are numerous other archaic Indo-European loanwords in Uralic (see Koivulehto 2016: 277–289), many earlier works (especially those of Joki [1973] and Rédei [1986c]) dealt with the early Indo-Iranian and Indo-European influence as identical, even though

the linguistic data pointed to a different conclusion. In this dissertation, it is important to show that it is possible to distinguish Indo-Iranian loans from other early loanword layers, such as possibly Proto-Indo-European or Proto-Balto-Slavic loanwords, and to provide the criteria to do so, but a more exact commentary on these other loanword layers has to be left to further research. Naturally, in the case of some etymologies, it is extremely difficult to determine whether they are loans from Indo-Iranian or some other branch of Indo-European, and in these cases the characteristics of other Indo-European branches will be dealt with more thoroughly.

The early contacts between Indo-European and Uralic are also connected with the old problem of the possible genetic relationship between the two families (cf. Čop 1974). Despite a long history of serious and less serious attempts to prove that Uralic and Indo-European are related (the so-called Indo-Uralic hypothesis), no conclusive proof has been offered, and the genetic relationship between the two families is not widely accepted (although it must be said that it is more widely accepted among Indo-Europeanists than Uralicists). Some of the recent studies in defence of Indo-Uralic are Kassian, Zhivlov & Starostin (2015), Zhivlov (2019) Kloekhorst (2008b), Kortlandt (2010), Hyllested (2009) and Kümmel (2009; 2015; 2019). For a detailed criticism of Indo-Uralic, see Kallio (2015a). Campbell (2006) and Campbell & Poser (2008) are good and critical overviews on the methodological difficulties in long-range comparisons in general. The criticism by Koivulehto (2001b: 257–259) of the Indo-Uralic comparanda and the lack of regular correspondences between the alleged cognates is still relevant. Also, the repeatedly stressed argument that the basic vocabulary items of Indo-European and Uralic are phonologically too similar speaks in favour of early lexical contacts rather than common ancestry (Rédei 1986c: 18; K. Häkkinen 1987: 242). The question of whether the two families are ultimately related is, however, not central to this study, since the Indo-Iranian borrowings in Uralic are usually so easily identified as loanwords that they cannot be mixed up with possibly inherited words.<sup>1</sup>

In what follows, I will give a concise overview of the research history of the topic and a description of the most important sources, discuss the methodology that is used in this work, and give an overview of the historical development of Uralic and Indo-Iranian phonology and lexicon.

## 1.2. RESEARCH HISTORY

The monograph *Uralier und Indogermanen* by Aulis J. Joki (1973) includes a copious and detailed treatment of the earliest research history on the subject, and because of this the

---

<sup>1</sup> The Uralic languages have also been connected to other macro-families, such as Uralo-Yukaghiric, Uralo-Altaic and Nostratic. While there is no space to deal with these views in detail here, it can be stated that no such hypothesis has serious followers in today's mainstream Uralic studies. Aikio (2014d) has critically evaluated the Uralo-Yukaghir hypothesis with the result that these two families are not related but the alleged lexical correspondences are early loanwords from Uralic to Yukaghir, and there are numerous critical views which show that the Nostratic hypothesis has no place in modern historical-comparative studies (Clackson 2007: 20–23; Ringe 1995; Vine 1991; 1998; Campbell 1998; Campbell & Poser 2008: 243–246).

research from the 1970s onward will be presented here. The research history from the time predating Joki's book will be only briefly commented on when it is necessary to add something to the information which Joki presented. The most important works from the earlier period will be referred to in the main part of the work (Chapter 2): Munkácsi (1901, and a series of shorter articles), Jacobssohn (1922), H. Sköld (1925) and shorter articles by Kalima, Setälä and others.

In addition to Joki's presentation, also Korenchy (1972) includes a brief but informative section on the history of loanword research, commenting also on the various different views on the stratigraphy of the Indo-Iranian loanword layers. Lushnikova's unpublished dissertation (see below) provides a short overview of the research history, too, although this has not been very useful for later researchers because the dissertation is not available in libraries or online.

The research history that Joki offers is not a perfect presentation of the topic. Although it contains the most important works by scholars working in either the Uralic or Indo-European field, some notable works are missing. Especially worth mentioning is the article by Jochem Schindler (1963), where several quite well-founded Indo-Iranian etymologies for Finnic words are presented (such as the etymologies for Finnic *ahnas* and *paksu*, later discussed by Koivulehto [1999c: 223–224, 225 = 2016: 226–227, 228] in more detail). Also, notably missing from Joki's work is any mention of the brief article by Benveniste (1949) on Indo-European animal names, where the Indo-European (and possibly Indo-Iranian) etymology of Uralic \**porćas* 'pig' (> Fi *porsas*) word is discussed (this is a word that Joki himself discussed in length in several publications).

Although the historical treatment of Joki's work is still extremely useful, the other main part of the book, the Wörterverzeichnis, has not stood up so well over time. Joki discusses the etymologies critically and offers valuable comments on many older etymologies, but his ability to distinguish different loanword layers based on phonological criteria is far from Koivulehto's later, much more sophisticated studies. From Joki's material one easily gets the picture that he has not been familiar enough with developments in Indo-Iranian historical linguistics, especially in historical phonology. Very often his estimates of the loanword layers are not based on any evidence but look like more or less educated guesses, and it is almost never expressed explicitly why some loan should be "Proto-Iranian", another should be "Middle Iranian" and so on. Although Joki in principle distinguishes several layers of loanwords (Proto-Indo-Iranian, Proto-Iranian, Old Iranian, Middle Iranian), it is not expressed in clear form what are the criteria for these layers. In defence of Joki one can argue that the material for the book was written and compiled over a long period of several decades, but this does not save the author from criticism concerning the methodological foundations of the loan etymologies.

This makes it a misleading source for etymologists, which is unfortunate because Joki's monograph is the largest synthesis of Indo-Iranian–Uralic contacts that has yet been published. In his otherwise generally positive review of the book, the Indo-Europeanist Oswald Szemerényi (1977a: 129–132) lists some of the shortcomings of Joki's etymologies.

Nevertheless, Joki must be given credit for the work he did in sifting the more convincing etymologies out of the earlier works and presenting them in a clear form. Joki's etymological wordlist serves mainly as a useful compilation of older material, but few of his conclusions on stratigraphy can be considered convincing (in this respect it bears a

resemblance to the UEW, which has been criticized for similar issues, see Zhivlov 2014: 113–114).

In his later handbook contribution, an article entitled “Geschichte der uralischen Sprachgemeinschaft” (1988), Joki also deals with the prehistoric loanwords, albeit rather briefly. Joki mentions that there are roughly 50 early borrowings from “Indo-Iranian” or Pre-Iranian (*voriranisch*). Joki’s definition of these terms and layers is not entirely clear to me as they are not explicitly described in the text. Joki also mentions that Hungarian researchers, such as János Harmatta, often see even the earliest Indo-European loans as “Iranian”. Joki divides the earliest loans into Proto-Indo-Iranian based on the reflexes of PII \*a, and the words which show \*e or \*o he considers Pre-Indo-Iranian. Joki also refers to Koivulehto’s then-new ideas of ancient contacts between archaic branches of Indo-European and the ancestors of Finnic in Northern Europe.

Slightly before the work of Joki was published, the monograph of Éva Korenchy about the contacts between the Iranian and Ob-Ugric languages (*Die iranische Lehnwörter in den obugrischen Sprachen*) saw publication in 1972. Although some of Korenchy’s views were taken into account in individual etymological entries of Joki’s monograph, a commentary on Korenchy’s monograph could not be included in the chapter on research history, which discusses the etymological research up to the year 1971.

Korenchy’s monograph is a significant overview of Iranian borrowings in the Ugric languages. Although only the Ob-Ugrians are mentioned in the name of the book, Korenchy also dealt with all the other (Indo-)Iranian loanwords that were found in Khanty and Mansi, including words that were acquired to earlier proto-languages. As many of the earliest loanwords into Proto-Uralic (or Proto-Finno-Ugric, as assumed by Korenchy) have reflexes in either Khanty or Mansi, Korenchy’s book features entries on such classic loanwords as \*sata (modern reconstruction: \*śiṭa) ‘hundred’ and \*asVra ‘lord’. While there is not much new information offered on these earlier loans, the loanwords that are shared by Khanty and/or Mansi and Hungarian offer more new insights, and Korenchy’s work can be held as the first reliable etymological reference work on the Iranian loans of Ugric as whole.

However, it is only natural that many of Korenchy’s results have become outdated by now. The material for Indo-Iranian borrowings in Ugric has been expanded by Blažek (1990), Helimski (1997b) and others, and many of Korenchy’s views on Iranian and Uralic historical phonology are now outdated due to significant studies by such authors as Mayrhofer (1989), Kümmel (2007), Lipp (2009) (on Indo-Iranian) and by Sammallahti (1988), Janhunen (1981), Aikio (2012, 2015), Zhivlov (2006) (on Ob-Ugric and Uralic). Also Korenchy’s views and ideas on prehistory are outdated, and her traditional ideas on the relationship of the Ugric languages have also been questioned in recent decades (Gulya 1977; Salminen 2002). It is especially worth mentioning that in Korenchy’s era, the historical phonology of Khanty and Mansi was more poorly known than it is today, as even the monograph of Honti (1982) had not yet been published, let alone other, more advanced works such as Tálós (1984) or Zhivlov (2006).

However, there are also shortcomings in Korenchy’s research. First of all, Korenchy relies heavily on the works of Harmatta, whose views often contradict other researchers. A shortcoming that partly stems from this is her clinging to the “Iranian” stratigraphy: even the earliest loans which have no Iranian characteristics are in almost every case explained as “Iranian” loanwords by Korenchy. Here it has to be mentioned that Korenchy took into

account also the newer research by Mayrhofer, which with justification refuted some alleged Proto-Iranian developments, but Korenchy has taken the idea of “only Iranian” too far. Further comments on this can be found within the individual entries of the present work, as well as in the Results chapter.

The work of the Hungarian philologist and Iranist János Harmatta also includes many articles (especially Harmatta 1977, 1992) that deal with the Iranian influence on Uralic. In almost every one of his contributions, the datings of the loans are so problematic that the conclusions are automatically misleading, although some of his ideas about individual etymologies are plausible. Another notable problem in Harmatta’s work is that he, like Korenchy, considers almost everything to be *Iranian*, and he presents even the most archaic loanwords as borrowed from an early form of the Iranian branch, which he (1977) dates to a period between 4500 and 1600 BC, assuming a pace of sound changes happening once in every 300 years.

Károly Rédei, the Hungarian linguist best known as the main editor of the *Uralisches Etymologisches Wörterbuch*, dealt with Indo-Iranian loanwords in several publications, the most important of which is his monograph *Zu indogermanisch-uralischen Sprachkontakten* from 1986. The same topic was already the topic of an article in 1983, and many of Rédei’s conclusions were repeated in a (1988) article in the handbook *The Uralic Languages* edited by Denis Sinor. Many loanwords are also mentioned in the entries of the UEW.

While Rédei’s work has become a kind of standard reference in Uralic studies (it is also referred to by Indo-Europeanists like Mallory & Adams 2006), it has been obvious from the start that Rédei’s work is deeply flawed, especially in the stratigraphy of borrowings. Rédei’s work was heavily criticized by Katz and Gippert. The most problematic thing in Rédei’s research is that his stratigraphy of the early loans to many different layers is ad hoc. One of the most problematic issues is that the oldest loans are considered to be *Indo-Iranian*, although they show no sign of any of the well-known Indo-Iranian phonological developments. This point was raised especially by Gippert (1985). Thus words like *\*weti* ‘water’ and *\*nimi* ‘name’ are considered Indo-Iranian, although there is absolutely no grounds for doing so (for example, the Indo-Iranian word for water, *\*udan-*, fits poorly as a source of Uralic *\*weti*, as there is no trace of the *\*e*-grade stem *\*wed-ŋ-* in Indo-Iranian). There are also similar problems in the division of Proto-Indo-Iranian and Proto-Iranian loanwords, where Rédei again presents no real criteria for this difference, except for the distribution within the Uralic languages.

Rédei was also accused of circulating Joki’s arguments (this has been noted by K. Häkkinen [1987] and especially by Gippert [1985: 158] who accused Rédei of plagiarizing Joki’s results), and it is true that his work fails to bring almost any new arguments to the discussion. The most important part of Rédei’s research is probably his study of Iranian loans in the Permic languages, as it is the first thorough presentation of these loans. Many of the arguments presented in that section have not been reworked or challenged, probably because of the complicated nature of Permic historical phonology. Regarding the older loanwords, Rédei notably accepted Koivulehto’s Pre-Indo-Iranian etymology for Uralic *\*kečră* ‘spindle’.

Many of Rédei’s views on loanwords can also be found in the UEW, although as K. Häkkinen (1987) observes, the UEW and Rédei sometimes give different and even contradictory information about the loanwords. Rédei (1986c) more clearly states that the

words are borrowings, whereas the UEW often expresses the connection of the Uralic and Indo-Iranian words in a rather enigmatic way, and the reader cannot be sure of what the exact relationship of the words across the two language families is. No serious attention has been paid to the stratigraphy of the loans in the UEW.

A notable work of Finnish etymology is the dissertation of Kaisa Häkkinen (1983), which deals with Finnish and Uralic etymology and the methodology of etymological research. The work also discusses Indo-European (including Indo-Iranian) loanwords and includes useful notes and tables on the distribution of early loanwords (op. cit. 218–235), although some of the etymologies that are listed in the book are outdated by now.

The most important work on Indo-Iranian loanwords since the pioneering studies of Munkácsi and Jacobsohn has been carried out by Jorma Koivulehto. This Finnish scholar started as a researcher of Germanic loans in Finnic, but later he wrote extensively on the early contacts between Proto-Indo-European and Uralic (see especially Koivulehto 1991) and also loans from Indo-Iranian to Uralic. While Koivulehto never published a larger synthesis of his views on Indo-Iranian loanwords, he discussed the topic in many long articles (especially 1999a, 2000a&b, 2001a&b, 2003, 2007, 2009), most of which have been republished in Koivulehto (2016). This scholar also planned to publish a monograph about the Indo-Iranian loanwords, and the results of his unpublished manuscript about the topic are referred to in this work.

Koivulehto employed a systematic study of sound substitution in the stratigraphy of the loanwords, and his studies gave more solid and reliable results than those of Joki or Rédei. In addition to those loanword layers that were known from earlier research (Pre-Indo-Iranian loanwords with retained *\*e* and possibly *\*o*, Proto-Indo-Iranian loanwords which show reflexes of PII *\*ć* and PII *\*a*), Koivulehto was able to convincingly show that the western Uralic languages borrowed words from Proto-Iranian. These words can be distinguished from earlier loans through the substitution of the Proto-Iranian affricates *\*ts* and *\*dz* (from PII *\*ć* and *\*ǰ*). Koivulehto (2007) also suggested at least one example of the same substitution in Ob-Ugric too, pointing to similar substitutions employed in different parts of Uralic language area (see the entry *\*päčäy* ‘reindeer’). Interestingly, similar substitutions of the Proto-Iranian affricates are found in early Iranian loanwords of both Tocharian (Isebaert 1991: 143–144; Tremblay 2005a: 423<sup>2</sup>; Peyrot 2018) and Armenian (Olsen 2005: 479–480), although the examples there are even fewer than in the Uralic material. However, together they give valuable evidence for the phonological reconstruction of Proto-Iranian. Koivulehto also suggested several new etymologies and offered detailed comments on many older ones.

In addition to Proto-Iranian loans, Koivulehto argued in favor of early Alanic-type loans in West Uralic (1999a: 226–227). Less convincing is Koivulehto’s argument that some of the Iranian loanwords still retained PII *\*e*; this view has not been widely taken up. Naturally, also some of Koivulehto’s new etymologies are unconvincing (all of these will be discussed in Chapter 2), but his impact on research into Indo-Iranian loanwords cannot be underestimated.

Koivulehto’s ideas on the stratigraphy of Indo-Iranian loanwords have been discussed and developed by Petri Kallio (see especially 2006, 2009) and Jaakko Häkkinen (2009).

---

<sup>2</sup> However, Tremblay also notes that Tocharian loan evidence gives support for the development PII *\*ćw* > *śś* that has possibly occurred in Sakan (without the intermediary phase *\*ts*).

Kallio (2006) argued for a later dating of the split of Proto-Uralic, using the Indo-Iranian loanword evidence as one key argument. J. Häkkinen (2009) attempted to develop Kallio's ideas further, discussing many Indo-Iranian etymologies and their distribution within Uralic in detail, but unfortunately many of his etymological remarks contain inconsistencies and errors (these will be analysed in Chapter 2).

In addition to Koivulehto's important work, also the Finnish Indologist Asko Parpola has suggested several new Indo-Iranian etymologies for Uralic words in his work from the 1990s onwards. At the centre of Parpola's research has been the connection of linguistic prehistory and archaeological cultures. Parpola's work was also influential on Koivulehto's idea of a Proto-Iranian loanword layer in western Uralic languages, distinguished from earlier loans through the substitution of the Proto-Iranian affricates *\*ts* and *\*dz*. Parpola has also contributed some new Indo-Iranian etymologies for Finnic and other Uralic words, as well as Uralic etymologies for Indo-Iranian and other Indo-European words in his more recent works (2010, 2015). A notable idea of Parpola's is the hypothesis that some of the loanwords stem specifically from Proto-Indo-Aryan (already Munkácsi discussed Indo-Aryan influence on Uralic, but Parpola offers new arguments to the discussion). Koivulehto has also discussed possible Indo-Aryan loans but otherwise the idea has not received widespread acceptance. Kümmel (2018a) is critical towards the Indo-Aryan loanword layer because of the lack of convincing evidence (see below in 1.5. for more discussion of Indo-Aryan features).

Another important Finnish scholar is Pekka Sammallahti, who has discussed the Indo-European (including Indo-Iranian) influences on the Saami languages in several articles (1999, 2001, 2002) and briefly also in his comprehensive monograph about the Saami languages (Sammallahti 1998). He has suggested several new Indo-Iranian etymologies for Saami words, some of which have since been refuted by Holopainen (2018a). Although Sammallahti took the latest results of loanword research into account in his works, some of the Indo-European sources he referred to (such as the IEW) and the reconstructions he used were outdated, making some of his etymologies quite uncertain.

Many scholars in the Soviet Union and later the Russian Federation also have researched this topic over the last five decades. Notably the eminent Eugene Helimski dealt with this topic in several articles: his 1992 article (republished in Helimski 2000) on the emergence of Uralic consonantal stems used Indo-Iranian and other Indo-European loans as key evidence, and it was one of the first serious attempts to stratify the loanwords, paying attention to the non-initial syllables as well. Helimski (1997b) discusses Indo-Iranian loanwords more generally, but it is especially notable for the introduction of the "Andronovo Aryan" idea: Helimski argues that some loanwords in Ob-Ugric and Permic are derived from an unattested, third branch of Indo-Iranian. Helimski's idea has been supported by at least Mikhail Zhivlov in a 2013 article, but otherwise it has not received wide acceptance. Helimski was also known for his criticism (see especially Helimski 2001) of Jorma Koivulehto's etymological work: although the main targets of Helimski's criticism were Koivulehto's writings on Proto-Indo-European and Germanic borrowings (which fitted poorly with Helimski's ideas of the Nostratic roots of Proto-Uralic and his other theories on Uralic linguistic prehistory), also some of his Indo-Iranian ideas received unnecessarily sharp criticism in Helimski (2001).

In general, the views of Koivulehto and his followers have not gained much recognition in Russia. Many Russian researchers ignore his work almost completely (Zhivlov 2013;

Napol'skikh 2014). While there is justified criticism towards some of the Koivulehto's Proto-Indo-European etymologies and their methodology and historical context, the Indo-Iranian loan etymologies presented by him can and should be regarded as the most valuable and methodologically high-class evidence that is available. This does not mean, naturally, that all of Koivulehto's Indo-Iranian etymologies should be correct, or that no criticism towards him could be allowed.

In Russian research on Uralic linguistics, Lushnikova's 1990 dissertation on the stratigraphy of Iranian-Uralic contacts has been an important source. Researchers such as Napol'skikh (2010, 2014) have referred to Lushnikova's work. A very detailed criticism of Lushnikova's monograph was written by Helimski (1991). Lushnikova's dissertation has never been published, however, and thus it has remained largely inaccessible for researchers outside of Russia. This means that references to her work are missing from practically all Western etymological works (for example, Koivulehto or Parpola never referred to Lushnikova). Helimski's review was published in the journal *Linguistica Uralica*, and it has thus been known to Western scholars too, but references to it have been made mostly by Russian researchers, such as Napol'skikh (2014) and Zhivlov (2013). For this dissertation I have been able to consult Lushnikova's work, and here both Lushnikova's results and their commentary by Helimski will be taken into account.

Regarding Lushnikova's own views, her work has to be hailed as a worthy attempt to stratify the contacts based on sound substitutions; her aims here are largely the same as those of this work. However, Lushnikova's work suffers from certain shortcomings, some of which are also pointed out by Helimski in his review. First of all, the work includes little original discussion of the etymologies – they are not reviewed in the way Joki or Korenchy have done, but earlier etymologies are presented with little to no criticism, and conclusions are then drawn based on them. Secondly, the author has not taken into account all the necessary research published outside of the Soviet Union. It is true that Koivulehto's main works on Indo-Iranian loanwords were published later in the 1990s and 2000s, but already before that Koivulehto had discussed Indo-Iranian loans in some articles, and he had written extensively about the earliest contacts between the Uralic and Indo-European families, and therefore this is something that ought to be referred to in a dissertation that aims at the stratigraphy of Indo-Iranian loanwords in Uralic. In spite of these shortcomings, Lushnikova's work includes valuable ideas about several etymologies and sound substitutions, which will be discussed in the main part of this work.

Vladimir Napol'skikh is another important Russian scholar who has written on several occasions about Indo-Iranian–Uralic contacts. His 2014 article is notable for its criticism on Helimski's Andronovo Aryan theory and his arguments in favour of Indo-Aryan loanwords. Napol'skikh also considered some of the traditional Indo-Iranian loanwords to be borrowings from Tocharian (see below) in some of his earlier works, an idea which has been criticized by Kallio (2004) and Widmer (2002) and which Napol'skikh himself has since dropped in later publications (2010, 2014), where many of these alleged Tocharian loans are again considered Indo-Iranian.

Some of the main characteristics of Russian research is that the earliest Indo-European loanwords are usually considered to represent an inheritance from the Nostratic proto-language (Helimski [2001]; Kassian, Zhivlov & Starostin [2015]), an idea which is not widely accepted by scholars of Uralic in the West. Although this often does not concern the Indo-Iranian loanwords at all, or it concerns only a part of them, the works of Jorma

Koivulehto, who dealt with both earlier Indo-European and Indo-Iranian loans, receive so much criticism from the Russian scholars that his important ideas are often totally rejected or left unmentioned in Russian research. This kind of rejection of central etymological research literature can be considered one of the most pressing problems in Uralic loanword studies, and it leaves a regrettable gap between Russian and Western European scholars in this perspective.

The Czech linguist Vačlav Blažek has also written several articles (1990, 2003, 2005) on Indo-Iranian–Uralic contacts and discussed individual etymologies in some other articles (2016, 2017a&b, 2018).

It is important to mention that these etymological works by Uralicists have in many cases also been reviewed within Indo-European linguistics. The Indo-Europeanist Oswald Szemerényi (1977a, 1988) reviewed in considerable detail both the works by Joki (1973) and Rédei (1986).

Rédei's work was also reviewed – rather negatively but in cautious detail – by Hartmut Katz (1987). Jost Gippert (1985) also reviewed Rédei's work in the 1983 memorial volume of the Finno-Ugrian Society very negatively in *Kratylos*, criticizing Rédei's view on the stratigraphy of Indo-European material in the book. Rédei (1986b) replied to this review in the pages of *Kratylos*, but he could not offer new arguments in favour of his ideas, many of which were then repeated in subsequent publications by Rédei (1986c; 1988). Rédei's 1983 work was also criticized by A. D. Kylstra (1984), who raised similar points of criticism as Gippert, criticizing Rédei's use of the label "vorurarisches" for his reconstructions that looked simply Indo-European. Rédei also published a reply to Kylstra, again circulating his earlier arguments (see Rédei 1986a).

Hartmut Katz himself made important contributions (see especially Katz 1983a & b) to the study of Indo-Iranian loanwords, but his main work, the Habilitationsschrift (1985) that was edited and published by Katz's students in 2003, is problematic because of its lax methodology and unusual views on both Indo-Iranian and Uralic historical phonology. It is important to list the main problems with Katz's results here, so that each of his etymologies do not have to be discussed in the Wörterverzeichnis in detail.

Several reviews of Katz (2003) have been written by researchers of historical Uralic and Indo-European studies, and almost none have been favorable, namely Aikio & Kallio (2005), Rédei (2004b), Anttila (2006), Pinault (2007). The comments below are based on the views expressed in these works. Blažek (2017a: 179) has also briefly criticized Katz's reconstructions, describing them as ad hoc, and Katzschmann (2005: 96–99) has analyzed the laryngeal etymologies of Katz, judging them as unconvincing.

The greatest problems of Katz's work can be summed up as follows:

1. The reconstructions do not follow the consensus on either Uralic or Indo-Iranian historical linguistics. His Uralic reconstructions follow a version of Wolfgang Steinitz's (1944) theory of paradigmatic vowel alternation in Proto-Uralic, which has generally been considered to be refuted in Uralic linguistics. Also the Indo-Iranian reconstructions of Katz are far from the commonly accepted ones. In addition, often the reconstructions bear more resemblance to Proto-Indo-European than Proto-Indo-Iranian and there is no substance to his claims that the words are borrowed from an early form of Indo-Iranian (in this Katz's ideas largely resemble those of Rédei).

2. Katz presents phonetically problematic sound substitutions, and the same Indo-Iranian consonant cluster can be substituted in a number of ways without Katz giving any conditions for this. This was criticized in detail by Aikio & Kallio (2005: 213–215).
3. In general, Katz does not offer detailed argumentation in defence of his views but seems to present many of his arguments *ex cathedra*.

In spite of all of this, there are some promising etymologies in Katz's work. These include the ones published in earlier works during his lifetime. The etymologies that fulfill the criteria of Indo-Iranian loanwords (as discussed below in 1.4.) are included and discussed in the Wörterverzeichnis section. As an Indo-Europeanist, Katz also offers some important arguments about the Indo-Iranian donor forms.

In addition to the works listed above, Indo-Iranian loanwords have been the topic of many smaller articles, and Indo-Iranian etymologies have received some attention in articles that have discussed other aspects of Uralic etymology. These include Janhunen's (1983) article about Indo-European influences in Samoyed, and many articles by Ante Aikio (2002, 2006a, 2012b, 2014a, 2015a&b).

Also in some wider studies of Indo-European linguistics the contacts between Uralic and Indo-Iranian have received attention. Burrow discussed the loanwords in his 1955 handbook and also commented on the topic in his review of Joki (1973) (Burrow 1976). The monumental work of Gamkrelidze and Ivanov (1984) has to be mentioned, although its presentation of the loanwords suffers from circular argumentation concerning the dating and locating of the contacts, as they do not accept Proto-Indo-Iranian but only Iranian loanwords, and some of the loanwords in Uralic are considered to represent *Tocharian* loans. (The idea that some early loans are from Tocharian was already suggested by Burrow 1976). The same kind of claims are also made by Napol'skikh (2001), who has since changed his views on the issue – see his later publication from 2014). In spite of the shortcomings, Gamkrelidze & Ivanov's work has the methodological advantage of strict substitution rules: unlike Joki (1973) and Rédei (1983; 1986c; 1988), Gamkrelidze and Ivanov attempt to argue that the various substitutions of Indo-Iranian *\*a* can be explained either as conditioned or chronological differences. Already in earlier times some Indo-Europeanists, such as Norbert Jokl (1921), paid attention to the contacts. Detailed information on this can be found in the research history provided by Joki (1973). Tischler (2002) refers to early Indo-European loans in Uralic as a key argument in determining the location of the Indo-European homeland. Tischler (2002: 483–484) mentions two different chronological layers of early Indo-European loans: many Indo-Iranian loanwords such as *\*mekši* and *\*sarwi* are included in the second layer, even though Tischler does not mention that these words are borrowed more specifically from Indo-Iranian.

The Indo-Europeanist Wolfgang P. Schmid (1979) also discusses Indo-Iranian loan evidence in one article. His main idea is that due to the possibly late occurrence of the sound change *\*s > \*h* in Iranian, many of the early loanwords in the Uralic languages could be relatively late. In most cases this contradicts the Uralic evidence.

The Finnish Indo-Europeanist Raimo Anttila also mentions the evidence of Indo-Iranian and other Indo-European loanwords in Uralic in his (2000) monograph about the Proto-Indo-European root *\*ag-*. The dissertation of Adam Hyllested (2014) includes several articles which deal with contacts between Indo-European and Uralic, and although

Indo-Iranian loanwords are not his main focus, some Indo-Iranian loanwords are discussed as well.

Mayrhofer mentions many loan etymologies in his etymological dictionary of Old Indo-Aryan (EWAia), referring to the work of Joki, Katz and Koivulehto. Koivulehto's etymology for Uralic \**kečrã* 'spindle' and some other possible early loans are also mentioned in the *Indogermanische Grammatik* of Cowgill and Mayrhofer (1986: 11, 68, 111), and Mayrhofer referred to the evidence of loanwords already in an article about Brugmann's law in 1952, as well as in his ground-breaking study of Proto-Iranian (1984).

Dolgopolsky (1989: 19–20) writes about the early contacts of Indo-European with other language families, including that of Uralic, referring mostly to Joki (1973) and not taking into account work published after that.

### 1.3. THE RESEARCH MATERIAL AND SOURCES

#### 1.3.1. The loanword material

The main part of this dissertation consists of a review of the 245 Indo-Iranian etymologies that have been presented in previous research. In addition, some new etymologies are presented.

The etymologies include all Indo-Iranian etymologies that Jorma Koivulehto has discussed in his works (1999, 2016, 1991, 2003a&b and other shorter articles, as well as in the unpublished handout from 2003 and an unpublished manuscript [Koivulehto ms.]). Also all Indo-Iranian etymologies presented or discussed by Asko Parpola, Ante Aikio or Petri Kallio are discussed. These together represent the most up-to-date research on Indo-Iranian loanwords in Uralic, which makes these the most important group for the purposes of this dissertation, the stratigraphy of loanwords and the study of sound substitutions. All the Indo-Iranian loan etymologies that have been suggested for Saami words by Sammallahti (1998, 1999, 2001) are included, as well, but these are discussed only briefly, as the reader is referred to Holopainen (2018b) where these etymologies have been reviewed in more detail.

This listing includes most of the Indo-Iranian etymologies that are mentioned in SSA and EES. All the Indo-Iranian etymologies mentioned in these etymological dictionaries are included in the material.

I also analyse and discuss all the etymologies from Korenchy (1972). Most of these etymologies fulfil the criteria of Indo-Iranian borrowings, making it obvious that these etymologies have to be included. Some of the loanwords that are limited to either Khanty and/or Mansi represent probably later borrowings and are technically less important for this work, but because so little work has been done on Ob-Ugric etymology and especially on Indo-Iranian influences, I consider it important to discuss all these etymologies as well.

From the works of Joki (1973, 1988) and Rédei (1983; 1986c; 1988), the UEW as well as Lushnikova, only those etymologies are included that can with justification be considered Indo-Iranian loanwords. As mentioned above, especially Rédei classifies many Indo-European words as early Indo-Iranian even though there is no reason for this. Also for the publications of Katz (1983a&b, 1985, 2003) which are discussed above, the same criteria apply. The most important criterion here is that the etymologies have to show Indo-Iranian

phonological developments: I present these developments in Chapter 1.5.1. Etymologies that have strong arguments in favour of an Indo-Iranian origin because they belong to a semantic group of words that include several obvious Indo-Iranian loanwords (see the case of \*meti) are also included.

In addition, etymologies from the smaller articles and presentation handouts by Blažek, Helimski, Rédei, Janhunen, Benveniste, Kümmel and Schindler are included, based on the same criteria.

Loanwords that are clearly late in Hungarian and Permic languages are left out. For the loans that appear only in Hungarian, only those are included that do not show the obvious and late Ossetic developments. They are listed below in Chapter 1.4.2.2. For the Permic loans, only those that can have predated Proto-Permic vowel changes are included; this is usually an easy task, thanks to the complicated vowel developments, which reveal the age of the words quite easily. The few Permic words that show \*s instead of Iranian h make an exception (see below on Andronovo Aryan developments): they are included even if their vocalism points to a later origin.

### **1.3.2. Sources on Uralic and Indo-Iranian etymology and phonology**

The main source for the Uralic etymological materials will be the wordlist of Sammallahti (1988) as well as the wordlists of Aikio (2015b) and Zhivlov (2014), which represent the most up-to-date views on the Proto-Uralic lexicon. Those parts of Aikio's forthcoming etymological dictionary that have been accessible for me (the handout Aikio 2018) are also used. These articles also provide the state-of-the-art description of the Proto-Uralic sound system and its developments in the various branches.

For Indo-Iranian etymologies, Mayrhofer's EWAia is the single most important source, by far. It is supplemented by the dictionary of Iranian verbs by Cheung (2007), the list of Ossetic etymologies by Cheung (2002) and the Indo-Iranian materials found in other recently published Indo-European etymological dictionaries (LIV<sup>2</sup>, NIL and the relevant entries published so far in the Leiden Indo-European Etymological Dictionary series: Kloekhorst 2008a; de Vaan 2008; Derksen 2008 & 2015; Martirosyan 2010; Beekes 2010; Kroonen 2013). The etymological dictionary of Ossetic by Abaev (1958–1989) is also used, although its statements will be double-checked against newer etymological dictionaries.

For the historical phonology of Indo-Iranian, the article of Mayrhofer in Schmitt (ed. 1989), Beekes (1997), Sims-Williams (1998) as well as Mayrhofer's *Indogermanische Grammatik* (1986) serve as the most important sources. These are supplemented by recent works that describe various developments of Indo-Aryan and Iranian in detail: Lipp (2009), Skjærvo (2009), Kobayashi (2004), Fortson (2010: 181–190) Kümmel (2007), as well as by the articles by Cantera, Lubotsky and Kobayashi on Indo-Iranian historical phonology in Klein et al. (2017–2018). References to earlier standard-works on Indo-Iranian phonology can be found in all these sources.

In addition, many details of Indo-Iranian etymology and historical phonology have been discussed in many shorter articles that cannot be listed here. They are referred to in the context of the relevant etymologies in Chapter 2.

## 1.4. THE URALIC LANGUAGES

### 1.4.1. THE TAXONOMY OF THE URALIC LANGUAGES AND THE EARLIEST SPLIT

Whenever the etymology and historical phonology of the Uralic languages is discussed, it is important to briefly comment on the taxonomy of the Uralic language family. This is not of minor importance to etymological research, as the postulated proto-languages often act as recipient (or donor) languages in the prehistorical language-contact situations. As within the Indo-European languages (whose taxonomy will be briefly dealt with below), there are several open and disputed issues in the taxonomy of the Uralic languages. Although it has been rightly stated that the Uralic languages form a geographical network where the neighboring languages usually share some isoglosses with each other and thus their relationships seem to be less complicated than those of the Indo-European family (see Mallory 2001; Salminen 1999; 2001), in reality the issue of the relatedness of the Uralic languages is still rather complicated, and there are several competing views.

The taxonomy of the Uralic languages is also crucial because the Indo-Iranian loanwords have been considered to represent one of the most important arguments in favor of an early binary split of the Uralic family into two branches, Finno-Ugric and Samoyedic. This traditional binary taxonomy is also the view of both the UEW as well as Janhunen (1981) and Sammallahti (1988). As there are indeed many Indo-Iranian loans shared by the so-called Finno-Ugric languages, and as there seem to be some early borrowings from Indo-Iranian to Samoyedic alone, at first glance this division might seem convincing. However, more recent research (such as Aikio 2015b) has discovered some potential loanwords that might be shared by both “branches”, and most of the other arguments in favour of this binary split have been convincingly rejected by several researchers (see below). It should also be emphasized that since the phonological level of the intermediary reconstructions is identical to Proto-Uralic, one is able to work on loanwords even in the situation where the exact taxonomy is unclear. This has been noted by Koivulehto (1991: 12–13).

Traditionally it was considered that after the split-up of Proto-Uralic, Proto-Finno-Ugric eventually split into two branches, Proto-Finno-Permic and Proto-Ugric. These then split further, Proto-Finno-Ugric into Proto-Permic and Proto-Finno-Volgaic, and Proto-Ugric into (Proto-)Hungarian and Proto-Ob-Ugric. Proto-Finno-Volgaic then underwent another split, again binary, into Proto-Volgaic and Proto-Finno-Saamic, and the binary splits of these then led to the well-reconstructed, well-attested and still living branches of Mari, Mordvin, Saami and Finnic. The Ob-Ugric branch also developed into the modern Khanty and Mansi branches through a similar binary split.

The taxonomy and distribution are also central for other questions than just the initial binary split, as also the Ugric unity has been supported by loanwords, and the distribution of loanwords in various branches has been seen as a key argument to the age of the borrowings (this has been later heavily criticized by Kallio [2012b: 227]).

Although many of the traditionally reconstructed intermediary branches are problematic and supported with very little evidence, most of them can be considered areal-genetic units, a term which Helimski (1982; 2003) invented to describe areally intertwined

units of closely related languages, whose shared features are not inherited but due to later convergence. At least Ugric, Ob-Ugric and Finno-Saamic can, due to the unclear situation of their true genetic status, safely be considered areal-genetic units. It is unclear whether the same can be said of Volgaic (Mari and Mordvin), but the “Finno-Volgaic” unity probably reflects a one-time areal-genetic unit, as many of the shared innovations and especially the shared vocabulary points to such a scenario.

Notions such as Finno-Volgaic or Finno-Permic can thus be used in Uralic etymology when the distribution of the word is described, and this has been done by e.g. Reshetnikov & Zhivlov (2011). This practice will be followed in this work too when distribution is concerned. The similarity of the reconstructed intermediary proto-languages and Proto-Uralic mean that Proto-Uralic can, with justification, be used as a term even when describing a word that is present in even a couple of branches, and many modern researchers (see Aikio 2015b for example) use the term quite freely in this sense.

Traditionally, lexical innovations have played the most important role in the genealogical classification of the Uralic languages. Salminen (2002) has deemed the use of the lexicon as the most important criterion to be outdated, as words frequently disappear from languages (see also Fox 1995: 289). Salminen is right in arguing that since no phonological or morphological innovations can be reconstructed for any of the “intermediary” proto-languages between Proto-Uralic and Proto-Finnic (or Proto-Saami, Proto-Mari, etc.), one cannot reliably assume the existence of proto-languages like Finno-Permic or Finno-Volgaic. Salminen has argued that the phonological criterion is the most crucial one in genealogical classification. This has also been the view of J. Häkkinen (2007, 2009). However, especially in Indo-European linguistics, it is the reconstructed morphology that counts the most (see Clackson 2007: 6–7; Clackson 1994: 60). Paradoxically, in Uralic linguistics the morphological innovations have been rarely used as evidence in linguistic taxonomy.

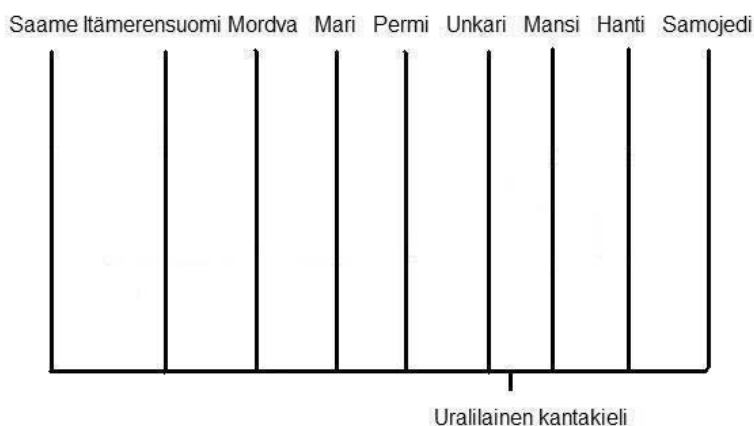


Figure 1. Uralic family-tree according to Salminen (1999). Source: [https://upload.wikimedia.org/wikipedia/fi/3/33/Sukupuus\\_3\\_Salminen.jpg](https://upload.wikimedia.org/wikipedia/fi/3/33/Sukupuus_3_Salminen.jpg)

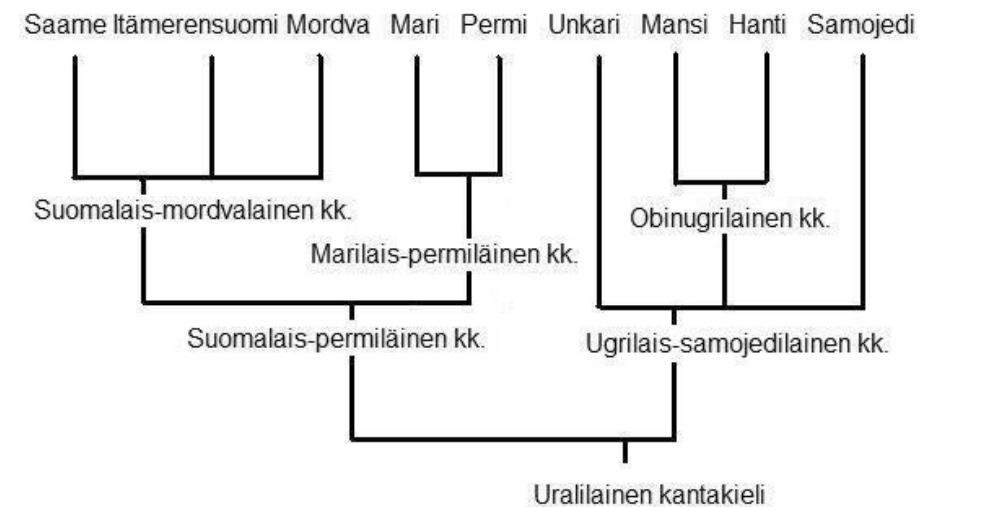


Figure 2. Uralic family-tree according to J. Häkkinen (2007). Source: [https://upload.wikimedia.org/wikipedia/fi/5/50/Sukupuu\\_6\\_H%C3%A4kkinen.jpg](https://upload.wikimedia.org/wikipedia/fi/5/50/Sukupuu_6_H%C3%A4kkinen.jpg)

Kaisa Häkkinen’s 1983 dissertation (see also K. Häkkinen 1984) questions most of the “intermediary proto-languages” and considers it methodologically false to suppose only binary splits. Salminen (1999, 2001, 2002) has continued to develop K. Häkkinen’s views, and has also criticized heavily the notion of Proto-Ugric and Proto-Ob-Ugric unity.

This criticism has gained a large following especially in Finland, where many researchers no longer operate with such levels as “Finno-Permic” or “Finno-Volgaic”. Some other researchers have offered competing taxonomies that differ more or less from the traditional “family tree”, such as Abondolo (1998: 2–6), who still retains most intermediary branches of the family but presents a different order of divergence from Proto-Uralic, or Janhunen (2009: 64–65), whose taxonomy includes disputed branches like “Mansic” (= Hungarian and Mansi) and Finno-Mordvinic (excluding Saami).

One of the most recent alternative taxonomies was presented by Jaakko Häkkinen in 2007 (see also J. Häkkinen [2009] and [2012]). He considers the earliest split of the family to have occurred between “East Uralic” = Ugro-Samoyedic and the rest of the family (either a binary split into Finno-Permic and Ugro-Samoyedic, or a tertiary split into West, Central and East Uralic). He reconstructs West Uralic (Finno-Saami-Mordvinic) and East Uralic based on several innovations, but part of these are based on outdated views of historical vocalism (Häkkinen operates with combinations of vowel and \*x reconstructed by Janhunen [1981], but Aikio [2012a] has questioned this aspect of Janhunen’s vowel reconstruction), and Häkkinen’s conclusions are thus far from certain. Zhivlov (2014: 115–117) has argued that some phonological isoglosses show a closer relationship between Mordvin and Saami, making the assumption of West Uralic more complicated. Of the traditional branches, Häkkinen finds support for Proto-Ob-Ugric but not for Proto-Ugric.

Even though the theories of Jaakko Häkkinen have gained some following and it might be correct to postulate West Uralic and East Uralic proto-languages, it is still clear that

these proto-languages can at best be considered dialects of Proto-Uralic and the “West” and “East” Uralic branches are not well-defined linguistic units like Indo-Iranian, for instance, but rather more distant macro-branches like Graeco-Armeno-Aryan or Italo-Celtic (see especially the works of Clackson [1994], Martirosyan [2013], Schrijver [2016] dealing with intermediary proto-languages in Indo-European).

The taxonomy of the Ugric languages is one of the most difficult questions in Uralic historical-comparative studies. The Ugric languages share numerous innovations, yet most of them seem to be areal or parallel developments, and some of the innovations are shared with neighbouring branches of Uralic, and some show strange gaps in Ugric.

Although the Ugric genetic node is a reality to many researchers, there has in fact been little work done on the actual reconstruction of the Ugric and Ob-Ugric proto-languages. For Proto-Ob-Ugric, there is the pioneering work by Honti (1982), whose failures were clearly pointed at by Tálós (1984) and Helimski (1985). The work of Honti is actually a collection of Ob-Ugric words, not a proper reconstruction of a proto-language itself. Sipos (2002, 2003) has attempted to improve Honti’s reconstructions by sifting out false cognates and parallel borrowings, but despite her good observations on many individual words, she does not offer much new to the reconstruction of Proto-Ob-Ugric phonology. Zhivlov (2006) offers a much better, yet unpublished, version of Proto-Ob-Ugric reconstruction, but like Honti it suffers from the lack of the inclusion of cognates from other branches of Uralic and systematic treatment of developments from Proto-Uralic to Proto-Ugric and Proto-Ob-Ugric. J. Häkkinen (2007: 85–88) also reconstructs a different version of Proto-Ob-Ugric, also presenting the developments from Proto-Uralic and Proto-East Uralic.

Sammallahti (1988: 499–520) represents the first attempt to trace the developments leading from Proto-Uralic all the way to Proto-Ob-Ugric and its daughter languages. His work is based on Honti’s reconstructions but differs from them. While Sammallahti’s work is partly outdated and suffers from the methodological weakness of a forced canonical reconstruction of several intermediary proto-languages, it nevertheless follows a Neogrammarian method of regularity. It is also the first attempt to reconstruct the phonology of the Ugric proto-language.

The second attempt to reconstruct Proto-Ugric is that of Róna-Tas (in WOT:1040–1069), mainly based on the etymological material of the UEW. The UEW also includes Proto-Ugric reconstructions, but these are ad hoc reconstructions and it is not shown how the editors of the UEW came to such conclusions. Bakró-Nagy (2013) is a synthesis of the work on Ugric historical phonology, but it offers no new evidence on the problem.

Tapani Salminen (1989, 2001, 2002) is the most vocal critique of Ugric unity. Salminen’s critical ideas on taxonomy have paved the way for consequent research about the taxonomy and divergence of the Uralic language-family: even if Salminen’s views on Uralic taxonomy are not necessarily correct, he has convincingly shown that the traditional views of Proto-Ugric and other *intermediary proto-languages* are not based on true linguistic innovations.

In summary, there is no consensus on the relationship of the Ugric languages or on the reconstruction of Proto-Ugric and Proto-Ob-Ugric. It is also obvious that in Korenchy’s era the knowledge of the reconstructions was even poorer. This single fact makes it necessary to review her etymologies. Here mainly Zhivlov’s views will be followed, though not without criticism. One might ask if it is possible to start scrutinizing Ugric etymologies if

the details of reconstruction are not clear, but here one has to keep in mind that although many details still remain unclear and subject to debate, we still have a quite coherent picture of the regular relationships between the words in Khanty, Mansi, Hungarian and the other Uralic languages, and the path leading from Proto-Uralic to these individual branches (= the development of Proto-Uralic sounds in Khanty, Mansi and Hungarian) is well-known, so we have enough knowledge to study the Indo-Iranian loans in Ugric. It is merely the debate on the exact prehistorical relationship between these branches that is not known. The same thing could be said of Finnic and Saami as well, but there are still numerous works on various Indo-European borrowings in these branches of Uralic (see Aikio 2012b). And after all, the study of loanwords can also point to new ideas in the reconstruction and relationship of the Ugric languages.

With regard to other controversial intermediary nodes, it is the current consensus that the Finno-Volgaic proto-language cannot be reasonably reconstructed and the hypothesis is now universally rejected (Bereczki 1988). The place of Mari and Permic within the Uralic family is still open, but there has not yet been any serious attempt to link these genetically with the other branches. Although J. Häkkinen (2007: 78) mentions the possibility of connecting them with the West Uralic one (and he also suggests the possibility of a Finno-Permic proto-language), not a single innovation can be adduced to support this view.

Regarding the internal taxonomy of Samoyedic, a split into Northern and Southern Samoyedic was usually postulated in earlier research, but this view has been challenged by Helimski (1998: 482) and Janhunen (1998: 458–459), as well as Janurik (2012), and there are thus different views on the internal taxonomy of Samoyed (see also Salminen 2002). It can be briefly stated here that South Samoyed unity is rejected by practically all researchers, and the closer relationship of Nenets, Enets and Nganasan is also not considered certain by most researchers, with Janhunen (1998) arguing for a very early split of Nganasan.

#### **1.4.2. Proto-Uralic phonology and its historical developments in the daughter languages**

The present work follows the current, prevailing views of Aikio (2012a, 2015b) and Zhivlov (2014), themselves based on the groundbreaking research of Janhunen (1981) and Sammallahti (1988). This reconstruction differs in rather many ways from that of the UEW, let alone pre-1970s sources, even though many of the latter views are still followed by many Uralic scholars, especially in Hungary (cf. Honti 2017, Csúcs 2005, Csúcs & Molnár 2009, WOT, Róna-Tas 2017, Bereczki 2013). There is no space here to examine the differences between the two etymological traditions in detail, but it can briefly be stated that the tradition followed here is based on the strict regularity of sound change, whereas the UEW follows a “pre-Neogrammarian” framework, as Zhivlov (2014: 114) has fittingly put it (see also Ponaryadov [2012] for criticism of UEW and the etymological tradition it represents). The UEW’s reconstructions are basically identical with the ones that Erkki Itkonen used in his works in 1940s and 1950s (see E. Itkonen [1954]). The biggest differences between the framework represented by UEW and the more critical modern studies involve the reconstruction and development of vowels, and notably also the

number of reconstructed etymologies (also the criteria by which the words are assigned to certain proto-languages differ).

According to the sources mentioned above, the reconstructed PU used here has the following phonemes:

Consonants

p	t			k
			č	ć
	s	š	ś	
	ð		ð'	ɣ
m	n		ń	ŋ
	r			
	l			
w			j	

Vowels (initial syllable)

i	ü	ĩ	u
e			o
a			ä

Due to the lack of space, the developments of the Proto-Uralic sound system in the daughter languages cannot be discussed here in full detail. The main sources for these developments are listed here. For general issues in Uralic reconstruction, especially the recent papers of Aikio (2012, 2015) are essential reading, and for Finnic the works of Kallio (2007, 2014) and Pystynen (2018). The Saami developments are presented by Sammallahti (1998). Mari and Mordvin developments are presented clearly by Bereczki (1988), with important additions by Aikio (2014c). For Permian and Ugric, the situation is more complicated because some of the vowel developments in these languages are poorly understood. Zhivlov (2006) and Kálmán (1988) are useful references for Khanty and Mansi, although not all of their conclusions can be accepted. Sammallahti (1988), Csúcs (2005) and Zhivlov (2010) are basic references for Permian vowel history. Sammallahti lists also the developments leading from Proto-Uralic to Proto-Samoyed.

Some peculiar problems of PU reconstruction that are especially relevant for this work will be discussed below in more detail.

#### 1.4.2.1. \*j

This phoneme has been reconstructed as \**e* (= rather mid than high vowel) by J. Häkkinen (2007, 2009), and this practice has been followed more recently by Aikio (2018). The UEW reconstructs \**e* in some cases, but traditionally most cases in which one reconstructs \**j* have been reconstructed as \**a*. This phoneme merged with \**a* in Finnic and in \**a*-stems in Mordvin and Saami as well as in Hungarian. See Aikio (2013a, 2015b) for recent discussions relevant for the reconstruction of \**e*/*ĩ*.

As \**e* often appears as the substitution of Proto-Indo-Iranian \**a*, it is a further argument for the reconstruction of this PU vowel as a non-high vowel. J. Häkkinen (2009) argued that unaccented \**a* was substituted by \**e*, but the material he based his claim on was very limited. This question will be examined in Chapter 2.

### 1.4.2.2. Non-initial syllables

Proto-Uralic had initial stress (preserved in most branches), and the system of non-initial vowels was much simpler than that of initial vowels. This difference has been maintained in practically all branches of the family, although the vowels themselves have changed profoundly.

For Proto-Uralic, a system of *\*i*- and *\*A*-stems is usually reconstructed (Janhunen 1981; Sammallahti 1988). However, there is no conclusion on the exact phonetic nature of the stem vowel *\*i*, and a variety of vowels have been reconstructed as the stem vowel in the history of Uralic linguistics. Most modern scholars reconstruct this stem vowel simply as *\*i*, though in earlier works (such as the UEW or E. Itkonen 1954) *\*e* was often reconstructed, based solely on the Finnic evidence. In the oblique stems of Finnic nouns the vowel appears as *\*e*, and it seemed that the *i* in nominative forms represents a later development. Some researchers have supported the view that the high vowel *\*i* was realized as a reduced vowel *ə* in Proto-Uralic (see Lehtisalo 1936, Janhunen 1982, and more recently Kallio 2012a). Korhonen (1988: 9–10) found support for Janhunen's ideas on reduced vowels in non-initial syllables of Proto-Uralic. According to Korhonen, Proto-Uralic might have been a stress-timed language instead of syllable-timed, and the modern vowel systems of Saami and Finnic would have developed to syllable-based later under influence from neighbouring languages.

The latest reappraisal of the classical system of the non-initial syllable vocalism is Kallio (2012a). Kallio supports the earlier views of Lehtisalo and Janhunen, according to whom the Proto-Uralic stem vowel *\*i* was in reality a reduced vowel *\*ə*.

The *A*-stems consisted of stems ending in *\*-a* and *\*-ä*, depending on vowel harmony. It is also possible that the high vowel *\*i* had front and back variants (*\*i* and *\*j*), and Janhunen (1982) and Kallio (2012a) have argued that the reduced vowel which they reconstruct also had different variants according to palatal harmony (*\*ə* and *\*ǰ*).

The issue of Indo-Iranian borrowings is intertwined with Uralic phonotactics, as it is debatable whether stems of the type *-as* or other kinds of consonantal stems existed in Proto-Uralic, or whether in Proto-Uralic only vocalic stems were allowed phonotactically. There exist several Indo-Iranian loanwords of this shape (where *-as* reflects the nominative masculine ending of Indo-Iranian), and according to Helimski (2000: 194–195), these kinds of consonantal stems were possible in the Finno-Permic proto-language. However, it is debated whether Finno-Permic actually is a valid node; reconstructing *-as* stems for Finno-Permic (but not Finno-Ugric) would, of course, count as a FP innovation, but in the face of the almost total lack of other FP innovations it is debatable whether the emergence of *-as*-stems can be counted as such, especially as most of the reconstructed stems in *-as* reflect Indo-European loans.

In some sources, such as the UEW, the stems in *-as* are reconstructed as *\*-asV*, but this is circular reasoning in order to make the words suit the reconstructed Uralic phonotactics in which only vocalic stems are permitted.

The *o*-stems are not usually reconstructed for Proto-Uralic, as the labial vowels in non-initial syllables in modern Uralic languages can in almost all cases be explained as secondary developments (see Kuokkala [2012 and 2018] for a thorough presentation of especially Finnic and Saami material). However, Aikio (2015b) has recently attempted to re-evaluate the idea of the Proto-Uralic age of the non-initial syllable labial vowels. The Indo-Iranian loanwords possibly reflecting PU non-initial syllable *\*-o* are discussed in

Chapter 2. Aikio bases his assumption on some Finnic and Saami *\*o*-stems in which the labial vowel cannot be explained as the result of a derivational process as well as on the correspondences of initial-syllable vowels in Khanty and Mansi.

The age of the Finnic *u*-stems is even more unsecure than that of the *o*-stems; no *\*u*-stems are reconstructed for Proto-Uralic in any recent studies. However, there seems to be at least one II loan (*paksu* ‘thick’, see below) with *-u* going back to Pre-Finnic and showing a surprisingly neat reflection of the *\*u* in the donor language. This of course represents limited evidence, but Holopainen, Junttila & Kuokkala (2017: 130–131) show that the early emergence of *u*-stems seems possible on the basis of Baltic and Germanic evidence, and that is why the idea of *u*-stems among Indo-Iranian loans has to be investigated with care. I know of no suggestions of reconstructing *\*-ü* for Proto-Uralic non-initial syllables.

A novel idea is presented also by Mikhail Zhivlov in his 2014 study on Uralic vocalism: on the basis of the twofold developments of *\*a*–*a*-stems in Khanty, Hungarian and Mari, Zhivlov reconstructs two different *\*-a* vowels for Proto-Uralic, dubbing them simply *\*a<sub>1</sub>* and *\*a<sub>2</sub>*. Zhivlov also cautiously suggests that *\*a<sub>2</sub>* might be reconstructed as *\*o*, but this idea is not based on real comparative evidence but only on the assumption that since *\*a<sub>2</sub>* does not seem to occur in words with *\*o* in the initial syllable, this might mean that *\*o*–*o* stems were not allowed in Uralic. Zhivlov’s *\*a<sub>2</sub>/\*o* has nothing to do with Aikio’s (2015b) recent idea of Proto-Uralic *\*o* (Aikio’s *\*o* is reconstructed for completely different words than Zhivlov’s *\*a<sub>2</sub>*).

The issue of the stem-types of the Indo-Iranian loans has lately been complicated by Aikio’s assumption that some “secondary *i*-stems” of Finnic reflect original Proto-Uralic *\*i*-stems which have secondarily become *\*a*-stems in Saami and Mordvin (for a detailed analysis see Aikio 2015b). This influences the reconstruction of at least two postulated Indo-Iranian loan etymologies, *\*šarwi* ‘horn’ and *\*šali* ‘intestine’ (*\*šorwa* and *\*šola* in the traditional reconstruction).

### 1.4.2.3. Alternative views on Uralic vocalism

In the Uralic linguistics of the of the 20th century, an “alternative” reconstruction of the historical vocalism emerged; the German scholar Wolfgang Steinitz (see, for example, Steinitz 1944) attempted to reconstruct a paradigmatic vowel alternation of full and reduced vowels in Proto-Uralic, based on the “key languages” Eastern Khanty (with its opposition of full and reduced vowels and paradigmatic vowel alternation) and Western (Hill) Mari (which likewise has a system of opposing full and reduced vowels). This view never became widely accepted in Uralic linguistics, and it was criticized especially by E. Itkonen (1946). However, Steinitz’s PU “ablaut” did convince some researchers to follow this theory, such as Katz, who also based his work studying loanwords on reconstruction of the ablauting Proto-Uralic vocalism.

Although the theory of Steinitz is widely considered to be obsolete, it has recently been praised by Eberhard Winkler (2017). Although Winkler has adhered to the “traditional” reconstructions of the UEW in other works (such as Winkler 2014), in his 2017 article on Uralic historical-comparative linguistics, produced for the *Handbook of Indo-European Linguistics* (Klein et al. (eds.) 2017), Winkler vehemently defends the practice of Katz and

Steinitz and accuses the mainstream scholars of Uralic of abandoning the “Indo-European” practice of historical linguistics.

Related to Steinitz’s views is the theory of Proto-Uralic vowel alternation by Abondolo (1996; 1998). Abondolo’s views were inspired by the ideas of Tálós (1984; 1987), who based his theory on the notion of natural phonology and the idea that vowel shifts are more likely to happen in certain directions. Abondolo and Tálós reconstruct only two tongue heights for Proto-Uralic, contrary to the usual system of three tongue heights (instead of *\*a ~ ä, e, i* Tálós and Abondolo argue for a system of *\*a ~ \*ä* and *\*i*: in this system some of the assumed qualitative oppositions in vocalism are explained as differences in vowel length instead).

Both Tálós and Abondolo have criticized the UEW and the earlier research tradition of Uralic historical vocalism, and even though they have not found general support within Uralic linguistics, their criticism includes many valuable points. Especially Tálós has stressed that his reconstructions are based on the Neogrammarian principle of the regularity of sound change.

Although the ideas of Tálós and Abondolo have not received mainstream support within Uralic linguistics, recently Martin Kümmel has in a number of conference presentations (2015, 2017, 2018) considered them a viable alternative to the mainstream reconstruction of Uralic vocalism. For example, Tálós’s idea that PU *\*e* should be reconstructed as *\*ä* fits Kümmel’s idea that traditional *\*e*-loans such as PU *\*mertä* ‘human’ (= Tálós *\*märtä*) could simply reflect different substitution of PII *\*a*.

### 1.4.3. Consonants

#### 1.4.3.1. The Uralic velar spirant *\*ɣ*

The velar spirant *\*ɣ* has a complicated research history. While this phoneme was reconstructed for Proto-Uralic in the UEW too, it was later connected to a laryngeal phoneme that was behind the long vowels of Proto-Finnic according to Janhunen (1981). Janhunen and Sammallahti (1988) denoted this laryngeal as *\*x*. In Sammallahti’s Proto-Ugric reconstruction the reflex of this spirant was denoted as *\*g*, although no reasons were given for this. As Aikio (2012a) has shown that this *\*x* has little to do with the development of the Finnic long vowels, it is more logical to mark the spirant by *\*ɣ* in the more traditional way. This is the practice most widely used (Aikio 2015b; Zhivlov 2014), and I will follow it in this work.

#### 1.4.3.2. The reconstruction and development of the PU affricates and sibilants

This issue is crucial for the stratigraphy and dating of the loanwords. Although the later Iranian layers can, in principle, also be distinguished on the basis of the vocalism, the differences between the Proto-Indo-Iranian, Proto-Iranian and earliest post-Proto-Iranian stages are so little that the development of Indo-Iranian affricates and their substitutions

in Uralic are often the best way to distinguish the Proto-Indo-Iranian loans from the later ones.

### 1.4.3.3. PU \*ć and \*ś

The two PU phonemes \*ć and \*ś are reconstructed by the UEW and many sources, but Zhivlov (2014) has recently noted that these two sounds might represent only one phoneme. Zhivlov bases his assumption on Sammallahti (1988), who reconstructs only \*ś because the etymologies with \*ć have no cognates in Samoyedic. Zhivlov argues that typologically \*ć would be more suitable than \*ś.

In the assumed Proto-Indo-Iranian loanword layer, both the traditionally assumed \*ś and \*ć appear as the substitute for PII \*ć (from PIE \*kʰ; see below for the reconstruction of the PII affricates). These instances are mostly cases where there are no minimal pairs between the traditional \*ś and \*ć. The reason for PII \*ć being substituted by PU \*ś can be the following: 1) \*ś was in reality an affricate \*ć; 2) there is no affricate \*ć in PU, so \*ś, a palatal sibilant, is the “nearest” sound to PII \*ć; 3) in certain positions (such as word-initial) there is no opposition between PU \*ś and \*ć, and this explains the high number of substitutions by \*ś; PII \*ć is not an affricate but a sibilant (unlikely, see below).

The development of \*ć and \*ś in Ugric and Samoyed is a further complicated issue. It is usually assumed that \*ś becomes \*s in all the Ugric languages as well as in Proto-Samoyed. This is a chain-like development, with \*s and \*š developing into Mansi, Samoyed *t*, Khanty *l* and Hungarian  $\emptyset$ , probably through an intermediary stage of \* $\theta$ . The reflexes of (traditional) \*ć are in Ugric, as in most languages, usually identical to those of \*ś, although a limited number of counter-examples can be shown, at least in Hungarian, where \*ć > *cs* [č] in \**śolmi* > *csomó* ‘knot’. Also some consonant clusters show apparent counter-examples, as Hungarian *s* [š] possibly reflects \*ć*k* or \*ś*k* in some cases, such as *mos* ‘to wash’ < PU \**mośki*-.

A number of apparent exceptions to the Ugric sound change can be found in all branches. Regarding Mansi, the few exceptions are probably based on false reconstructions only (Ms *säj* ‘Eiter’ < PU \**sąji* is an irregular Uralic etymology; for Ugric \**sąptä* ‘seven’, see the entry \**sąptä*) and it can be argued that the change \**s* > \**t* was complete in this branch. In Proto-Mansi the phoneme \**s* seems to have split into \**s* and \**š* under unknown conditions. Honti (1999: 125–130) reconstructs this development already for Proto-Ob-Ugric, but because Khanty does not provide any kind of evidence for this assumption, it is probable that the development \**s* > \**š* is a late development in Mansi (as argued also by Vértes 2003: 275, 278).

In Khanty, it seems clear that the sibilant-assimilation \**s*–*ś* > \**ś*–*ś* (PU \**sükśi* ‘autumn’ > Pre-Kh \**sükśi* > PKh \**söyəs*) took place *before* the Ugric shift \**s* > \* $\theta$ , which heavily points to the conclusion that the latter has to be an areal change. In Hungarian, the reflex of PU \**pesä* ‘nest’ (> Hu *fészek*) is a well-known example of irregular development, but it might mean that in Hungarian \**s* was retained in word-internal position.

For Samoyed, the situation seems to be more or less the same as with Mansi, with \**s* > *t* being a regular change. Because \*ś develops to \**k* in Mator, Mikola (2004: 27–36) has attempted to establish a separate proto-phoneme \*ć, which would produce Mansi *š* and

Samoyed *k*, but this idea is based only on a very few (irregular) etymologies,<sup>3</sup> and this idea should be rejected.

Many loanwords in Ugric and some in Samoyed were clearly acquired before this change took place, as they show similar reflexes as inherited words. Some loans may have been acquired while the change took place, which has led to anomalous correspondences between the Ugric forms (see *\*ačtVrV*, *\*säptä*).

If the sibilant changes in Ugric and Samoyed are late, areal innovations, then these changes have less relevance for the dating of the loanwords. If Proto-Uralic split up into two dialects, as J. Häkkinen (2009) assumed, a loan in Mansi which predates the change *\*s* > *\*t* must have originally had a cognate in Samoyed too. If the change spread areally, the loans that show the change *\*s* > *\*θ*, *\*ś* > *\*s* in Khanty, Mansi, Hungarian and Samoyed can have been acquired independently by these branches.

#### 1.4.3.4. Reflexes of PU *\*č* and *\*š*

Contrary to earlier views that were still followed by Koivulehto in his studies of Iranian loanwords, Aikio (2015b: 44) has shown that Finnic *h* cannot reflect Proto-Uralic *\*č* (cf. also Sammallahti 1999: 76; Saarikivi 2007: 340–344). This means that the reconstructions of many Proto-Uralic or West Uralic words in which Koivulehto reconstructed *\*č* are reconstructed with *\*š* here, if Finnic shows *h*.

For the development of *\*š* in the Ugric languages, see above.

## 1.5. THE INDO-IRANIAN LANGUAGES

### 1.5.1. The taxonomy of Indo-Iranian languages

Indo-Iranian is a clearly defined entity within Indo-European: Iranian and Indo-Aryan languages share numerous common innovations in the sphere of phonology, morphology, syntax and lexicon, and no researcher would disagree with the idea that Iranian and Indo-Aryan descend from a common intermediary proto-language. In addition to Iranian and Indo-Aryan, also Nuristani languages belong to the same branch, but it is questionable whether Nuristani forms a third branch of Indo-Iranian or whether it is more closely related to Iranian or Indo-Aryan (Mayrhofer 1984; Blažek & Hegedűs 2012). A special position for the Dardic languages as a separate branch within Indo-Iranian has also been suggested (Kogan 2005; Kulikov 2007: 175, footnote 4; 2017: 206). Traditionally the Dardic languages were considered as part of the Indo-Aryan branch, and Kogan's arguments for its early split have been briefly criticized by Kümmel (2018a); see also the discussion by Liljegren (2008: 29–36). While the status of Nuristani and Dardic is

---

<sup>3</sup> One of these etymologies is TN *šer* 'Sache', a possible Indo-Iranian loan which will be discussed in the entry *\*čerV*. Mikola's other etymologies can be rejected either as obsolete (TN *šex* < ? PU *\*šakerV*-, cf. SaN *čegar*; Selk *šara* 'hard' ~ SaN *čares* 'karhea' < PU *\*čarV*), or the Samoyed sibilant can be explained as secondary (TN *šij* < PU *\*šüδi* 'coal'; E *šudo* < PU *\*šüinsi* 'breast'; Kam *šē* 'it' < ? PU *\*še*).

important for many aspects of Indo-Iranian linguistics *per se*, from the point of view of the Indo-Iranian–Uralic contacts this question is not maximally important, as the phonological systems of both language groups can be derived from Proto-Indo-Iranian as it is reconstructed in modern standard sources, such as Mayrhofer (1989).

The Indo-Iranian languages share a number of features which distinguish them from the related Indo-European languages, but in some cases it is difficult to determine whether a word has been borrowed from Indo-Iranian or some other branch. Especially difficult is the distinction between Indo-Iranian and Balto-Slavic, as both branches have been in contact with Uralic, and share several isoglosses. These include the RUKI rule and palatalization of velar stops before front vowels, as well as affrication/assibilation of PIE palatal stops (although the reflexes in Balto-Slavic and Indo-Iranian are not exactly identical).

### 1.5.2. Indo-Iranian phoneme system

#### Vowels

i, ī

u, ū

a, ā

r, ṛ

#### Consonants

p

b

b<sup>h</sup>

w

m

t

d

d<sup>h</sup>

r

n s

ć

ǰ

ǰ<sup>h</sup>

č

ǰ̇

ǰ̇<sup>h</sup>

y

š

k

g

g<sup>h</sup>

H

### 1.5.3. INDO-IRANIAN HISTORICAL PHONOLOGY

As stated above, the phonological criteria is the most important criterion in this work for evaluating the etymologies and for establishing substitution rules. Because of this, a sketch of Indo-Iranian historical phonology will be given here. The development from PIE to PII, and after that from PII and PI and the attested Old Iranian languages (including notes on the developments that are attested in the Alanic language of the steppes).

As Proto-Indo-European, also Proto-Indo-Iranian clearly had a mobile stress (somewhat reconstructable on the basis of the stress patterns in Old Iranian and Old Indo-Aryan) and probably a pitch accent, which still survived in the Vedic language (Fortson 2010: 190). Thus the stress system was very different from the Proto-Uralic one, and Indo-Iranian phonotactics did not limit the occurrence of vowels in unstressed syllables in the same way as the Uralic phonotactics did. However, because the vowel system of PII was so simple in general, it is reasonable to suppose that the vowels were realized in a number of ways, which makes the tracking of the substitution rules harder.

#### 1.5.3.1. FROM PIE TO PII

There were numerous sound changes leading to Proto-Indo-Iranian, but because of the simple consonant system of Uralic, not all of them are relevant for research into Indo-Iranian–Uralic contacts. The presentation here is based on the sources that were listed above.

#### Consonants

PIE  $*k'$  > PII  $*ć$

PIE  $*g'$  > PII  $*ǰ$

PIE  $*g'h$  > PII  $*ǰh$

PIE  $*k^w, *g^w$  >  $*k, *g$

PIE  $*h_{1-3}$  > PII  $*H$ ;  $*h_2$  must have remained distinct at first, as it causes aspiration of the preceding stops (see below)

PIE  $*p, t, k + h_2$  >  $*p^h, t^h, k^h$  (= aspirated stop)

PIE  $*s$  ( $*z$ ) >  $*š$  ( $*ž$ ) /  $r, u, k, i$ \_

---

<sup>4</sup> This is now the most probable development, and it is a widely held view that a palatal affricate can be reconstructed for PII. However, especially in earlier works, the PII (and even PI) reflex of PIE  $*k'$  was often reconstructed as a palatal sibilant  $*ś$ . This practice was also still used in Uralic loanword studies by Joki (1973) and Rédei (1986c). From the point of view of the stratigraphy of Uralic loanwords this question is central, and this will be dealt with more thoroughly in the entries for the individual etymologies.

PIE *\*l* > *\*r* (see below for more details)

PIE *\*sk* > *\*ś* before front vowels (Lubotsky 2001b; Cantera 2001: )

Other consonants stayed intact (e.g. *\*PIE p* = PII *\*p*, *\*PIE s* = PII *\*s*, except in the environments described above).

### Notes on the consonantal changes

The three laryngeals did not merge at once in Proto-Indo-Iranian, as can be seen from the laryngeal aspiration (caused by *\*h<sub>2</sub>*). It is unclear what kind of reflex the laryngeals really had in Proto-Indo-Iranian and Proto-Iranian; some kind of consonantal reflex can certainly be accounted for on the basis of the Vedic and Avestan “laryngeal hiatus” (cf. Gippert 1997; Malzahn 2006; Kümmel 2014), but this gives us little evidence for the nature of the laryngeal. The laryngeals merged also in Proto-Balto-Slavic, probably resulting in a glottal stop (Kallio 2008: 267; Kortlandt 1989: 46). A similar assumption for PII could also explain the lack of consonantal substitutions for the Indo-Iranian laryngeal in the loanwords; a glottal stop could not explain the aspiration of stops in Indo-Iranian, but if the aspiration was caused by only *\*h<sub>2</sub>*, there is nothing to prevent us from assuming a glottal-stop like realization for the *\*H* that developed from the merger of *\*h<sub>1</sub>* and *\*h<sub>3</sub>*. Kümmel (2018b: 169–170) assumes that *\*h<sub>3</sub>* probably became a weak fricative early in Proto-Indo-Iranian and was eventually lost in Proto-Indo-Iranian already, as it has not left any clear reflexes, whereas *\*h<sub>1</sub>* and *\*h<sub>2</sub>* retained a stronger articulation and survived longer.

Gotō (2013: 172) notes that “the laryngeals remained as consonants or a consonant in PII”. According to him, *\*h<sub>2</sub>* caused the aspiration of the preceding consonant in PIE already (this is usually assumed to have happened in PII, cf. Kümmel 2018b; Cowgill & Mayrhofer 1986: 135–138). The interconsonantal laryngeals were vocalized as *i* in Indic and with varying reflexes of *i* and zero in Avestan. As noted by Gotō, the laryngeals developed in varying ways in such sequences as PIE *\*NHC* > PII *\*āC* > OI OI *āC* or PIE *\*m<sub>h</sub>H* > OI *ān* before *t* with analogical *n*, *\*NH* > OI *ām* before *s*.

Parpola (2015: 51) has argued that the laryngeals were lost in Core PIE (after the Anatolian languages had split off from the Indo-European unity), and that the reflexes of the laryngeals in the Indo-Iranian languages are “archaizing” traces of a substrate from some unattested branch of Indo-European. This kind of argumentation is complicated, and Parpola can hardly be right, as the reflexes of the laryngeals show divergent developments in various branches of “Core Indo-European”, meaning that the laryngeals were not yet lost by the time Core Indo-European split into different daughter languages, and the Vedic and Avestan laryngeal hiatus speaks in favour of a more simple solution of retained laryngeals rather than some complicated substrate scenario.

The so-called “thorn clusters” present a special problem in the history of Indo-European languages, including Indo-Iranian. The name “thorn” comes from an outdated view of the reconstruction of a separate fricative phoneme, but nowadays it is the consensus that they represent a cluster of dental stop and velar, which then developed in various ways in various branches (see Fortson 2010 for more information). The most recent and most thorough survey of these developments in Indo-Iranian can be found in Lipp (2009). It has been often argued that the Indo-Iranian result of the *\*tk*, *\*dk* clusters

was \*kš in Indo-Iranian, but Lipp (2009: 477–478) argues that this was not the case and the Proto-Indo-Iranian reflexes of these clusters were the palato-alveolar clusters \*tš and \*dž<sup>h</sup>, which differed from the true affricates in that they were not monophonemic. There are not many instances of loanwords where the question of the “thorns” would be relevant, but at least the loanword \*mekši ‘bee’ is possibly borrowed from an Indo-Iranian form which included a reflex of a so-called thorn cluster (but this is disputed; see the entry for a more detailed discussion).

The Proto-Indo-Iranian age of the sound change \*l > \*r is disputed by some, because Old Indo-Aryan sometimes shows (retained) l, and Schwartz (2008) has argued that \*l was retained even in some Iranian languages, based on evidence from modern Iranian (Avestan shows solely \*r while l is completely absent in the phoneme system, and Old Persian has only a very few loanwords in which l appears, with the Indo-European \*l reflected by \*r in all cases and even many loanwords that originally had \*l, such as *Bābirauš* ‘Babylon’). According to Mayrhofer (2004), all these instances of \*l instead of \*r can be explained as secondary: l developed in some Middle Indo-Aryan forms (which were spoken at the time when Sanskrit was used as a written and religious language), and from these l diffused to the Epic and Classical Sanskrit texts, even into some Vedic texts. This is also the view of Kümmel (personal communication, September 2017). Also the Indo-Aryan loanwords in the Hurrian texts of Mitanni show r in cases where Sanskrit shows secondary l (for example, *babrunnu* ‘brown’ = OI *babhrú- id.* [Parpola 1999: 193; 2015: 87; Mayrhofer 1974]).

Parpola (2015: 299) argues that in the original Rigvedic dialect of Indo-Aryan, the PIE \*l had merged with \*r, but the \*l forms were borrowed from an “Atharvavedic” dialect that resulted from a more recent immigration of Indo-Aryans to India, centuries later than the original Rigvedic Indo-Aryans arrived. The l forms then spread secondarily to the original Rigvedic dialect too. While the dialectal origin of the l forms probably are due to borrowing from dialectal forms, it seems unlikely that Parpola’s scenario of several waves of migrating Indo-Aryans is true. The mainstream view that in certain dialects \*l was later generalized for PII \*r seems more plausible, and one can simply assume that dialectal \*l forms were borrowed into the language of the Vedas and into Epic Sanskrit. In any case, in the oldest Vedic texts l is very rare, and becomes frequent only later.

## Vowels

PIE \*e, \*o, \*a > \*a<sup>5</sup>

but: PIE ablauting \*o (not \*o from \*h<sub>3</sub>e) > \*ā in open syllables (the development known as Brugmann’s law)<sup>6</sup>

<sup>5</sup> It is dubious whether there was a vowel \*a in PIE as an independent phoneme, but from the point of view of PII it is not very relevant, as all the non-high vowels merged as \*a (ə) anyway. Parpola (1999) and Koivulehto (1999; 2001) have argued that the vowel merger could have taken place independently in Iranian and Indo-Aryan after the end of PII unity, but this is unlikely in the view of the numerous other innovations that are clearly common to both branches.

<sup>6</sup> The accuracy of Brugmann’s law has been disputed at times but the law is now generally supported; see Lubotsky (1990) with references.

PIE  $*\bar{e}$ ,  $*\bar{o}$ ,  $*\bar{a}$  ( $< *eh_{1-3}$ )  $> *ā$

PIE  $*r_ṣ$ ,  $*l̥$   $>$  PII  $*r_ṣ$ <sup>7</sup>

(PIE  $*r_ṣH$  has different outcomes in Iranian and Indo-Aryan than the simple  $*r_ṣ$ , but in PII  $*r_ṣH$  was probably still intact; see below for the Iranian developments)

PIE  $*n_ṣ$ ,  $*m_ṣ$   $> *a$  (but  $*an$ ,  $*am$  before vowels)

PIE  $*i$ ,  $*u$  stayed intact (as did their semi-vowel allophones)

As PII had only three vowels, it is obvious that the vowels could be realized in a wide variety of ways. This phenomenon in Old Indo-Aryan was noted by the classical Indian grammarians. The vowel that is usually denoted with  $*a$  was usually a schwa, and in the loanwords it is substituted in at least three different ways. It is interesting that in Greek loanwords from Sanskrit the vowel  $a$  is denoted by  $\alpha$ ,  $\epsilon$  and  $o$  (for example, the placename Μέθορα ← *Mathura*), and in loans into Tamil the  $a$  is substituted by  $e$  after palatal consonants, for example Tamil *celam* ‘water’ ← OI *jalām* id. (Oberlies 2017: 360).

### 1.5.3.2. FROM PII TO PI

Of the reconstructed Indo-Iranian languages with which we operate here, Proto-Iranian is clearly the most problematic, for various reasons. While it is usually taken for granted that the attested Iranian languages can be derived from a unitary, reconstructable Iranian (and not directly from Proto-Indo-Iranian). Some aspects of Proto-Iranian are debated, especially the phonetic value of some of the reconstructed phonemes, and some authors such as Tremblay (2004) have questioned many of the sound changes that are traditionally considered to have happened in Proto-Iranian. Most of these debated issues relate to the reflexes of PIE palatovelar and velar stops in Iranian, and these are directly relevant for the study of Iranian-Uralic contacts, as well as the reflexes of PIE  $*s$ , which is likewise closely connected to the stratification of the loanwords in Uralic.

Here the other Proto-Iranian sound changes will be presented briefly, and the sibilants and palatals will be dealt in more detail below. Only those changes will be commented on which show Iranian diverging from PII; the retentions are not specifically mentioned.

## Consonants

(based on Fortson [2010], Beekes [1997])

---

<sup>7</sup> Gotō (2013: 54, footnote 147), referring to Darms (1978) and Kuryłowicz (1948), argues that the starting point for Indic lengthened grade vrddhi is syllabic  $r_ṣ$ . At the time when the Indo-Iranian vrddhi system developed, PII  $*r_ṣ$  was (phonetically)  $*ar$ , and its vrddhi grade had a long vowel:  $\bar{a}r : r_ṣ$ . This was analogically spread to the vrddhi of  $*i$  and  $*u$  in Indo-Aryan (producing the alternation  $\bar{a}i : i$  and  $\bar{a}u : u$ ), whereas this did not happen in Iranian. This is an interesting argument in determining the phonetics of PII  $*r_ṣ$ . But the fact whether it had an epenthetic vowel or not is not that relevant from the point of view of Uralic contacts, as  $*r_ṣ$  was in any case substituted by a sequence of a vowel and a resonant. But if it is possible to show that the reflexes of  $r_ṣ$  were clearly different in, say, Proto-Indo-Iranian and Proto-Iranian (or Old Iranian), this would give us tools to distinguish different layers of sound substitution. The same can be said of the sequence  $*r_ṣH$ .

The deaspiration of PII *\*b<sup>h</sup>*, *\*d<sup>h</sup>*, *\*g<sup>h</sup>* (> PI *b*, *d*, *g*) and *ʒ<sup>h</sup>*, *\*ʒ<sup>h</sup>* (> PI *\*dz*, *\*dž*).

Spirantization of PII voiceless stops *\*p*, *\*t*, *\*k* in consonant clusters and before the laryngeal:

> PI *\*f*, *\*θ*, *\*x*

*\*r<sub>l</sub>H* > ? *\*ar* or *\*ər* (Cantera 2001; 2017: 18–19)

### **PI affricates and sibilants: PIE *\*ć*, *\*ǰ*, *\*ǰ<sup>h</sup>* > PI ?**

Various researchers have widely different views on this, but the view established by Mayrhofer (1989) can be considered to represent the mainstream. The main problems of the reconstruction of the Proto-Iranian reflexes of PIE palatal stops will be presented below.

The development of these phonemes in Proto-Iranian is a rather complicated issue. The reason is that the reflexes in Old Persian (*θ*, *δ*) and in Avestan and the other Iranian languages (*s*, *z*) are very different and it is difficult to reconstruct a convincing Proto-Iranian intermediary phase.

A traditional way was to assume that these sounds, the reflexes of the PIE palatal stops, were *\*ś*, *\*ž* in Proto-Iranian and remained as such in Proto-Iranian (this was also the view of Uralic loanword researchers until the 1980s; see the material in Korenchy 1972, Joki 1973 and Rédei 1986c).

Since the 1980s, several different scenarios for the Proto-Iranian development have been presented, and the development of these ideas has been crucial for Uralic loanword studies too (see Koivulehto 1999c, 2001b; Napol'skikh 2014). Here the different views are listed below, and after that they will be evaluated and briefly commented on. More on their relevance for the Uralic-Iranian contacts will be provided in the Wörterverzeichnis and in the results about substitution rules.

Here the different scenarios that are found in the relevant research literature are presented:

PI *\*ts*, *\*dz*: the PII affricates are depalatalized and the aspiration of *\*ʒ<sup>h</sup>* is lost (Mayrhofer 1989; Sims-Williams 1998; Windfuhr 2009). Mayrhofer considers Iranian and Nuristani originating from the same post-PII branch, and sees the Proto-Nuristani depalatalized affricates as suitable for explaining the Iranian reflexes too, thus postulating the development PII *\*ć*, *\*ǰ*, *\*ǰ<sup>h</sup>* > PI *\*ts*, *\*dz*. This view has been supported by Koivulehto (1999c) and Pärpola (1999), based on loanword evidence.

PI *\*ć*, *\*ǰ*: the PII affricates are retained, only aspiration is lost, as with the aspirated stops (Beekes 1997).

PI *\*ś*, *\*ž*: PII sibilants (!) are retained in Proto-Iranian (Edel'man 1986: 14, 38–45; 1999: 11–12); Kogan (2005: 27) also considers the PI reflexes to be palatal sibilants, but he assumes affricate reflexes *\*ć*, *\*ǰ*, *\*ǰ<sup>h</sup>* for PII.

PI \*s, \*z: the reflexes in Avestan and the majority of Iranian languages are projected back into Proto-Iranian, and it is somehow assumed that the Persian reflexes  $\theta$ ,  $\delta$  developed from these sibilants (Lubotsky 2001; Cheung 2007). A slightly different but rather unlikely view is held by Katz (2003: 40), who has considered the Proto-Iranian affricate \*ts to be a secondary development from \*s.

Lipp (2009) reconstructs also the Proto-Indo-Iranian reflexes differently:

PII \*č, \*j, \*jh > PI \*ts, \*dz

The PII cluster \*sć (from PIE \*sk- before front vowels) developed into Proto-Iranian \*ss according to Lubotsky (2001a: 49). Its reflexes in Avestan are similar to those of PI \*ts, so a development PII \*sć > PI \*ts (Av > s etc.) would also be possible, but in Old Persian the situation is more complicated, as \*sć is reflected by  $\theta$  word-initially and by s inside the word.

It has been argued that in Sakan PII \*ć remained palatal (Sakan śś) before \*w, meaning that \*ć > \*ts did not happen in all environments (Tremblay 2005a, b). Kümmel (p.c.) has argued that Sakan śś is secondarily from \*ts before w. The Uralic loanword material does not give additional evidence for this change, though for the Tocharian loanword arguments in favor of \*ćw > \*śś, see Tremblay (2005a: 423, see also p. 7, footnote 2 above).

### PI (?) \*s > \*h

Szemerényi (1966: 191–194; 1968) considered this change to be later than Proto-Iranian, and his arguments are accepted and supported by Mayrhofer (1989: 7–8). Szemerényi's argument in favor of the late occurrence of this change is that in Elamite written sources transparent Iranian (Persian) loans can be found, which show retained \*s instead of Iranian h. As the Persians first appear in the written sources of the Near East in the early first millennium BC, the loans in Elamite cannot be much older than that, and the appearance of the \*s-words in Elamite shows that the change has to have happened in the Iranian plateau. This means that the sound change must have then spread secondarily to all the other Iranian languages, spoken in a very wide area across Eastern Europe and Central Asia.

The Elamite evidence includes the divine name *as-sa-ra-ma-za-aš*, which could be borrowed from Iranian \**Asura.mazda-*, and the Old Persian name of Elam, *U-v-j/(H)ūža-*, which was probably borrowed with \*s- which then regularly developed into h- (Mayrhofer 1989: 7; the name for Elam is etymologically the same as the name of the Persian city of *Susa*)<sup>8</sup>.

However, not all scholars of Iranian historical phonology agree with this problematic dating. At least Cheung (2007) reconstructs \*h for his Proto-Iranian forms already. The issue is important from the point of view of stratigraphy, as the sound change \*s > \*h is one of the major isoglosses clearly separating the Iranian languages from Indo-Aryan.

---

<sup>8</sup> h is not marked in the Old Persian script, so it is difficult to say for certain whether this sound had disappeared by then, but this has no relevance for the question of the Proto-Iranian status of \*h.

Hintze (1998) has analysed the linguistic evidence for and against the late occurrence of this sound change and attempts to evaluate it also in the light of archaeological evidence, and although her conclusions (op. cit. 148–49) are rather cautious, she thinks it possible that *\*s > \*h* occurred around 1500 BC already, that is, the change could be Proto-Iranian. Hintze is cautious of connecting the Zoroastrian divinity *Ahura.mazda* with the *as-sa-ra-ma-za-aš* that is attested as a divine name in neo-Assyrian test, and even if the two names are connected, the Assyrian form might represent an earlier, fossilized loanword. If Hintze’s dating of the sound change is correct, this would mean that the ambiguous loanwords with retained *\*s* should be considered Proto-Indo-Iranian, although this would still mean that the Iranians and Assyrians were in contact before the change had fully operated. It is also possible that the change *\*s > \*h* was operative for a long period of several hundred years, although Hintze finds this problematic. A good argument in favour of the early date of the sound change is that it had to operate before the time of Zarathustra and the compilation of the Old Avestan canon; and it is not at all clear that Ahura Mazda was a major deity before Zarathustra’s time. This makes it disturbing to find a very archaic form of this divine name attested in the neo-Assyrian documents.

In a personal communication (January 2018), Martin Schwartz has also expressed doubts on the late operation of this change. Schwartz is likewise doubtful of the connection of *Ahura.mazda* and *as-sa-ra-ma-za-aš*, as well as of the connection of *Sūša* and *Hūža*.

Hintze (1998: 145, footnote 43) also correctly notes that the relative chronology of the change *\*s > \*h* and other changes of sibilants in Iranian is problematic: *\*s > \*h* was clearly pan-Iranian, whereas the development of the PII *\*ć, \*ǰh* (from PIE palatal stops) shows clearly divergent results, clearly separating Persian from the rest of the Iranian languages, and in certain cases (such as cluster *\*ćw*) dividing also the rest of the languages into sharp isoglosses.

Connected to this issue is the interesting problem that if some kind of “laryngeal” reflex of the three Indo-European laryngeals was still retained in Proto-Iranian (as is usually assumed, also by Cheung 2007), then what was the difference between the two sounds *\*H* and *\*h*?

On the other hand, Hintze notes that it is unlikely that the sound change in Iranian would have operated independently in the various sub-branches of the family (Hintze speaks of “Eastern” and “Western” Iranian), because the outcomes within Iranian are identical. In my opinion it is also very doubtful that this sound change could have spread areally from the Iranian plateau all the way to Eastern Central Asia without leaving any exceptions.

Carpelan & Parpola (2001: 134) has attempted to connect this Iranian sound change with the well-known similar changes in Armenian and Greek, suggesting an areal connection between these phenomena. While this idea is interesting, it is chronologically and geographically not self-evident, and here one has to note that this sound change is typologically so frequent that it can have operated independently in these three branches of Indo-European. It is also worth noting that the results of the sound change are not identical in Iranian and Greek, as the word-final *-s* is retained in the latter.

To sum up, the absolute dating of the change *\*s > h* is a very complicated problem, and no satisfactory solution to it exists. This means that one has to be cautious when dealing with retained PIE *\*s* in the loanword material, and no far-reaching consequences based on

it can be made. Korenchy (1972) considered many loanwords specifically Iranian because of this sound change, but all of these cases are re-examined in Chapter 2.

## Vowels

The vowel system of PII was retained in PI. In Avestan, the vowels have gone through a number of changes (de Vaan 2008), but these probably post-date PI; the Avestan notation is phonetic, and probably shows different phonetic realizations of the vowels, especially of \**a* (see above for the Sanskrit situation). An important difference from Indo-Aryan is that the PII diphthongs \**ay* and \**aw* were retained in Iranian.

### 1.5.3.3. From PI to Old Iranian

Korenchy (1972) as well as Joki (1973) have used the label “Old Iranian” to describe the source of some of the loanwords in Uralic, but it remains largely unclear what is actually meant by this. In the attested Old Iranian languages there is a major difference between the phonology of Avestan and Old Persian in the reflexes of PI \**ts* and \**dz*. It remains largely unclear whether there is sense in using the same chronology (Old, Middle and Modern Iranian) for the unattested languages of the steppe, but based on the very limited evidence from the loanwords it can be said that the unattested languages were more of the Avestan type in this respect. This is not surprising, as most of the later attested Iranian languages are closer to Avestan in this point.

However, in many other respects it is difficult to draw a line between Proto-Iranian and “Old Iranian” based on the loanword evidence. Other Old Iranian developments such as \**r̥H* > *ar* are very hard to distinguish in loanwords because of the sound substitutions, and in general there are not many innovations separating the attested Old Iranian languages from the reconstructed Iranian proto-language.

About the Old Iranian (?) development \**s* > \**h*, see above.

### 1.5.3.4. Middle Iranian

While reconstructed stages such as Proto-Indo-Iranian and Proto-Iranian can be determined/described unproblematically, it is more troublesome to use terms like “Old Iranian” or “Middle Iranian” in the context of the unattested Iranian languages of Eurasia. These terms have been developed to describe the chronological development of the attested Iranian literary languages, but it is difficult to assume that the developments in the remote steppe would have followed the literary languages so closely that these kinds of unitary terms can be used.

The term Middle Iranian is even more troubling than Old Iranian, as at the time of attested Middle Iranian (roughly 500 BC – 500 AD) the Iranian languages were already very divergent, and the term simply describes a period, not a reconstructed stage. In Uralic loanword research the use of the term “Middle Iranian” has never been properly determined, and it is possible that its use simply points to a relatively late donor language.

At the same time, developments typical of some attested Middle Iranian languages, such as the development of \*a to \*æ, are ascribed to a Middle Iranian donor language.

The term “Middle Iranian” has been used as a similar imprecise cover term for donor languages of Iranian loanwords in the Tocharian languages by Isebaert (1980), but Isebaert’s practice has been criticized by Tremblay (2005a: 421–422). The Tocharian situation is different from the Uralic one in that the donor languages of many of the loanwords in Tocharian can be more securely identified with the attested Middle Iranian languages of Central Asia, unlike the unattested Iranian forms spoken in the Eurasian steppe.

### 1.5.3.5. Alanic developments

Because some of the loanwords in Uralic are supposedly from an unattested Iranian language, which bore a resemblance to Alanic as it is known from the (scanty) ancient and medieval sources and from the evidence of modern Ossetic (Koivulehto 1999c; Helimski 1997b), it is important to list the characteristics of “Alanic” here as precisely as possible. Even some of the early loanwords point to Alanic characteristics. The “Sarmato-Alano-Ossetian” developments that separate this branch from the rest of Iranian according to Bielmeier (1989: 240–42) are the following (it is important to note that the dating and relative chronology of these developments is highly uncertain):

\*p > p/f (Oss. f)

\*fri > fli/li (Oss. li-)

\*θρ > rθ/rt (Oss. rt-)

\*gr > rγ

\*syāw- > saw

\*fra- > ra-

\*xš > xš/xs (Oss. xs)

\*pt > vd (Alanic βδ)

\*rš > rs, rss (Cheung 2002: 32)

\*rt(a) > art (Alanic αρδ)

\*ha- > a-

\*ri/ry > l

\*-a(i)θnī > -in, -inæ

\*sp > Alano-Sarmatian sp, later in Ossetic fs

\*br > \*br, later in Ossetic rv

\**pr* > *pr*, later in Ossetic *rf*

\**a*, \**ā* > Alano-Sarmatian *a* (for later conditioned developments in Ossetic, see Cheung 2002: 12–15)

\**č*, \**š* > Ossetic *c*, *s*, a more recent development

In addition, Windfuhr (2009: 18) lists the development of PII \**ćw* and \**źw* (an important isogloss within Iranian) in Alanic as follows:

\**ćw* > \**tšw* > \**šf* > \**fs*

\**źw* > \**žw* > \**žw* > \**vz*

### 1.5.3.6. Andronovo Aryan

The hypothetical Andronovo Aryan language that has been advocated by Helimski (1997b) and Zhivlov (2013) would differ from the other Indo-Iranian languages in the ways that are listed below. The name Andronovo Aryan stems from the archaeological culture east of the Urals. Note that all the assumed developments are based on loanword evidence, and they will be discussed critically in the context of the loanwords. The developments are few, and not a single one of them is not attested either in Indo-Aryan or Iranian. The argumentation is also made weaker because the alleged Andronovoisms 1, 2 and 3 listed below are archaisms, not innovations, and the rest can be explained as Uralic sound substitutions, not necessarily as true changes in the phonological system of the donor language of the loans. This makes the existence of the Andronovo Aryan language extremely uncertain. The archaeological culture is also so early that it is dubious whether it could have provided relatively late loanwords to the Samoyed and Ob-Ugric languages, not to speak of Permic that is also areally farther from the area of the Andronovo culture. The Andronovo culture probably was inhabited by speakers of Indo-Iranian but this has more plausibly been connected with the Proto-Indo-Iranians than the later offshoot that Helimski and Zhivlov assume.

Napol'skikh (2014) has also criticized this loanword layer, but his conclusions differ from mine. References to Napol'skikh's arguments concerning individual etymologies will be provided when the etymologies are analysed.

The following are the developments argued by Helimski and Zhivlov:

1. The palatal affricates (PII \**ć* etc.) are retained, unlikely in Iranian.
2. The Proto-Indo-Iranian cluster \**ćt* (from PIE \**k't*) is retained in Andronovo Aryan (this became \**št* in both Proto-Iranian and Proto-Indo-Aryan; one could thus assume \**št* for Proto-Indo-Iranian already, but this probably developed via intermediary \**ćt*).
3. Proto-Indo-Iranian \**s* is retained, unlike in Iranian (see above).
4. The cluster \**r<sub>l</sub>H* becomes \**ar*, as in Iranian (vs. *ir*, *īr*, *ur* in Indo-Aryan).
5. Furthermore, Zhivlov argues that the substitution of Indo-Iranian \**a* by Pre-Permic/Pre-Mansi/Pre-Khanty \**ä* points to a development \**a* > \**ä* in the Andronovo Aryan language.

## **1.6. GEOGRAPHICAL AND CHRONOLOGICAL CONTEXT OF THE CONTACTS**

Recently it has been argued that the correlation of linguistic and archaeological data has been approached with too much optimism, and that there is no decisive proof for the idea that archaeological cultures correlate one-to-one with language communities (Saarikivi & Lavento 2012; Aikio 2012b: 97–102; cf. also Erdosy 1995; Gosselain 2016). Due to the long history of archaeological research on the Uralic and Indo-Iranian homelands, one can approximately determine where these languages were spoken and where the contacts happened, but absolute certainty can never be gained.

Archaeological cultures have been studied especially by Parpola, most recently in (2012). Proto-Uralic has been connected with the Sejma-Turbino archaeological network which originated in Central Russia; this was argued already by Kallio (2006: 17–18) and it is supported by Parpola (2012: 156–162). Parpola (2012: 137–141) argues that Indo-Iranian evolved from Proto-Indo-European in the Pontic-Caspian steppe in connection with the late Yamnaya culture.

Based on linguistic and archaeological evidence, Kallio (2006) assumes that Proto-Indo-Iranian and Proto-Uralic probably disintegrated at approximately at the same time, probably around 2000 BCE. Earlier research usually dated Proto-Uralic much earlier (around 4000 BCE), but this would make it chronologically impossible for Proto-Uralic to have contacts with the much later Indo-Iranian.

## **1.6. METHODS**

### **1.6.1. General notes on loanword research**

In this work I follow the Uralic loanword research tradition from recent decades as described above. The well-argued stratigraphy of Germanic and Baltic, as well as Slavic loanwords in Finnic (and Saami) serve as a model of this work (cf. Thomsen 1869; 1890; Mikkola 1894). The former loanword layer is well-presented in LÄGLOS which was compiled by a team led by A.D. Kylstra in Groningen. Methodological questions in connection with the Baltic loanwords have been discussed in detail by Junttila (2015). Also Koivulehto's own works (2001a&d; 2016: 210–211, 356–358) include a useful presentation of the methodology of research into prehistoric loanword layers.

Although the reasons for borrowing of words have been rather heavily discussed in handbooks of historical linguistics, some general methodological remarks are provided here, especially as methodological introductions to loanword studies are lacking in many earlier presentations of the topic (such as Joki 1973 or Rédei 1986c). The reader is referred to Campbell (1999: 59–60), Hock & Joseph (2009: 258–262), and Anttila (1989: 154–177), Durkin (2009: 132–178) as well as Haugen (1950) for more comprehensive theoretical introductions on borrowing in general. It is, however, worth noting that research into

prehistoric language contacts, contacts between reconstructed proto-languages, is in many cases fundamentally different from the loans in attested, literary languages, and many of the handbooks discuss mostly the latter. The methodological literature of Uralic linguists such as Koivulehto and Junttila mentioned above is often the best source.

Borrowing very often results from the need of new names for new things and concepts (the so-called “Wörter und Sachen” loans; see Seebold [1981: 194]; Algeo [1990b: 409]), and many of the loanwords discussed in this work indeed fall into groups of such cultural words. Important cultural terms have been borrowed from Indo-Iranian into Uralic, such as names for some domestic animals, grains, metals and tools. A different kind of cultural borrowings are the substrate borrowings, which represent “local” vocabulary often relating to nature and geography (Saarikivi 2016: 72–74).

On the other hand, as many have pointed out (e.g. Campbell 1999: 59–60; Seebold 1981: 194–197; Hock & Joseph 2009: 258–262), another main reason for borrowing is prestige (= the words are borrowed from speakers of a culturally superior or attractive language without a real need to borrow), and many of the loanwords discussed here have to be considered prestige loans. Kinship terms, some elementary adjectives and verbs can be considered these kinds of prestige borrowings, as they represent concepts that already existed in the language. K. Häkkinen (1983: 101–102) has argued that many early Indo-European loanwords in Uralic and the objects that they denote are often of different ages (this is of course based on the assumption that prehistoric linguistic areas and archaeological cultures correlate in some way), pointing to the fact that it is a very simplified picture to assume that the loanwords have been acquired with the new material concepts that they describe.

Seebold also notes that affection can be considered a reason for borrowing: when negative meanings develop for earlier words, newer words are borrowed to avoid those.

It has been repeatedly stressed (K. Häkkinen 1983; Koivulehto 1991: 12–13; 1999b: 4–7; Junttila 2015: 246–247) that new words do not just appear, i.e. *Urschöpfung* is a very unlikely way to increase the lexicon of a language (naturally this is not true of modern, conscious language planning, such as the case of the Estonian, Finnish or Hungarian language revivals). In addition to the disputed *Urschöpfung*, onomatopoeia, borrowing, compounding and derivation are considered as the main ways to increase the vocabulary of a language.

For recognizing borrowings, Koivulehto lists both language-internal and language-external criteria. Language-internal criteria are the phoneme systems of the contact languages (here Koivulehto lists the differences in Finnic and Germanic consonants and vowels and the substitutions stemming from these differences: Finnic’s richer system of vowels and vowel harmony, the richer sibilant system of early Proto-Finnic). Phonotactic criteria can also point towards borrowing (some combinations can have risen with borrowing and are only found in loans, not in inherited vocabulary, and the phonotactics always influence the substitutions). Also morphological arguments are mentioned by Koivulehto.

As language-external criteria Koivulehto mentions cultural innovations that are visible in loanwords, and related to this, “attraction centres” where words with certain meanings are easily borrowed, and some words are borrowed several times in different periods.

In addition to the reasons for borrowing, also different types of loanwords have been listed in the research literature. Algeo (1990b: 409–410), following Haugen (1950: 214–

215) differentiates between loanwords, loanshifts and loanblends. Loanshifts are more often called *calques*. Although all these three types are important in loanword research, in this work mostly the first one is relevant. For loanwords that have been acquired in the prehistoric period, the notion of loanshifts or calques, as well as loanblends is very difficult to use.

Methodological questions about recognizing loanwords have also been discussed extensively in the literature (Seebold 1981: 210–212; Junttila 2015: 53). As is the case with etymologists in general, researchers of loanwords have to be able to judge the arguments on phonology, morphology and semantics (Durkin 2009: 2–3). In principle, it is crucial to be able to distinguish between borrowing and code-switching, but in the study of prehistoric language-contacts, there is in practice no danger of mixing these two topics. Lushnikova (1990: 8) lists phonetic (= phonological), morphological, geographical, and semantic (“realno-semantičeskij”) criteria for the identification of loanwords. The latter notion implies that the assumed loans have to denote objects that existed at the time of contacts. For example, for Proto-Indo-Iranian various names for metals can be reconstructed, and it is logical that some of these were borrowed into Uralic, while on the other hand, the vocabulary denoting agriculture in Proto-Indo-Iranian is more scarce, cf. Kümmel (2017), making it more unlikely to search for the origin of Uralic agricultural terms in Indo-Iranian.

The importance of word history (tracing the development of the loanword and its meaning in context with the cultural development of the object that it denotes) in the study of borrowing has also been stressed (Polomé 1990: 428–429). Although this is important when the contacts in historical periods are studied, this is less useful when prehistoric contacts are discussed, and especially when the exact datings and locations of the contacts are uncertain. Nevertheless, word history is important in the context of the Avestan, Old Persian and Sanskrit words that reflect the Indo-Iranian source material. Without knowing the context of the texts where these words appear, one could easily mistake later derivations, analogical formations and semantically divergent forms as reflecting earlier stages of linguistic history. One can also attempt to follow the word history on the basis of archaeological cultures, but this is difficult as languages and archaeological cultures do not always correlate (Saarikivi & Lavento 2012).

Junttila’s (2015) presentation of prehistoric Baltic loanwords in Finnic includes the most comprehensive discussion of methods for loanword research within Uralic linguistics. Junttila raises several important points especially on the arguments in favor of loanwords and in the stratigraphy of borrowings. According to Junttila (2015: 136–139), in the study of Baltic loanwords in Finnic the following arguments have been used to argue in favor of the Baltic origin of the etymologies: formal criteria (*muotoargumentit*): sound substitution, sound structure; semantic criteria: semantic correspondence, semantic category; evidence of the source word (*lähtökieliargumentit*); evidence of the loanword layer (loanword layer arguments): existence of the loanword layer, statistical probability; homonymy argument (a different kind of sound substitution is used to avoid homonymy); exceptionality arguments (some words or word classes, such as pronouns, are treated differently from other loanwords); morphological and semantic nativization; etymological nativization and alienization (see the chapter below for a description of this); phonotactic arguments (phonological traits pointing to a borrowing); distribution in the target

language family: inclusive and exclusive distributional criteria. Junttila also argues that the distinction of language-internal derivatives and parallel borrowings is important.

Anttila (1989: 158–160) also lists criteria for identifying the direction of borrowing. As mentioned above, the direction of borrowing is usually very easy to determine with regard to the etymologies that are discussed in this work, but it is still essential to keep the basic criteria in mind, as not all cases are so obvious, and especially in older research one can find contradicting statements on the direction of borrowing. Distribution in either language family is, of course, a crucial criterion. Cultural influence and the more advanced culture of the Indo-Iranians, as well as known historical circumstances such as the wide areal distribution of the Indo-Iranian languages, often make the direction of borrowing obvious.

### 1.6.2. On sound substitution

While in general comparative linguistics the sound law is the basic tool and principle according to which a researcher must work (see Zhivlov 2014: 113–114; Rivila 1966: 38), in loanword studies the same role is filled by the substitution rule (this is emphasized by Junttila 2015: 139). Furthermore, in studying prehistoric language contacts, without mastering the sound laws of the language families in question one cannot work on reliable sound substitutions.

Koivulehto's research is one of the best examples of the use of sound substitutions in the stratigraphy of loanwords. He has been followed by scholars like Saarikivi (2007), Aikio (2009), J. Häkkinen (2009) and Kallio. Although the importance of sound substitutions was stressed already by Thomsen (1869, 1890), LÄGLOS is the first major dictionary where Koivulehto's strict methodology is followed, and it also gives an overview of the methods of sound substitution. Joki (1973) and Rédei (1986c) give sound substitution a rather minimal role in their stratification of the loanwords, and they do not employ the evidence of substitution systematically.

It should be borne in mind that sound substitution and regular sound correspondence are in many ways fundamentally different, although their usage and importance in linguistic stratification is similar. The main difference is that in sound correspondences between genealogically related languages, the regular correspondences are usually one to one: for example, Finnish *k* corresponds to Hungarian *k* in the inherited Uralic vocabulary. There are exceptions to this, but these are because of different environments. In front of old velar vowels, the correspondence set is Finnish *k* ~ Hungarian *h*.

In loanwords, the situation is often more complicated. Although in principle the substitution rules operate in a regular way, very often the loanwords represent different chronological layers, which also manifest results of different substitution patterns, giving almost endless possibilities for different correspondence sets to emerge. This is true also of the Indo-Iranian loanwords in Uralic. A very good example is provided by the Germanic loanwords in Finnic: for example, in early loans Germanic *\*st-* (in word-initial position) corresponds to Finnic *\*s-* (for example, G *\*starrā-* > Fi *sara*; see LÄGLOS I: X and s.v. *sara*). In later loans, the correspondence is Germanic *\*st-* ~ Finnic *\*t-* (for example, G *\*stangō* > Fi. *tanko*). In Indo-Iranian loanwords, the substitutions of the Indo-Iranian reflexes of PIE palatal stops provide a good example of different correspondence sets: Old

Indo-Aryan *ś* corresponds to Finnish *s* in some words (usually from earlier \**ś*), in some words to *h* (from earlier \**ś*). LÄGLOS (I: XVII) notes that in the long-running language contact between Germanic and Finnic, both contact languages have gone through phonological changes, and this makes research on sound substitution difficult. This is probably true of other long-lasting contact situations, too.

In addition, within the same loanword layer one can find various substitutions of the same phoneme of the donor language. The Indo-Iranian \**a* is a notorious example of this, as the loanwords in Uralic show at least \**a*, \**o*, \**i* and \**ä* as its substitutions, and there is no consensus among the researchers about the exact environments or chronological differences of these occurrences.

Even more troubling is that in principle, one cannot expect a similar kind of regularity from sound substitution as from sound law. The relationship between sound laws and sound substitutions has been discussed in detail by Heidermanns (2001). When the vocabulary of related languages is studied, it is generally impossible that in some cognate sets, the correspondences are different purely because the speakers intentionally decided, for some odd reason, to develop some sound differently. This kind of bizarre situation would make the comparative method totally useless. However, in loanword studies one can, in principle, encounter such bizarreness. Even in the most well-known case studies, such as the Germanic loanwords in Finnic, the “regular” substitution rules (such as the correlation of Finnic *t* ~ Germanic *d*, Finnic *tt* ~ Germanic *t*) leave room for exceptions.

The different substitutions occurring in the same layer of loanwords can give an impression that in the study of substitution rules, anything goes. This is the main weakness of Katz (1985, 2003), criticism of whom is discussed above.

These examples should make it clear that research into sound substitution is often challenging. A handbook devoted to the study of sound substitution and its use in the stratigraphy of loanwords would be a very welcome tool for those who work with Uralic (or Indo-European) etymology.

Regarding the theory of substitution, it is often argued that substitution means the use of the nearest phonetic equivalent of the language to replace the foreign sound (Campbell 1999: 61; Hock & Joseph 2009: 247–252; LÄGLOS I: XVII–XXII). LÄGLOS notes that very rarely are the sound systems of the contacting languages similar, and this leads to the substitution of sounds with different ones. It has been repeatedly noted (Haugen 1950: 214; Hock & Joseph 2009; Aikio 2009: 12–17) that there is no real definition for phonetic nearness. Gussenhoven & Jacobs (2005: 35–37) note that there are two phases in substitution: the perceptual phase (the speakers perceive the foreign sound) and the operational phase (the speakers use their native sound to replace the foreign sound).

Algeo (1990b: 409) states that borrowing involves either substitution or importation. The latter denotes the situation where instead of substituting the foreign sound, the new phoneme is acquired by the recipient language.

In practice, one should provide detailed arguments about the phonetic similarity in the case of every substitution in order to convincingly argue in favor of them. For example, Koivulehto’s idea that the Proto-Iranian dental affricate \**ts* was substituted by the Uralic (cacuminal?) affricate \**č* is convincing, because both are affricates and are thus “phonetically near” to each other. However, Koivulehto’s argument that Iranian \**ts* or \**dz* could be substituted also by Uralic \**ks* is much less convincing. Here the affricate *ts* and

the cluster *ks* are phonetically much more different in the way of articulation, but again, the place of articulation is quite “near”.

Koivulehto’s idea of *\*ks* as the substitution of the Iranian affricates is an example of a cluster of sounds substituting for a single sound of the donor language (although phonetically the affricate includes both a stop and a sibilant). This is not a rare phenomenon. Hock & Joseph (2009: 260) provide an example of the English substitution of French nasal vowels: although the French phoneme is one sound, in English it is perceived as two sounds and is substituted by a cluster of a vowel and a nasal stop. A similar substitution pattern has been employed in both Finnic and Hungarian to substitute for the nasal vowels of Slavic loanwords.

A crucial factor in sound substitution is phonotactics, which influences the way foreign sounds can be acquired into a language (Hock & Joseph 2009: 249–250; Campbell 1999: 62). In Indo-Iranian loans phonotactics plays a central role, because Proto-Indo-Iranian and Proto-Iranian had various consonants and consonant-clusters that did not exist in Uralic, and various types of substitutions had to be employed when the loans were acquired. In Proto-Uralic, word-initial consonant clusters were not allowed, and in the loanwords word-initial consonant clusters of Indo-Iranian were always simplified. Phonotactics also determined the substitution of vowels in the non-initial syllables: because fewer vowels were allowed in non-initial syllables, the substitutions were different.

Morphological substitution likewise plays an important role in loanword research (Hock & Joseph 2009: 253). LÄGLOS notes that morphological substitution often makes the stratigraphy of loanwords complicated, and it often involves derivational suffixes. In the study of Germanic loanwords in Finnic, the material is so much larger than it is easier to analyse the morphological substitutions that have been used. Also the derivational morphology of Proto-Finnic is reconstructed in much more detail than that of Proto-Uralic. For these reasons, in the research of Indo-Iranian loanwords the idea of morphological substitution has not really been used.

A special case of sound substitution is etymological nativization. This term is used by Hock & Joseph (2009: 249) and by Aikio (2007; 2009: 12–17). In Aikio’s dissertation the notion of etymological nativization is applied to the lexical contacts between Saami and Finnic. Aikio convincingly shows that due to a large number of lexical cognates and long-lasting contacts, the speakers of Saami and Finnic are aware of the phonological correspondences that exist between their languages, and they substitute the phonemes of loanwords in the way that would be expected for inherited words.

Due to the relatively scarce use of the concept of etymological nativization in Uralic loanword research, it has not been really applied to the study of Indo-Iranian loanwords. It is naturally true that in the study of unrelated contact languages, etymological nativization does not have same kind of relevance as in the study of intensive contacts of related languages, but cases like the Basque-Spanish contacts which Aikio also mentions prove that if the contacts are intensive enough, etymological nativization can also operate in such situations. It is unlikely that the contacts between Uralic and Indo-Iranian were so intensive at any point, but words that are originally from Iranian can have spread between Uralic languages such as Khanty and Mansi that have been in intensive contact with each other, and the archaizing effect of nativization could have played a role. Theoretically, it is also possible that some words with limited distribution in Uralic may appear more archaic

because of etymological nativization, but this has not been suggested as far as I know, and for the time being there is no evidence to support this kind of claim.

### 1.6.3 Distribution as a criterion in the dating of loanwords

Although it may not be explicitly stated in many works, intuitively the distribution of a given word has been used as a major criterion in determining its age. This is done most explicitly by Rédei (1986c), whose stratigraphy of the Indo-Iranian loanwords relies heavily on their distribution in the Uralic languages and gives much less weight to the evidence that the phonetic shape of the Indo-Iranian words gives.

Distribution is naturally an important principle when one has to determine the direction of borrowing (cf. Anttila 1989: 158–160; Junttila 2015: 230–236; Campbell 1999: 64), but within the Indo-Iranian loanwords in the Uralic languages this has in most cases been convincingly done by previous research, and there are very few cases in this work where this kind of consideration would be necessary. Principles for the determination of the direction of borrowing include: the distribution of the word in the contact languages and phonological and phonotactical traits, as well as the morphological complexity and transparency (if the word is analyzable morphologically in one of the contact languages and opaque in the other, the direction of borrowing looks obvious).

Many scholars of Uralic etymology have criticized the use of distribution. Kaisa Häkkinen (1983: 204–205) has stated that although distribution is (along with the phonological and derivational arguments) one of the main determiners of the age of a word, it is a very weak criterion when used alone. She nevertheless goes on to admit that, in practice, distribution is often used as the main criterion. Häkkinen's notions concern mainly the determining of the age of the inherited words, but these same arguments can well be used in prehistoric loanword research, too. She further states that the early loans are not distinguishable from the inherited (“original”) words. Also, the distribution criterion in loanwords is more explicitly criticized (1983: 102–103).

In her 1987 review of Rédei's 1986 book, Kaisa Häkkinen makes a very detailed and specific critique of Rédei's distribution criterion. She is astonished by Rédei's methodology, which gives contradictory results in using both Indo-Iranian phonological developments and Uralic distribution as main criteria. Because of this, Rédei's material lists Pre- and Proto-Indo-Iranian loanwords in both Finno-Ugric and Finno-Permic layers of borrowings, meaning that in the earlier Finno-Ugric period, speakers borrowed words from two chronologically different layers, which remained chronologically different even later, when the Finno-Permic linguistic period started. Similar remarks about Rédei's method are also made by Katz (1987). This is obviously a grave methodological error, and shows in practice that distribution is a very ill-suited criterion in loanword research. Even Rédei himself implicitly accepts this, as he considers some words that appear only in Finnic as parts of the earlier Finno-Volgaic layer. To be consistent with Rédei's own views, one should consider these words borrowings from Indo-Iranian into Proto-Finnic, but the truth is of course the opposite. Words disappear from languages, and distribution is not a trustworthy criterion.

Koivulehto (1999a: 208–209) has criticized the idea that the word's distribution tells more of its age than its phonological shape. Koivulehto's arguments are first and foremost

connected to his ideas of the very early independent contacts of various branches of Uralic with early forms of Indo-European, and he mainly argues that even an archaic word which is present in one branch of Uralic, like Mordvin or Finnic, can result from a very early borrowing. Nevertheless, Koivulehto is certainly right in criticizing the idea that a word with a very limited distribution could not be early. Words do disappear, and within the Uralic family (or any family of languages) there are branches which are lexically more innovative (such as Samoyed) than others, and some that are more conservative. Because of the instability of the lexicon, the number of shared words is considered to be an unreliable criterion in the genealogical classification of languages (Fox 1995: 289), and this same caution should apply to the distribution of early loanwords as well.

Koivulehto's views have been supported by Kallio (2012b: 227). Also Aikio (2009: 21–25) states that distribution is not as central a criterion as meaning and phonological shape, and the distribution criterion is discussed critically by Junttila (2015: 150–151). Aikio & Aikio (2001: 10–11) have criticized the use of distribution in the study of early Germanic loans into Finnic, arguing that because the southern Finnic languages are generally more innovative in their lexicon, the northerly distribution of many loans cannot, in reality, tell much about the true circumstances of the prehistoric contacts.

However, even this is not as simple as it sounds: some of the Indo-Iranian loans seem to have a wide distribution, but upon a closer look it becomes clear that they include phonological irregularities, which can only be explained by assuming that they are parallel loans. The ability to recognize parallel borrowings is extremely important in Uralic loanword studies, and it has been developed with success in the research of Germanic and Baltic loanwords (see Junttila 2015).

To mention an example among the Indo-Iranian loans, perhaps the most well-known Indo-Iranian loanword, Proto-Uralic *\*śata* '100' (in the traditional reconstruction) includes exactly these kinds of problems (this is noted by Kallio 2006: 12, footnote 10). This word will receive a more detailed analysis in the Wörterverzeichnis below, but in order to illustrate the problem, the details of the problematic vowel correspondences will be presented here:

The traditional reconstruction of the word is *\*śata*, but it has been the mainstream view in Uralic studies at least since Janhunen (1981) and Sammallahti (1979; 1988) that some Uralic etymologies which traditionally were reconstructed with *\*a* should be reconstructed with the PU high central vowel *\*i* instead. The idea that PU *\*i* should be reconstructed is much older, and its reconstruction was convincingly supported already by early-20st-century Uralicists such as Gombocz (for example 1920: 112; 1940: 73), but it is not systematically reconstructed by many mainstream sources like the UEW. The opposition between *\*a* and *\*i* has become lost in several branches of the Uralic family (completely in Finnic, and in *\*a*-stems in Saami and Mordvin as well as in Hungarian) but it is retained most clearly in Mansi and Samoyed (Aikio 2013), as well as in many contexts in Khanty, Permic and Mari.

Thus, on the basis of, for example, Mari *šüδö* (PMA *\*šüdə*) and Mansi *\*šitV*, it is clear that this word should be reconstructed as *\*šita* in Proto-Uralic (there are also numerous other loans where PII *\*a* is reflected as *\*i* in loanwords). However, in the Permic languages, where the opposition between *\*a* and *\*i* is regularly retained in old *\*a*-stems, the vowels in the word '100' (Udmurt *śu*, Ko *śo* < PP *\*śo*) are irregular and probably point to a separate borrowing into Pre-Permic. On the basis of Finnic, Saami and Mordvin,

either *\*a* or *\*i* could be reconstructed, and Hungarian offers no evidence here, as *\*a* and *\*i* cannot be distinguished in the reflexes of PU *\*a*-stems in Hungarian. The situation is further complicated by the fact that the Permic vowel correspondence (Ud *u*, Ko *o*) usually reflects the old *\*e-ä* vowel sequence. So although here we seem to have a very early, Proto-Indo-Iranian borrowing, it looks likely that the word has to have been parallelly borrowed into early dialects of Proto-Uralic languages.

Fortunately, in a word like *\*š̌ita ~ \*š̌ata* there are other features that point to its old age (for instance, the regular development of *\*š̌* in Saami, Hungarian, Khanty and Mansi). But it is clear that distribution is not a good criterion to stratify the Uralic loanwords. It is also illustrative that there are a couple of very archaic-looking loans that are found just in Finnic, but then there are clearly later loans, such as *\*warsa*, that are shared by Finnic and Mordvin. This, again, shows that distribution is a bad criterion, as words get easily replaced by newer loans, and especially when dealing with prehistoric language contacts with a huge time depth, it becomes obvious that distribution as a criterion is obsolete.

Junttila (2015: 73–75) also discusses the need to distinguish parallel borrowings from derivatives: it is possible that a word has been borrowed several times, or that the same word exists in the language as several derived forms (or in related languages in forms that include various derivational suffixes, making the recognition of regular cognates difficult).

Interestingly, K. Häkkinen (1983: 207) argues that although words disappear from languages, the most basic words often remain stable and are maintained for longer periods. Although this is probably true, here the notion of “basicness” is something that is open to different interpretations. Many central concepts in culture and livelihoods are often described with prestige words that are borrowed, and these central words can be very easily replaced. In determining the age of the loanwords one has to always keep in mind that a reflex of a very early cultural borrowing from Indo-Iranian to Proto-Uralic/Proto-West Uralic etc. can easily have been lost in some daughter language, if a later prestige loan for the same concept has been borrowed from some later contact language (such as from some form of Germanic or Baltic into Finnic or from some Turkic language into Udmurt, Mari or Mordvin).

In Uralic linguistics the common loanword layers shared by some intermediary proto-language have often been seen as giving support to the reconstruction of these stages, but K. Häkkinen (100–108) considers this problematic. It should also be noted that the distribution of Indo-Iranian loanwords very rarely matches the assumed taxonomic divisions: there are some loanwords confined to the Finno-Permic, Finno-Volgaic or Ugric languages, but very few loanwords that would be Finno-Permic, Finno-Volgaic or Ugric in the way that the word is found in all the languages that belong to the branch.

In Uralic linguistics also the existence of early parallel borrowings has been long since noted, and this has been emphasized especially by Jorma Koivulehto (1990). In loanwords, sound correspondences that are irregular in relation to inherited vocabulary can point to parallel borrowings, although this is not always necessarily the case, as many early loans were probably borrowed into Proto-Uralic already, and these words display similar kinds of apparent irregularities as inherited Uralic words, irregularities that can often be solved with the discovery of new sound laws and conditioned developments.

#### 1.6.4. Semantics

In research on loanwords, semantics obviously plays an important role, as rarely are the meanings of the donor and recipient form exactly identical. However, the role of semantics can be seen as less central than that of phonological criteria described above: it is futile to search for Uralic and Indo-Iranian words with similar or identical meanings and then try to work out the substitution rules. Usually the method works the other way around.

Lexical typology and the study of semantic change has seen great leaps starting from the 1990s, thanks to such studies as Haspelmath & Tadmor (eds. 2008) and Koptjevskaja-Tamm (2008), but the results of semantic and lexical typology have not been utilized to the full extent in Uralic loanword research. The recent developments have also been included in useful databases, notably WOLD and DatSemShift. Even though their scope and material is far from perfect, they provide useful tools for loanword researchers, and they have been utilized in this work as far as possible.

It is often noted that compared to phonology, morphology and other areas of historical linguistics, semantics is a more challenging sub-discipline because it has no “laws” or absolute rules that would guide semantic development in a similar way (Campbell 1999: 267; Zimmer 2017: 76–77). Algeo (1990a: 399–401) notes that following the Neogrammarian model in phonology, attempts have been made to establish laws or universals of semantic developments (Algeo refers especially to Bréal 1883; see also Durkin 2009: 243–253). However, it remains uncertain how well these universals work in practice. Campbell (1999: 255, 267) notes that in semantic change one can assume tendencies; this seems a more fruitful approach than the quest for semantic “laws”. An important notion is the view that semantic change usually develops through the initial stage of polysemy, and reconstructing the earlier polysemy on the basis of the divergent meanings can help in tracking down the semantic developments. Furthermore, although one cannot establish precise rules for semantic developments, in order to argue for a particular semantic path in loanwords one should always be able to provide semantic parallels (Durkin 2009: 254).

Campbell (1999: 266) further notes that sociocultural historical facts are often relevant for semantic change, and because of this some have argued that generalizations are useless. However, some generalizations can and have to be made.

Despite the problems described above, one has to make use of the theory of linguistic semantics, which identifies various types of semantic correspondences. Algeo (1990a: 399–400) mentions the following kinds of changes:

1) substitution, i.e. “a change of a referent or in our knowledge of it or in our attitude towards it”: for example, ‘ship’ in modern English means a different kind of ship than the word’s predecessor denoted in Old English;

2) analogy (including popular etymology): a word changes its meaning because it resembles another word (here one could mention useful examples from the Uralic languages, such as Finnish *ahtaa* < PU \**ikta-*, which has changed its original meaning ‘to hang’ to ‘to stuff, squeeze into’ due to influence from the homonymous Baltic borrowing *ahdas* ‘narrow’ ← Baltic \**ankštas*).

3) shortening: English *the blue devils* ‘melancholy’ was shortened to *the blues*.

4) nomination: the intentional naming of a referent with a term not previously used for it.

5) transfer: the use of a word for a new referent that somehow is similar to older referents of the word (Algeo lists examples such as *sharp* taste, *sweet* sight).

6) permutation: a meaning change resulting from a shift of focus (English *want* ‘to lack’ > ‘to desire’).

7) adequation: a shift in which characteristics of the referent are considered central to the meaning of the word: generalization, restriction, amelioration and pejoration. Connected to this category are the different types of semantic change that are listed by Campbell (1999: 254–266): widening (in Durkin’s 2009 terminology *broadening*), narrowing, metaphor, metonymy, synecdoche (inclusion), degeneration (= pejoration), amelioration, taboo replacement, hyperbole (exaggeration) and litotes (exaggeration through understatement). Campbell also mentions semantic change due to contact: for example, the Lake Miwok word *kó:no* originally meant ‘bow’, but it has acquired the meaning of ‘gun’ after contact with European settlers who used guns.

It is easy to observe that not all of these categories are equally relevant in loanword studies. Various types of adequation take place regularly, but rarely is the entire “package” of meanings borrowed, so restriction or narrowing is very frequent. Some useful examples of this are listed by Anttila (1989: 155–156).

Among the Indo-Iranian loanwords in Uralic, one can easily mention examples that follow the classification of semantic change as described above. For widening or generalization, *vasara* ‘hammer’ is a good example: the Indo-Iranian original denotes ‘the weapon of the god Indra’ in Indic and ‘the weapon of the god Mithra’ in Avestan, whereas Finnish ‘hammer’ (and the Mordvin meaning ‘axe’) are more general meanings of tools. Fi *huhta* is a good example of narrowing: Iranian *\*tsuxta-* means simply ‘burned’, whereas in Finnic *huhta* means specifically ‘a burned patch used in slash-and-burn agriculture’. Metonymy has taken place in Mordvin, where *čuvto* denotes simply ‘tree’; this probably developed through the meaning ‘wood burned for agriculture’. Khanty (South) *wārəs* denotes ‘horse’s mane’, but its Iranian original probably had a more general meaning of hair (cf. Avestan *varəsa-* ‘hair of human and animal, mostly hair of the head’).

An interesting example of degeneration is the etymology of Finnic *orja* ‘slave’, probably borrowed from the Indo-Iranian ethnonym *\*(H)ārya-* ‘Aryan’ (for the original semantics of this word, see the entry *\*orja* in Chapter 2). A similar development is seen in English *slave* which is etymologically connected to the ethnonym *Slav*.

Campbell (1998: 279–280) also notes that obsolescence and taboo are important sources for semantic developments (cf. the discussion above in the context of lexical loss).

Regarding semantics, one also has to realize that when working with reconstructed vocabulary and especially with lexical contacts between two or more reconstructed and non-attested languages, one has to be especially careful. In this case *all* the semantics are reconstructed, and although they are based on the attested meanings from the daughter languages, it is always possible that something turns out to be wrong. The difficulties in dealing with reconstructed lexicon are emphasized by Ringe (2004: 1023–1024), who notes that the reconstructed proto-language lexicon is always incomplete, and semantic reconstruction is often challenging. On the Indo-Iranian side one can rely both on the evidence of the attested Indo-Aryan and Iranian languages as well as that of the cognates in other branches of the Indo-European family, but still the matter of semantic error is a considerable threat. Avestan and Vedic texts are hymns which employ rich poetic language, and not all of the meanings are well understood. Due to the very long and rich attestation

of Sanskrit, the meanings of the words abound. Especially the meanings of Classical Sanskrit words that might differ from the meanings known from the earlier Vedic texts have to be approached carefully.

## 1.7. SOUND SUBSTITUTIONS

The substitutions that have been suggested in the literature are briefly presented here. The table below is modelled after the presentation of substitution rules in the Germanic loanwords of Finnic in LÄGLOS I. Only the substitutions that are argued in recent works by Koivulehto (1999a&b, 2001a&b, 2007, 2009), Parpola (2010, 2015, 2017, 2019), J. Häkkinen (2007, 2009), Helimski (1997, 2000), Napol'skikh (2010, 2014), Aikio (2015b), Kallio (2009, 2015), Zhivlov (2013) are presented here, because in earlier works attention has not been paid to substitutions (Joki 1973; Rédei 1986) or the ideas of different donor languages (such as Munkácsi's 1903 layering of Iranian and "Indic" loans) is clearly obsolete. The substitutions suggested by Katz (1985; 2003) will be discussed individually in the context of each etymology.

The chronology and environments of these substitutions will be commented on in more detail in the analysis of the individual etymologies, and in the Results chapter it will be shown which substitutions can stand up to critical review and which have to be abandoned. Here the suggested factors causing the possible substitutions will be very briefly presented.

It is usually assumed (Joki 1973; Rédei 1986; Koivulehto 2016) that the loans which show Uralic *\*e* vocalism reflect the retained PIE *\*e*, not a substitution of later (PII) *\*a*. Regarding the *\*o* vocalism, this can reflect either the retained PIE *\*o* or a substitution of PII/PI *\*a* (a similar substitution is also found in the Germanic and Baltic loanwords of Finnic). Various environments for these "o-loans" have been suggested, including the substitution of *\*o* ← *\*a* in word-initial position (Minissi 1970; Gamkrelidze & Ivanov 1984; Lushnikova 1990), before *\*r* (Gamkrelidze & Ivanov 1984), or before *\*rC* (Lushnikova 1990), as well as before a nasal (Heikkilä 2014: 46). Also the labial environment (after *\*p* or *\*w*) has often been considered as a factor that causes the *\*o*-substitution. Distinguishing retained *\*o* from the *o*-substitution of *\*a* is a major task of Uralic and Indo-Iranian etymological research. The idea of *\*o* loans is being revived to some extent because of the new vowel reconstruction by Aikio (2015b), which shows that a part of the *\*o*–*a* words have to be reconstructed as *\*a*–*i*-stems.

For the substitution *\*i* ← *\*a*, only J. Häkkinen (2009) has attempted to establish clear conditions, assuming that the unaccented Indo-Iranian *\*a* was substituted by Uralic *\*i*. Aikio's new idea of the divergent development of *\*a*–*i* and *\*i*–*i*-stems in West Uralic also has an impact on the ideas regarding the *\*i* loans.

The substitutions of Indo-Iranian and Iranian long *\*ā* are largely the same as those of the short *\*a*. Although in Indo-Iranian there probably was a notable difference in the realization of the long and short *\*a*, with short *\*a* being rather *ə* (Lubotsky 2018: 1875; Oberlies 2017: 360), this difference does not manifest itself in loanwords. However, it is often problematic to determine whether the Uralic word has been borrowed from a form which had long *\*ā*, as especially in borrowed verbs it is very difficult to determine the precise donor form, due to the Indo-Iranian ablaut and Brugmann's law. Many of the long

vowels in Indo-Iranian also result from earlier *\*VH* sequences, and as a reflex of the laryngeals probably was in some form present in Proto-Indo-Iranian and Proto-Iranian, many of the alleged long-vowel donor forms can be reconstructed with *\*aH*.

The substitutions of Iranian *\*r̥* and *\*r̥H* are likewise diverse, but less attention has been paid to the conditions that cause the different substitutions here. Gamkrelidze & Ivanov (1984) and Lushnikova (1990) assume that *\*r̥* was usually substituted by Uralic *\*ar*, but only part of the material fits this rule, and numerous other substitutions are found, as can be seen from the table below. It is important to note in general that Joki (1973), Gamkrelidze & Ivanov (1984), Rédei (1986c) and Lushnikova (1990) do not reconstruct the opposition of *\*a* and *\*i* for Proto-Uralic, and a large number of cases that they consider *\*a* loans could in reality be reconstructed equally well with *\*i*.

Only the substitutions in the initial syllable of Uralic are presented here systematically because of the more limited vocalism of the unstressed non-initial syllables and because not much has been written about this earlier. The substitutions in non-initial syllables will be discussed in the results.

The substitution of consonants is less problematic. The main problems concern the possible substitutions of the Indo-Iranian reflex of the laryngeals, and the different substitutions of Iranian *\*ts*, *\*dz*.

The Proto-Uralic (PU) here denotes all the reconstructed stages (Finno-Permic, Finno-Ugric, Finno-Saamic, etc. in the traditional sense) unless noted otherwise.

Dubious cases (substitutions involving forms with controversial reconstructions, uncertain Indo-Iranian etymologies) are marked with “?”. In some cases it is impossible to show which vowel should be reconstructed for the proto-form, due to the vowel mergers that are shared by different branches. These ambiguous cases (such as PU *\*a* or *\*i*) are listed separately.

Following the practice of LÄGLOS (I: XVII), completely implausible etymologies or etymologies that show implausible substitutions are not listed in the table, but they will be commented on in the text.

### 1.7.1. Vowels

PII/PI <i>*a</i>	→ U <i>*a</i> or <i>*i</i>	<i>*ačwa</i> , ? <i>*ajsa</i> , ? <i>*ajwa</i> , ? <i>*apV</i> , <i>*arwa</i> , <i>*asma</i> , <i>*aśnas</i> , <i>*asora</i> , ? <i>*aštira</i> , <i>*jama</i> , <i>*čara-</i> , ? <i>*karšV</i> , <i>*kata-</i> , <i>*mača</i> , ? <i>*maks-</i> , ? <i>*mana-</i> , ? <i>*martas</i> , <i>*pakas</i> , <i>*para(s)</i> , <i>*śaḍa-</i> ? <i>saja-</i> ,? <i>*sammV-</i> , <i>*sampas</i> , <i>*sasarV</i> , <i>*śara-</i> , <i>*śasra</i> , ? <i>*śata</i> , <sup>9</sup> ? <i>*šaji-</i> , ? Pre-P <i>*šata</i> (> PP <i>*šud</i> ) ? <i>*šatas</i> , Pre-Md <i>*čaya</i> , <i>*tajwas</i> , ? <i>*tarwas</i> , <i>*wajna</i> , ? <i>*wan̄ka</i> , <i>*wasa</i>
	→ U <i>*a</i>	<i>*ajša</i> , ? <i>*čača-</i> , Pre-S <i>*jawī</i> (> PS <i>*jāā</i> ), Pre-P <i>*kada-</i> ‘to steal’, <i>*kana-</i> , <i>*kara-</i> ,

<sup>9</sup> Both *\*śata* and *\*šita* are listed because of the controversial vocalism of the words for ‘100’ in the various branches of Uralic. This question will be dealt with in greater detail in Chapter 2.

		Pre-P *karka, ? PUG *manći, *parta, sayi-, ? *śali, *sara, ? *talas, ? PUG *tala, ? Pre- Sa *waćim, *waćara, Pre-S *wakša-, ? *watsa
	U *a or *o	*anti/onta, *śami/śoma-, Pre-Md *jawi/jowa, *wajni/ojna-, Pre-Sa *pari/pora, Pre-Sa *sami/soma
	U *o or **i	*poši, *soŋka
	→ U *o	? *kota, POUG *meja (< ? *moja) ? *oćra, ? *omtVrV, ? *ońća, ? *oŋki, ? *orpa, Pre- P *poći, *poćaw, *pora, ? *porćas, ? *portta, Pre-Sa *pottik, *šoži-, ? *śolki, *śoma, *somas, PUG *sora, *śorsa, ? *wermen, *wojna-, *woraśi
	→ U *i	*jini, *kiŋtaŋ, ? *miyi, *piŋka, *śiŋta, PUG *sirańa
PII/PI *a	→ U *ä	PMS *ārV, ? *äkštärä, POUG *ćäkV, ? Pre- OUg *ärkä (> POUG *ēräy), ? *kärtänä, PKh *kärtV Pre-Ms *māńci, POUG *päčäy, Pre-Kh *päntä, Pre-Ma *pänti-, PMA *pärcä, *pärtä, *pätäri-, PMS *rāsVnV, PUG ?*śāŋkV, ? PUG *säptä, *täjV, *täjVnV, POUG *wāńV, POUG *wāŋkVrV, POUG *wārV, PKh *wār(ə)sV, *wärkä, Pre-P *āŋkVrV, PMS *ārV
PII *a < Pre-II *e	→ U *e	? Pre-Sa *ćerti, *ertä, ? *jewä, *kečrä, *kekrä, *kert(t)ä, *kert(t)ä- *mekši, ? *mertä, *meti, *peijas, *pejmä, *pertä, ? *rećmä, *repä, *sejti, ? *sentü-, ? *šerä-, ? *serä, ? *šewći-, ? *tejni(š), ? *ternV, *weni(š), ? *wersV
PII/PI *a < Pre-II *o	→ U *o	? *ońća, *orpa
PII/PI *ā	→ U *a	Pre-Ms *aćtira, ? POUG *ńātV-, PUG *taja, ? *takra, ? POUG *wātV-
	→ U *o	*ora, *orja, ? *tojwV-, ? *tora-
PII/PI *i, *ī	→ U *i	? *iä, ? *iäntä, *iša (*iša <sup>2</sup> ), *išta, ? *iso, ? *(j)išV, Pre-OUg *itra, ? *niska, *šišta, ? *šiwa, *wiša (*wiša <sup>2</sup> ), *wirä
PII/PI *u, *ū	→ U *u	*juma, ? Pre-Sa *kupa, *mura-, ? *puntas, *šukta, ? *śuka, *surV, ? PP *sumiš ? PP *sur, Pre-P *tura, *uška, ? *utarV, Pre- Mari *wuđarV
	? → U *ü	? *mükkä
PII/PI *r	→ U *ar	*śarwi
	→ U *ar or *ir	? *martas, *tarna, *warna

	→ U * <i>ir</i>	* <i>siṛna</i>
	→ U * <i>er</i>	* <i>kertä</i> , ? * <i>mertä</i> , Pre-Md * <i>sernä</i> , Pre-Md * <i>serä</i> , ? Pre-Md * <i>werkas</i> , ? Pre-FI * <i>wesšV</i>
	→ U * <i>ir</i>	* <i>wirta-</i>
	→ U * <i>är</i>	* <i>wärk(k)ä</i>
	→ U * <i>ür</i>	* <i>kürtńV</i>
	→ U * <i>ur</i>	? Pre-S <i>wurka</i> (> PS * <i>wā rkā</i> ^)
PII/PI > Iranian * <i>ər</i> ~ * <i>ar</i> * <i>r</i> <sub>o</sub>	→ U * <i>ar</i> or * <i>ir</i>	* <i>warsa</i>
PII/PI * <i>r</i> <sub>H</sub> (> ? PI * <i>ər</i> )	→ U * <i>or</i>	* <i>tora-</i> , * <i>toraksi</i>
	→ U * <i>er</i>	? Pre-Sa * <i>ćero</i>

### 1.7.2. Consonants

PII * <i>ć</i>	→ U * <i>ś</i> ~ <i>ć</i>	? * <i>aćnas</i> , ? * <i>ćara-</i> ‘brown; to dawn’, * <i>ćara-</i> ‘to shit’, ? Pre-Sa * <i>ćero</i> , Pre-Sa * <i>ćerti</i> , * <i>ońća</i> , ? * <i>porćas</i> , * <i>śaŋka</i> , * <i>śiṭa</i>
PII * <i>ǰ</i> , * <i>ǰ<sup>h</sup></i>	→ U * <i>ś</i> ~ <i>ć</i>	* <i>waćara</i> , * <i>woraći</i>
PII * <i>ć</i> > PI * <i>ts</i>	→ U * <i>š</i>	? * <i>šiwä</i> , * <i>šukta</i>
	→ U * <i>č</i>	* <i>mača</i> , * <i>počaw</i> , * <i>päčäy</i>
	? → U * <i>ks</i>	* <i>-teksä</i>
PII * <i>ǰ</i> , * <i>ǰ<sup>h</sup></i> > PI * <i>dz</i>	? → U * <i>s</i>	? * <i>sentü-</i> , ? * <i>serä</i> , ? * <i>sewči-</i>
	? → U * <i>č</i>	? * <i>čača-</i>
	→ U * <i>š</i>	? * <i>šatas</i>
	? → U * <i>ks</i>	* <i>maksa-</i> , * <i>paksu</i>
PII * <i>ts</i> > Iranian * <i>s</i>	→ U * <i>s</i>	? * <i>asma</i> , * <i>wasa</i>
PII * <i>ǰ</i> , * <i>ǰ<sup>h</sup></i> > Iranian * <i>z</i>	→ U * <i>s</i>	Pre-Md * <i>sernä</i> , PUG * <i>siṛna</i>
PII/PI * <i>č</i>	? → U * <i>ć</i>	POUG * <i>ćäkV</i>
PII/PI * <i>ǰ</i>	? → U * <i>š</i>	? * <i>šerä-</i>
PII/PI * <i>š</i> , * <i>ž</i>	→ U * <i>š</i>	* <i>aštira</i> , * <i>ašja</i> , * <i>išta</i> , *(j) <i>išV</i> , * <i>karšV</i> , * <i>karsi</i> , * <i>mekši</i> , Pre-S * <i>wakša-</i> , Pre-S * <i>täši</i> , * <i>wiša</i>
	? → U * <i>s</i>	* <i>kärsä</i> , * <i>niska</i>
	? → U * <i>ć</i>	? Pug * <i>mańćV</i> , PMA * <i>pärća</i>
PII * <i>š</i> > Alanic * <i>s</i>	→ U * <i>s</i>	? * <i>kärsä</i> , * <i>warsa</i>
PII/PI * <i>b</i> , * <i>b<sup>h</sup></i>	→ U * <i>p</i>	* <i>pakas</i> , ? * <i>piŋka</i> , * <i>puntas</i> , * <i>sampas</i>
PII/PI * <i>p</i>	→ U * <i>p</i>	* <i>apV</i> , * <i>parta</i> , * <i>pakas</i> , * <i>para(s)</i> , * <i>pänti-</i> , * <i>pärća</i> , PFI * <i>perna</i> , * <i>peträ</i> , * <i>počaw</i> , * <i>pora</i> , ? * <i>porćas</i> , * <i>puntas</i> , * <i>päčäy</i> , * <i>pärtä</i> , * <i>pottik</i> , * <i>repä</i>
PII/PI * <i>p</i>	? → U * <i>p</i>	? * <i>parta</i>

> Iranian *f		
PII/PI *g, *g <sup>h</sup>	→ U *k	? *kertä
	? → U *w	*arwa
PII/PI *g, *g <sup>h</sup> > PI *γ	? → U *w	*arwa
Pre-II/PII/PI *k	→ U *k	*kečrä, *kekrä, *kerttä, *pijka, *wärkä
PII/PI *k > PI *χ	→ U *k	*šukta
PII/PI *d, *d <sup>h</sup>	→ U *t	*takra, *tejniš, *tejV, *tejVnV, *tervV, ? *tora
PII/PI *t	→ U *t	*kertä, *martas, *mertä, *meti, *parta, *peträ, ? *port(t)a, *pärtä, *sejti, *sukta, *tjrna, *tora, *toraksi
PII/PI *t > PI *θ		*iträ (> POUg *eträ)
PII/PI *m	→ U *m	*juma, *mača, *maksä-, *mana, *martas, *mekši, *meti, *mertä, *mura-, *mökkä, *sampas
PII/PI *n	→ U *n	*mana, PFi *perna, *puntas, *pänti-, PMS *räsVnV
	→ U *ñ	? POUg *ñätV-, *oñca, ? PMS *pñjtV, Pre- Md *serñä, PUG *siraña
PII/PI *r	→ U *r	*arwa, ? *aštira, *asora, *ćero, *ćerti, *itra, ? *karšV, *kečrä, *kekrä, *kertä, *kärtä, *martas, *mertä, *mura-, *ora, *parta, PFi *perna, *pertä, ? *port(t)a, *repä, PMS *räsVnV, *šara-, PUG *siraña, *šorsa, *surV, *tjrna, *tora, *toraksi, *wersV, *wirä, *wärkä, *woraći, *worna, PS *wärkä, *wärkä
PII ? *l <sup>10</sup>	→ U *l	? *ola, ? *šali, ? *talas, ? PUG *talV
PII/PI *y [i]	→ U *j	*jama, *juma, *jini, *sejti
PII/PI *w [u]	→ U *w	*wačara, *wajna-, Pre-S *wakša-, *warsa, *wasa, POug *wātV-, *weniš, ? *werkas, *wersV, *wojna-, *woraći, *wiša, *wirä, *wirtä, *worna, POug *wānV, POug *wāñkVrV, POug *wārV, PKh *wār(ə)sV, PS *wärkä, *wärkä
PII/PI *H	→ U *γ	*sayi-
	→ U *k	? Pre-Fi *karšV
	→ U *Ø <sup>11</sup>	

<sup>10</sup> It is highly uncertain whether \*l was a separate phoneme in Proto-Indo-Iranian or Proto-Iranian, and there are no convincing loan etymologies that would support the reconstruction of \*l. See above for a discussion of the problem of the Indo-Iranian \*l > \*r change.

<sup>11</sup> The cases where a laryngeal was not substituted are too numerous to be listed here. The problem is that it is often very difficult to determine whether the laryngeal was still present at the time of borrowing. There

PII/(?PI) *s	→ U *s	*asora, *sayi-, *saŋka, *surV, *sejti, *säptä
	→ U *ś	*pośi

### 1.7.3. Word-initial consonant clusters

PII/PI *dy-	→ U *j	*juma
PII/PI *st	→ U *s	*sampas
PII *ść	→ U *ś ~ *ć	Pre-Md *ćaya, ? PFi *sooja ? Pre-Sa *ćerti
PII/PI *sk	→ U *k	*kiŋtaŋ

## 2. THE ETYMOLOGIES: A CRITICAL ANALYSIS

### 2.1. Introductory notes

In this chapter the etymological material is reviewed. The etymologies are presented in alphabetical order according to the Proto-Uralic reconstructions (here Proto-Uralic is meant simply in the sense of a phonological stage of reconstruction and it does not imply a Uralic distribution). New etymologies by the author are presented separately in Chapter 2.2.

The presentation of the etymological entries and the argumentation is largely based on the model of LÄGLOS: the data are presented, sources are given, earlier argumentation is reviewed and new arguments are presented, and at the end of each entry it is clearly stated whether a given word is an Indo-Iranian loanword or not, or whether this remains uncertain. It is also stated at the end of each entry if the words that are discussed are rather parallel loans, or if only some of the words that have been derived from the proto-language word are Indo-Iranian loans.

The following information is given in each entry:

The bibliographical sources that deal with the etymology are listed. Here priority is given to those sources that have appeared after 1973, as earlier bibliographical information can be easily found in the Wörterverzeichnis in Joki (1973). Of these earlier sources, only those are mentioned that are not mentioned by Joki, or include arguments that are crucial for the etymology. The first source where the etymology is suggested, is usually mentioned.

If the Uralic word is attested only in one branch of the family, a reconstruction of the lowest proto-language level (e.g. Proto-Finnic, Proto-Mordvinic etc.) is given, and a possible pre-form reflecting the Proto-Uralic reconstruction stage (for example: Proto-Finnic \*ahnas < Pre-Finnic/Early-Proto-Finnic \*ačnas). In the case of Hungarian, also dialectal as well as Old and Middle Hungarian forms are given in these cases when possible.

---

are no convincing cases of the substitution of the word-initial laryngeals, but because most of the words that begin with vowel originally began with a laryngeal, almost all of the cases of vowel-initial words should be listed here.

The Uralic words are usually shown in their nominative form. The Finnic *i*-stems discussed here mostly belong to the so-called old *e*-stems, and it is not separately marked that their oblique stem ends with *-e-* because this is basic knowledge in Uralic historical linguistics. For the Proto-Indo-European and Proto-Indo-Iranian/Iranian as well as Old Indic nouns the stems are given (following the practice of EWAia). Nominative forms and other inflected forms are given if they provide relevant information that is crucial for the etymology. The Avestan and Old Persian words are mostly given in the forms are given by EWAia.

For the Indo-Iranian words, the Pre-Indo-Iranian, Proto-Indo-Iranian or Proto-Iranian reconstruction (depending on the word) is given, as all relevant forms from the earliest attested languages (Vedic, Classical Sanskrit, Avestan, Old Persian, Middle Persian languages, Greek glosses of Scythian, Sarmatian or Alanic names). Also, relevant modern forms are given, especially from Ossetic; in loans that show Alanic features, specific detail is paid to the modern Ossetic forms. For those Indo-Iranian words which can be derived from Proto-Indo-European (as most of the words discussed here can), a Proto-Indo-European reconstruction is also given.

In cases of competing Uralic or Indo-Iranian reconstructions, explanations and references are given in favour of the more convincing option.

In cases of poorly attested forms or forms with dubious etymology, more specific information will be provided on the word's earliest attestation (this is crucial, as both the Indo-Aryan and Iranian languages have a huge corpus, which include also words without etymology, later borrowings and *hapax legomena* that are difficult to interpret).

## 2.2. INDO-IRANIAN LOANWORDS IN URALIC

**Ko, Ud *už* ‘stallion’ < PP \**už*** (Csúcs 2005: 392)

< **Pre-P \**ačwa* ~ \**ičwa***

← **PI \**átswa*- ‘horse’ > Avestan *aspa-* id.** (EWAia I: 139–140, s.v. *ásva-*)

(Koivulehto 1999a: 226 = 2016: 228)

This etymology is convincing but open to different interpretations, as Permian *ž* can also reflect PU \**-š-*. There is no need to reconstruct \**w*, as Napol'skikh (2010: 237) has pointed out: the reconstruction of *w* based solely on Iranian evidence, the Permian words could also reflect simply \**ača* or \**iča*. Permian *ž* can also reflect earlier \**š* in an intervocalic position. Also in addition to Proto-Iranian \**atswa*, a form with \**šw* could also reflect the \**šf* that is reconstructed for Pre-Ossetic (cf. Windfuhr 2009: 18–19). Mari *ožo* ‘horse’ (< PMa \**ožə*), which is connected to the Permian word by Koivulehto, has to be a loan from Permian according to Metsäranta (2012: 27–29). However, Metsäranta states that this can be the case only if the Permian original had an affricate.

As noted already by Koivulehto, the Permian vowel in the initial syllable can reflect either \**a* or \**i*; in \**-a*-stems, their reflexes are indistinguishable in modern Permian. Both vowels

are also frequent substitutions for Iranian and Indo-Iranian \*a, meaning that it is almost impossible to decide which vowel should be reconstructed for this word.

Cheung (2002: 245) has proposed another Iranian etymology, deriving the Permic word from a reflex of Iranian \*wɾ̥ʂan- (> Ossetic *wyrs* ‘stallion’) which can hardly be correct because the cluster \*rs would not yield Permic \*ʒ.

An Indo-Iranian etymology

**SaSk *vuä́d'd* ‘root (esp. a pine root gathered for extracting root fibers for rope-making)’ (< PSa \*oantē), P *âatties*, L *oattes* ‘primary root of a tree’, Sk *vuḗddes*, K *vuenḏes* T *viḗnḏes* ‘root (esp. of a tree)’ (< PSa \*oantēs : \*oantāse-); Md E *undoks*, M *unks* ‘root’ (< PMd \*-undaks). M *-unda* in *t’išä-unda* ‘plant, (medical) herb’ (< PMd \*-undā); Ma *oḏar* ‘поросль, побег; ветвь, ветвистый’ (< PМа \*oḏar); Ko *od* ‘весенняя зелень на лугах; всходы, росток’, Ud *ud-* in the compounds *ud-ñañ* ‘Wintersaat, Winterkorn’, *ud-murt* ‘Udmurt (ethnonym)’ < PP \*ɔd**

< PU \*anti or \*onta ‘grass’ (Aikio 2018: 8; UEW s.v. *antə* (*ontə*))

← PII \*(H)ándhas- > OI *ándhas-* ‘Kraut, Grün, Somakraut, Somatrank’; possibly from the PIE s-stem \*h<sub>2</sub>éndhos-, cf. Gr. *ἄνθος* ‘Blume, Blüte’, East Frisian *andel*, Old Frisian *āndul* ‘eine feine salzhaltige Graspflanze’ (? < PG \*andla-) (Nowicki 1976: 13; EWAia I: 79; Stüber 2002: 192–193; RIVELEX s.v. *ándhas*!-; Kroonen 2013: 27, s.v. *andla*-; Beekes 2010: 104–105, s.v. *ἄνθος*)

(Katz 1977; 1985: 269; 2003: 218; Rédei 1986c: 49–50; EWAia I: 79; Aikio 2018: 8)

This etymology was originally suggested by Katz (1977) and has in general been accepted since then, but some of the details presented by him and other older works are outdated by now. Katz (as well as the UEW and Rédei 1986c) considered the Uralic word to be confined to Permic and Mari, but Aikio (2018) has presented cognates in further Uralic languages (Saami and Mordvin). The old idea of a Mari-Permic word is found also in Bereczki (2013 s.v. *oḏar*), who does not mention the Indo-Iranian etymology. The wider distribution does not have much effect on the Indo-Iranian etymology itself, but the etymology is complicated because of the uncertain background of the Indo-Iranian word. Although phonologically and semantically plausible, the Indo-Iranian etymology of the Uralic word becomes less certain when one consults etymological sources of other Indo-European languages.

Although some standard etymological works, such as EWAia, consider the Indic word to be an inherited Indo-European word with cognates in Greek and Germanic, Beekes (2010) does not support the connection of the Greek and Indic words, and rules the Frisian cognate out for certainty. Kroonen (2013: 27) mentions the possible PIE etymology of the Frisian word but considers the etymology and connection with the Indic word to be semantically weak. Beekes also mentions Armenian *and* ‘field’ as a possible cognate of the Greek word, but Martirosyan (2010: 72–73, s.v. *and*) does not support the connection of

Armenian with these Indo-European words because also alternative, more convincing etymologies for the Armenian word exist.

Nowicki (1976: 13) connected the Armenian, Indo-Iranian and Greek words with Tocharian A *ānt*, B *ānte* ‘surface’, and this cognate set is mentioned also by Martirosyan, who however notes that these Tocharian words are derived rather from PIE *\*h<sub>2</sub>ent-* ‘front, forehead’. Frisk (1954–1972: 108–109) considers the relationship of the Indo-Iranian and Greek forms to be convincing, but rejects the comparison with the Albanian and Tocharian words presented in the IEW (4off.). Frisk mentions the possibility that Old Frisian *āndul* ‘Marschgras’ might be related. Chantraine (2009: 85–86) is sceptical of the etymological connection between the Greek and Indic words, due to the uncertain meaning of *ándhas-* (Chantraine makes no comment on a relationship to the other possible cognates within Indo-European).

Within Indo-Iranian, the noun *ándhas-* is securely attested only in Old Indo-Aryan, and it has no derivational relationship with any verbal forms (Nowciki 1976: 13; Stüber 2002: 192–193), which may indicate that it is an old word. EWAia also mentions Bailey’s idea that an Iranian cognate would have been attested in the Khotanese word *mījidā* ‘amaranthus plant’ (the latter part of this compound could reflect Iranian *\*anda-*). This is highly uncertain, however, as no further Iranian attestations exist and it is not at all obvious that the Khotanese word should be segmented like this. The hypothetical Proto-Iranian form would be phonologically equally good source for the Uralic borrowing, but because of the wide distribution within Uralic it is more likely that the word was borrowed from Proto-Indo-Iranian already, rather than from Proto-Iranian.

Regarding the semantics of the Indo-Aryan word, RIVELEX (s.v. *ándhas<sup>1-</sup>*) gives ‘shoot (of the Soma plant)’ as the basic meaning, and notes that metonymically the word is used for the Soma drink more generally. In the dual the Vedic word is also used to denote both *soma-* and *sūrā-* (a beer kind of drink) as well as *soma* and milk.

In a recent study, Pinault (2016: 90–91) refers to an old idea that was published by Narten (1986), influenced by her personal communication with Karl Hoffmann: the element *-ándhi-* attested in the compound *púramdhi-* ‘abundance’ is etymologically related to *ándhas-*. Pinault (2016: 90, footnote 33) supports the etymological connection of the Indo-Iranian and Greek forms, and he assumes that they are ultimately derived from a PIE root *\*h<sub>2</sub>en-* ‘to sprout, blossom, flourish’ with a *-dh-* enlargement.

Pinault (2016: 101) argues that *ándhas-* originally started with a laryngeal: this can be seen from the compound *púramdhi-* (< *\*p<sub>r̥</sub>-Hand<sup>h</sup>i-*, with regular development of PII *\*r<sub>h</sub>* to OI *ur* in a labial environment). The form *ándhi-* that is attested in this compound is an *i*-stem nominal abstract according to Pinault (2016: 94, footnote 48), not a Caland formation<sup>12</sup>. Pinault notes that this *i*-stem has no parallel in any other Indo-European language. The possible word-initial laryngeal is not substituted in this word in any way, and this is the case with most of the other loanwords.

According to Stüber (2002), the original meaning of the word in Indo-European has been ‘a plant, from which intoxicating drink is prepared’. The attested meaning of the word in Greek is ‘grape’, and according to Stüber this semantic development can be explained

---

<sup>12</sup> In the Caland system, which can be reconstructed for PIE and is reflected in several daughter-languages, among them Indo-Iranian, different nominal and verbal stems are derived from roots with adjective meanings (Bozzone 2016; Rau 2009; Caland 1892, 1893).

through the importance of the grape as the source of the most important intoxicating drink, wine. Because Stüber sees the notion of intoxicating drink as semantically central for the IE noun, she argues that the word cannot be derived from the IE root *\*h<sub>2</sub>andh-* which denotes ‘blooming’, contrary to what is claimed by LIV. However, the recent survey of PIE vocabulary by Kölligan (2018: 2243) considers this word as possibly the original PIE word for ‘flower’. Kölligan mentions the aforementioned Greek, Indic and Armenian words as well as Albanian *\*ëndë* as derived from PIE *\*h<sub>2</sub>endh<sup>os-</sup>*.

Dunkel (1995: 12) argues that the name of the Dionysiac wine-festival *Ἀνθεστήρια* points to the assumption that Greek *ἄνθος* originally denoted the soma plant (“a relic of the time when the intoxicating drink was not yet fermented grape juice but still the bitter extract of the stems of the psychedelic *\*andh<sup>os-ø</sup> par excellence*”). According to Dunkel, in Europe the *soma* drink was replaced first by mead, and then in the Mediterranean by wine. This is reflected by Greek *μέθυ* ‘wine’ and Luvian *maddu-* ‘wine’, words which originally denoted ‘honey’.

Janda (2000: 272–273) supports the idea of Dunkel (1995) that the Vedic word denotes precisely the soma herb and its juice, and that the Greek word means ‘Blüte der Weinreben’. Janda assumes that since the Indic and Greek words are closely connected to gods and rituals, the word had strong mythological connotations already in Proto-Indo-European.

The remarks above serve to illustrate that the exact Indo-European connections are open to different interpretations. Although the Indo-European background of the Indo-Iranian word is somewhat uncertain, the word does not appear in the list of substrate words in the Indo-Iranian languages compiled by Lubotsky (2001b). Many other words related to the *soma* rituals are found in the list.

Semantically, the Indo-Iranian etymology of the Uralic word is plausible, even if the precise semantic development in the Indo-Iranian side is partly obscure. The semantics of the Uralic word are reconstructed as ‘junges Gras, Schößling, Rasen [young grass, sprout, lawn]’ by the UEW, whereas Aikio (2018) reconstructs simply ‘grass’. The Sanskrit word has a variety of meanings, denoting both ‘sprout’ in a more general sense as well as ‘the sprout of the *soma* plant’. If the suggested Indo-European cognates of the Indic word are correct, the word’s original meaning was probably a more general ‘sprout’, and this fits the meaning of the Uralic words better. On the other hand, a basic word like this is more easily borrowed if it had more specific cultural significance and meaning in the donor language.

Bereczki (1994: 122) has argued that the Mari word is likely to be borrowed from Permic because it reflects the denasalization typical of Proto-Permic. However, Metsäranta (forthcoming) argues that denasalization of the cluster *\*nt* is regular in Mari *\*i*-stems (contrary to what was claimed by Bereczki, who assumed that denasalized forms are always borrowed from Permic), so the Mari word could also be inherited from Proto-Uralic, if the word is indeed an original *\*a-i*-stem.

Aikio (2018: 8) has presented cognates of the Permic and Mari word in Saami and Mordvin. In the UEW the Saami and Mordvin words are presented under a different entry (*\*ontə*), but because they can be regular cognates of the Permic and Mari words, Aikio assumes that they all reflect a single Proto-Uralic form. Aikio is uncertain of the reconstruction of the Proto-Uralic word, as all of the daughter languages could reflect either PU *\*anti* or *\*onta* – it is shown by Aikio (2015b: 29–39) that the reflexes of these two Proto-Uralic stem-types merged in Saami and Mordvin, and it is well-known that Mari *o*

can also reflect either PU *\*a* or *\*o*, regardless of the stem type. Either of the Uralic stem types would fit the Indo-Iranian origin, although the *\*a*-stem *\*onta* would be expected, as most of the loanwords were acquired as *\*a*-stems, but some loans were acquired as *\*i*-stems. Regarding the initial syllables, both *\*o* and *\*a* are well-known substitutions of PII *\*a*.

Even though the Indo-European background of the Indo-Iranian word remains uncertain, the Indo-Iranian etymology of the Uralic word can be accepted.

An Indo-Iranian etymology

Fi *ahma* ‘wolverine’, *ahmia* ‘to wolf down, eat quickly’, *ahmatti* ‘greedy, voracious’ (cognates in Karelian, Ludic and Estonian); SaN *vuosmmis* ‘the morbid attitude shown by some pregnant women towards food’, *vuosmmestit* ‘(of a pregnant woman) to detest certain dishes; to vomit’; cognates found in most Saami languages, missing from Kld, T, A (SSA s.v. *ahma*; Álgu s.v. *vuosmmis*)

< PU/PWU *\*asma-* or *\*isma-*

← PII *\*Haćman-* or PI (unattested) *\*Hatsman-*, cf. OI *aśman-* ‘eater’, root *HaćH-* > OI *aśi-* ‘sich nähren/sättigen, zu sich nehmen, essen’ < ? PIE *\*h<sub>2</sub>ek h<sub>3</sub>-* ‘(auf)essen’ (EWAia I: 136; Werba 1997: 274–275, No. 259; LIV<sup>2</sup>: 261–262; Mayrhofer 2005: 21; RIVELEX I: 587– s.v. *aśi-*)

(Schindler 1963: 205; SSA s.v. *ahma*; Aikio 2015b: 44; Holopainen 2018b: 151–152)

See also *\*aćnas*

The etymology of Schindler is convincing but has not received much discussion in the literature since its publication. SSA mentions the etymology only in the bibliographical list at the end of the entry *ahma*, which means that SSA does not consider it probable (this is not explicitly stated, however). The Indo-Iranian noun is derived from the root *\*HaćH-* ‘to eat’, whereas the Finnic noun *ahma* is etymologically opaque. The Indo-Iranian verb is probably derived from Indo-Iranian, even though RIVELEX considers this explanation as uncertain. Fi *ahmia* ‘to devour’ and *ahmatti* are derived from *ahma*. In the meaning of ‘wolverine’, *ahma* is probably an original taboo formation. Finnic *ahnas* ‘greedy’ has also been considered etymologically related.

The Saami word *vuosmmis* is a phonologically regular cognate to the Finnic stem, and also semantically it is possible to connect the two. The Saami word has a wide distribution, and phonologically it would be quite unlikely to assume that it could be a borrowing from Finnic, so there is no reason to assume that the Finnic and Saami word families are not cognates. Proto-Saami *\*āsvā* ‘wolverine’ (> North Saami *āspâs* + cognates in Sk, K, T, L, P, S) is a later borrowing from Finnic according to Lehtiranta (2001: 16).

Also, a possible Komi cognate has been suggested: in SSA the Komi verb *adžni* ‘to gulp down’ is mentioned with a question mark. Due to the vowel correspondences (Finnic *a* = Komi *a*) it looks very unlikely that the Komi word could be derived from the same PU word as the Finnic and Saami words, and also Komi *dž* as the reflex of PU *\*ś/ć* would be irregular. It is also not probable that Komi would have borrowed this word later from some Iranian language, as later Iranian languages would have had *s* in the place of

PII \*ć, and this cannot have resulted in the Komi affricate *dž*. The etymology of the Komi word remains unclear.

It seems probable that the Finno-Saamic word is of Indo-Iranian origin, as both semantically and phonologically the predecessor of OI *aśman-* is a perfectly suitable origin. The dating of the loan is difficult on the basis of Finnic historical phonology. It is difficult to determine whether this is a Proto-Indo-Iranian or Proto-Iranian loan, as Aikio (2015b: 44) has shown that the development of clusters of sibilants and nasals in Finnic is more complicated than was previously assumed, and that several possible predecessors can be posed for the Pre-Finnic form of *-hm-* here.

As noted by Holopainen (2018b), if the Finnic and Saami words are indeed cognates, this would mean that the Proto-Uralic/Early Proto-Finnic consonant cluster should be \*-*sm-* or \*-*šm-*, as \*-*ćm-* would not be reflected by *-sm-* in North Saami (Sammallahti 1998: 191, 200; Korhonen 1981: 176). \*-*ćm-* would result in North Saami *-šm-*. If the word had been borrowed from Proto-Indo-Iranian, we would expect \*-*ćm-* on the Uralic side. \*-*sm-* or \*-*šm-* point to later borrowing from Iranian, perhaps from Proto-Iranian, where the word would have been \*(*H*)*atsma-*.

It is interesting that the root \**HaćH-* is poorly attested in Iranian, as noted by de Vaan (2000), so we have no direct evidence for the reconstructed PI form \*(*H*)*atsman-*. This is, nevertheless, not an obstacle to derive this Finnic word from Iranian particularly, as the root was present in Proto-Indo-Iranian anyway, and it is plausible to think that it was retained until Proto-Iranian. (Another loanword connected to the same root could be Mansi \**šārkvšv* ‘eagle’, cf. Korenchy 1972: 69; the Indo-Iranian original, cf. Avestan *kahrkāsa-* ‘vulture’, is often explained as a compound ‘chicken-eater’, but Lubotsky (2001a: 315) considers this a folk etymology, and notes that the Ossetic form *cærgæs* ‘eagle’ – the most probable source of the Mansi word – cannot be regularly connected to the Iranian forms with *k*.)

The existence of the root \**HaćH-* in PII remains an open question, but one can assume that an exact cognate of the Indic noun could have existed in Proto-Iranian: *-man-* was still a productive suffix in Proto-Iranian (Schneider 2010), and the formation \**aćman-*, if formed already in Proto-Indo-Iranian, could have been inherited into Proto-Iranian. It is also well-known that the corpus of Old Iranian texts is very small compared to the massive corpus of Old Indic, and our knowledge of Old Iranian and the reconstructed Proto-Iranian lexicon has huge gaps because of that. But still, there is no way to prove that the root or this noun existed at the Proto-Iranian stage.

As is the case with most Indo-Iranian loanwords in Uralic, also this etymology shows no reflexes of the Indo-Iranian laryngeal in Uralic (cf. the entry \**anti* above). It is clear that the Proto-Indo-Iranian and Proto-Iranian forms still had some kind of reflex of the PIE laryngeals present (cf. Kümmel 2007: 430), but this is not reflected in most of the loanwords.

Regarding the vowel in the initial syllable, it remains unclear whether the Finnic and Saami words reflect earlier \**a* or the central-vowel \**i* in this word, as these two phonemes merged always in Finnic and in \*-*a-*stems in Saami. \**i* is found as the substitute of \**a* in a number of convincing Indo-Iranian loanwords, for example PU \**šīta* ‘hundred’ ← PII \**ćata-*.

Regarding the possible other etymologies for the Finno-Saamic word, none is convincing. Koponen (1999: 510–512) has derived the words *ahnas*, *ahne* ‘greedy’ and

*ahma* ‘wolverine’ as well as *ahvakka* ‘greedy, eager’ from the “onomatopoeic root *ahk-* ~ *ahv-*”, but this etymology can hardly be correct. This is based on an earlier idea by Nikkilä (1991: 4), who already connects *ahma* and *ahvakka*. Nikkilä’s etymology is also mentioned by Hahmo (1994: 64–65, footnote 39). Koponen’s root etymology is referred to in SSA. There are no such “roots” in Finnic, and the other words mentioned by Koponen are not likely to be related to *ahma*. Koponen seems to be unaware of Schindler’s Indo-Iranian etymology. While it remains unclear whence the word *ahvakka* comes from, the Indo-Iranian etymology is a much more credible solution than an obscure “onomatopoeic root” (for general criticism against onomatopoeic arguments and Koponen’s root etymologies, see Koivulehto 2001b: 59–63). Finnish *ahnas* (< Pre-Finnic *\*ačnas*) is clearly a parallel borrowing from Indo-Iranian *\*(H)ačnas*.

It remains enigmatic whether Finnish *osma*, *osmo* ‘wolverine; bear’ and its cognates in Karelian (*osma* ‘wolverine [?]’) and Votic (*ozmut* ‘small bear’) have something to do with the *ahma* words. In a personal communication Janne Saarikivi has suggested that Finnic *osma* is borrowed from Saami. According to SSA the word is attested also in Votic and is found even in place names in Estonia, which makes the borrowing from Saami even more unlikely for obvious geographical reasons. However, VKS does not mention the Votic word in this meaning but only gives the meanings ‘dear, good’. The word *osmo* also has the meaning ‘bridegroom’ in Finnish folk poetry, and SSA notes that this group of words might consist of originally unrelated words which have been contaminated by folk etymology. The Votic word is thus probably unrelated to Finnish *osma* ‘wolverine’.

On the other hand, *osma* ‘groom’ could be a similar taboo name for wolverine as *ahma*. It is notable that *osma* also denotes the bear, and it is well-known that in Finnic there are numerous taboo expressions for bear (expressions such as *metsän kuningas* ‘king of the forest’ and *metsän omena* ‘the apple of the forest’). It is not difficult to imagine a word denoting a male relative becoming an euphemism for beasts like the bear and wolverine. On the other hand, it has been suggested that *osma* is a contamination of *oksi* ‘bear’ and *ahma*. This is also possible, although the problem with explanations involving contamination is that it is very difficult to prove that they are correct.

An Indo-Iranian etymology

**Fi *ahnas*, *ahne* ‘greedy, voracious’ (< PFi *\*ahnas*, *\*ahneh*; cognates in all Finnic languages)**

< *\*ačnas* or *\*ačnas* or *\*ičnas* or *\*ičnas*

← PII *\*(H)ačHnas* > OI *ačnas* ‘voracious’; also a name of a demon; or ← PI (unattested) *\*(H)atsnas-*; PII root *\*HačH-* > OI *ači-* ‘sich nähren/sättigen, zu sich nehmen, essen’ < ? PIE *\*h<sub>2</sub>ek<sub>3</sub>-* ‘(auf)essen’ (EWAia I: 136, s.v. *ači-*; Werba 1997: 274–275, No. 259; LIV2: 261–262, *\*h<sub>2</sub>ek<sub>3</sub>-*; Mayrhofer 2003: 14; RIVELEX I s.v. *ášna*<sup>-1</sup>)

(Schindler 1963: 205; Koivulehto 1999a: 225; 2001b: 256 = 2016: 227, 293; SSA s.v. *ahnas*; Aikio 2015b)

See also *\*asma*

The Indo-Iranian etymology for this Finnic word is convincing (the semantics and phonology are both transparent). The etymology was first presented by Schindler (1963), and later discussed by Koivulehto. Initially, Koivulehto assumed that the Finnic word is borrowed from Germanic (1974: 123).

LÄGLOS discusses two possible Germanic etymologies for this word. The more convincing Germanic origin would be the etymology proposed by Koivulehto (1974: 123): PG *\*axna-z* (> ON substantivized adjective *agn* ‘Köder’): this is formally identical to the Indo-Iranian *\*ačnas-*. Koivulehto was inspired by the Indic cognate of the Germanic word, which is semantically closer to the Finnic word, but apparently he did not think it possible that a word confined to Finnic could be borrowed from Indo-Iranian. Later Koivulehto considered the Finnic word to be an Indo-Iranian loan. LÄGLOS considers the Germanic etymology possible but not certain (“? germ. Lehnwort”).

The other possible Germanic origin is considered less likely by LÄGLOS: Karsten (1943/44) has derived the Finnic word from PG *\*asnez/\*asnaz*, which is not attested properly but would be related to Gothic *asneis* ‘Tagelöhner’. Because it is uncertain whether the form *\*asnaz* ever existed in Proto-Germanic, this etymology cannot be considered a serious alternative to the Indo-Iranian etymology.

Although the etymology itself is convincing, the more precise dating of the loanword is open to different interpretations. Koivulehto’s arguments that the word has to be a loan from specifically Proto-Iranian *\*atsnas-* have been doubted by Aikio (2015b), who claims that modern Finnish *-hn-* could result from a number of different early Proto-Finnic consonant clusters, and thus both PII *\*ačnas-* and PI *\*atsnas-* would be suitable origins. Koivulehto reconstructs *\*ačnas* and derives the word from Proto-Iranian *\*asnas*, but according to Aikio, Proto-Indo-Iranian *\*ačnas-* (whence Pre-Finnic *\*ašnas*) would be equally likely.

The Finnic word has no cognates in the related languages, but this does not mean that the word could not be an early borrowing from Indo-Iranian as there are many other Indo-Iranian loans that are confined to Finnic alone. Distribution as a criterion has been frequently criticized (see the references in the Introduction). Finnish *ahne* (with derivational parallels in Karelian, Veps and Votic), another adjective with almost identical meaning, is probably a derivation from *ahnas* (*\*-eš* > Fi *-e(h)*) is a productive derivational suffix).

Fi *ahnas* is one of the loans in which the Indo-Iranian masculine nominative ending *\*-as* is probably reflected. Helimski (2000: 194–195) has noted that these loans do not belong to the oldest layer of Indo-Iranian borrowings in Uralic, as consonantal stems probably were not yet possible in Proto-Uralic times, and there are no loans resulting in consonantal stems that have a “Uralic” distribution. The late appearance of the consonantal stems on the Uralic side does not, of course, automatically imply that a loan has to be from a later Indo-Iranian language, but because Proto-Uralic and Proto-Indo-Iranian seem to have been roughly simultaneous languages, loans acquired after the split of Proto-Uralic were then probably borrowed from the Iranian branch. For many loans this can be quite convincingly shown, as Koivulehto (1999a; 2001b) has argued.

If this and other *\*-as* stem words are borrowed from Proto-Iranian, this gives external evidence that *\*-as* was still retained at the Proto-Iranian stage; in the attested Indo-Iranian languages, the word-final *-s* had already disappeared, and a similar development (visargation) happened before the oldest attestations in Indo-Aryan as well, as is well

known. However, here one should remember that we often have very little knowledge on which form exactly served as the basis for the loan, and both in Iranian and Indo-Aryan word-final -s survived much longer in various sandhi positions.

The same ending -as is also reflected in many Baltic and Balto-Slavic loans in Finnic, but there is no suitable origin for this word in the Balto-Slavic languages, as the root is not attested there. The root *\*h<sub>2</sub>ekh<sub>3</sub>-* is reflected by few Indo-European languages: LIV (261–262) mentions ON *ēja* ‘weiden lassen’ (< PG *\*ahjan*) in addition to the Indo-Aryan forms; a relation to Greek *ἄκολος* ‘bouchée’ is rejected by both LIV and Chantraine (2009: 46).

The words *ahma* ‘wolverine’ and *ahmatti* ‘glutton’ have been compared to *ahnas*, but these words most probably represent a parallel loan from Indo-Iranian (see above). While *ahne* and *ahnas* can be explained as derived in Finnic from the same source word, it is more economical to consider *ahma* (belonging together with *ahmatti*) a separate loan, as it suits Proto-Indo-Iranian *\*ácmas* perfectly. The words *ahnas* and *ahma* are thus etymologically connected, but through the (Indo-)Iranian side, where they are derived from the same root. The idea that *ahne* and *ahma* are derived from the same Finnic stem is originally an idea of Sivula (1986: 291), and it is also mentioned in LÄGLOS, but it is impossible to prove.

Koponen (1999) has suggested an onomatopoeic root etymology for *ahnas* and *ahne*. For criticism of this, see the entry for *asma* above.

An Indo-Iranian etymology

### **Mordvin *ažija*, *ažja* ‘Deichsel’; Ud *vajiž* ‘Femerstange, Gabeldeichsel’, Ko *vož* ‘Deichsel’**

< PU ? *\*ajša* (UEW s.v. *ajša*)

← PI(I) *\*Hay(H)ša-* > Av *aēša-* ‘plough?’, PIE *\*h<sub>2</sub>/<sub>3</sub>oy(H)s-o-*, cf. OI *īša-* ‘shaft’ < *\*h<sub>2</sub>/<sub>3</sub>i(H)s-éh<sub>2</sub>-* (EWAia I: 208; Peters 1980: 95; RIVELEX II s.v. *ī ša-*; Gotō 2013: 51)

(Toivonen 1928: 99; FUV: 129; Joki 1973: 253–254; Gamkrelidze & Ivanov 1984: 922; Katz 1985: 315–316; 2003: 252; Rédei 1986c: 49; Parpola 2007: 8)

This etymology was first suggested by Toivonen (1928: 99) and has generally been accepted since by the other sources listed above, but it is a rather complicated one. Koivulehto only mentions it briefly, and it has not been discussed critically in modern sources. The UEW simply mentions the Indo-Iranian etymology (borrowing from PII *\*Hayš(H)a-*) without analyzing it in any detail. The Uralic cognates feature irregularities, and it might be possible that some of the forms mentioned here have nothing to do either with the Iranian word or with each other. Sammallahti (1988: 552) lists the Mordvin and Permic words under a reconstruction *\*ooši* in his list of Finno-Permic words, but this reconstruction cannot be correct in light of modern views on Proto-Uralic vocalism. Aikio does not mention these words in his works on the development of Uralic back vowels. Here all the forms are analysed in detail:

The Mordvin words can regularly reflect earlier *\*ajša* (UEW), if a metathesis of -jš- is assumed. The UEW explains the Permic word-internal consonantism as *\*-jš- > \*-jž- > \*-již-*. The irregular vowel correspondences between the Komi and Udmurt forms makes the reconstruction of a common Proto-Permic form difficult and the etymology more

complicated. Probably the Mordvin and Permic words are separate borrowings from Iranian.

In addition to these forms, Finnic *aisa* ‘shaft’, sometimes connected to these words, is now usually considered a Baltic borrowing (Joki 1973: 254, 293); see below for a more comprehensive discussion of the Finnic word. Gamkrelidze & Ivanov (1984: 922) list “Proto-Finnic *\*ajša*” as a cognate to the Mordvin and Permic words, but this is probably a misunderstanding.

This etymology has in some sources (such as Gamkrelidze & Ivanov 1984) been considered specifically Iranian, which is mainly because of the *\*aj* diphthong (Indic reflects the zero grade) which is attested in Avestan *aēša-*. The meaning of the scarcely attested Avestan word is uncertain, and also its exact relationship to OI *īṣā-* ‘shaft’ (<*\*h<sub>2i</sub>-Hs-a-*) and its Indo-European cognates *οἴηξ* ~ *οἴᾶξ* ‘Griff des Steuerruders, Steuerruder’ and Hittite *hišša-* ‘shaft’ has remained uncertain (Melchert 2000: 235; Höfler 2017: 3, footnote 2). For details on the etymology of the Hittite and Greek words, see Kloekhorst (2008a: 403, s.v. <sup>GIS</sup>*hišša-* (c.)) and Beekes (2010: 1052, s.v. *οἴᾶξ*).

Joki has argued that borrowing from Proto-Indo-Iranian is more likely than a borrowing from Iranian because of the semantics: according to Joki the Avestan word denotes ‘plough’, while Joki assumes that the meaning of the OI word, ‘shaft’, can be reconstructed for Proto-Indo-Iranian. Because the meaning of the Avestan word is highly uncertain, Joki’s argumentation is not very convincing. On the other hand, since the meaning ‘shaft’ is also found in related languages like Hittite, this meaning can be assumed for the Proto-Indo-Iranian word as well. For the time being, it is difficult to determine whether the Mordvin and Permic words are borrowed from Proto-Iranian or Proto-Indo-Iranian.

Kloekhorst (403, s.v. <sup>GIS</sup>*hišša-* (c.)) mentions the usual cognates of the Hittite word in Indo-Aryan, Slavic and Greek, but does not mention or comment on the Avestan word. The PIE preform is reconstructed as *\*h<sub>2</sub>i<sub>h</sub><sub>1</sub>/3s-eh<sub>2</sub>-*, but Kloekhorst mentions that the pre-forms of the Slavic and Greek forms are less clear (Kloekhorst reconstructs an *s*-stem with *e-* or *o*-grade in the root).

Also Beekes (2010: 1052) reconstructs an *\*s*-stem from which the Greek and Slavic words have been derived and he considers the Indo-Aryan and Hittite forms later thematicized forms. The origin of the *\*o*-grade in Greek is unclear according to Beekes. No mention is made of the Avestan form, but Beekes also mentions the Baltic origin of Fi *aisa* and the Indo-Iranian origin of the more eastern Uralic forms.

Chantraine (2009: 750–751) mentions that Greek *οἴηξ* : *οἴᾶξ* is an *\*-āk-* derivation (the suffix is used to derive instrument nouns). Chantraine also mentions the unattested Baltic parallel *\*aisō* or *\*aisa* that is reflected by the loanword in Finnic. Also the Slavic (*s*-stem *\*oios/es-*) and Indo-Iranian cognates are mentioned, as well as the Hittite word.

An Indo-Iranian etymology (probably parallel borrowings in Permic and Mordvin)

### **Fi *aisa* ‘wagon shaft’, Veps (deriv.) *aižaz* id. (cognates in all Finnic languages)**

← PII or PI *\*(H)aiša-* > Av *aēša-* < PIE *\*h<sub>2</sub>/3oy(H)s-o* (See above for references)

(Mayrhofer 1964: 185–186; SSA s.v. *aisa*; Peters 1980: 95; Koivulehto 1991: 97–98, footnote 53; 2001a: 362, footnote 3 = 2016: 260, footnote 3; EES s.v. *ais*; Parpola 2007: 8)

Although references to the possible Indo-Iranian origin of the Finnic word can frequently be found in the literature (for example, Peters 1980: 95 notes that the Finnic word is borrowed from either Baltic or Iranian, and SSA mentions the Indo-Iranian origin as one possible etymology for the Finnic word), it is now universally accepted that the word is borrowed from some other Indo-European language, most probably Baltic or Balto-Slavic (Katz 1983a: 118; Junttila 2016; noted already by Lidén [1897: 60–63]).

Koivulehto (2001a: 362, footnote 3 = 2016: 260, footnote 3) mentions this among the possible borrowings which do not show Uralic \*š as the substitution of Indo-Iranian \*ś (which developed from \*s in the RUKI environment). Koivulehto argues that this can be because of phonotactic reasons, as he notes that the cluster \*-jś- might have been impossible in Pre-Finnic or Proto-Uralic (Koivulehto also notes that the metathesis of the Mordvin word *ažija* that was discussed above points to a similar conclusion: -jś- underwent metathesis to -šj- because of phonotactics). This is an unlikely possible explanation: the word can be explained as a Balto-Slavic borrowing, as is shown below, and so there is no need to assume such developments. Furthermore, Aikio (2015b: 57) has suggested at least one Uralic example of a Uralic \*-jś- cluster, namely \**kajšV* ‘misery’ (> Fi *kaiho*).

Parpola has recently argued that the Finnic word could nevertheless be borrowed from Indo-Iranian, and the seemingly absent RUKI reflex (Pre-Fi \*s instead of \*ś) has parallels in at least two assumed loanwords, \**kärsä* and \**niska*. However, it is shown here that both *kärsä* and *niska* are problematic etymologies (see the respective entries). The Baltic or Balto-Slavic origin is thus much more convincing.

However, Junttila (2012: 280) considers the Baltic borrowing dubious, because the cognate is missing from Baltic (it would be \**aisā* or \**aisa* according to him). Later Junttila (2016) has considered the word to be borrowed from a lost dialect of Balto-Slavic. Junttila (2016) reconstructs the Baltic word differently from earlier research as \**ajes*, corresponding to the Slavic *s*-stem \**oje*, and he assumes that the Finnic word is borrowed from the plural/collective \**ajesa* or a neuter form based on it (with regular contraction of -e-). Koivulehto (2001a: 362, footnote 3) has considered the Finnic word a possible loan from Slavic \**oje(s)* (see also the entry \**ojas* below), but because of the vowel \**a* it is more plausible to assume that the Finnic word is an earlier borrowing from Balto-Slavic or some lost dialect of Baltic. As there are other specifically Proto-Balto-Slavic borrowings in Finnic, *aisa* could well belong to the same layer of loanwords.

Not an Indo-Iranian etymology

### Fi *aivan* ‘whole, exact’ (also in Ka, Lu, Vo, Est) < PFi \**aiva-*

< PII \**aywa-* > Av *aēva* ‘one, only, lonely, some, OI *evá* ‘so, just’ < PIE \**óy-wo-* ‘*einzig, allein*’ (EWAia I: 270, s.v. *evá*; Rivelex II: 470–, s.v. *evá/evá*; Gotō 2013: 148; Dunkel 2014: 587–, s.v. \**oy-*)

(FUV: 129; Joki 1973: 247; Gamkrelidze & Ivanov 1984: 922; Katz 1985: 362; 2003: 287; Rédei 1986c; SSA s.v. *aiva*)

This etymology is semantically and phonologically convincing. The Indo-Iranian etymology of the Uralic word is also mentioned by Dunkel (2014), who refers to Katz (2003). According to Dunkel, PII *\*aywa-* (< PIE *\*oy-wo-*) has cognates in Greek and Tocharian. Because of the vocalism, only the Indo-Iranian and no other Indo-European donor form can come into question.

However, also a Germanic etymology for this Finnic word has been suggested: the Finnic word could be a borrowing from PG *\*aiwīna-* (> Gothic *aiweins* ‘eternal’) or PG *\*aiwan* (> ON *ey* ‘always’; LÄGLOS I: 18–19). Many of the earlier sources are undecided about whether the word should be considered an Indo-Iranian or a Germanic borrowing. Both the Indo-Iranian and Germanic origins are mentioned as possible etymologies by SSA. Also FUV considers the Finnic word either an Indo-Iranian or a Germanic loan. The Indo-Iranian etymology is phonologically acceptable, but semantically less transparent. Since the Indo-Iranian word has a variety of meanings, it is difficult to demonstrate that the meaning ‘so, just’ was present already in Proto-Indo-Iranian.

The word *\*aywa-* is confined to Indo-Iranian. EWAia is uncertain of the possible relationship of this word and the Indo-Iranian numerals ‘one’ (PIA *\*ayka-*, PI *\*aywa-*) which continue a PIE stem *\*(H)oy-* (EWAia I: 262–263, s.v. *éka-*).

In Finnish dialects and old written Finnish there is also a word *aiva* ‘ganz, völlig; nur; (obsolete) lauter, bloß, rein’. LÄGLOS refers to Hahmo (1988: 82, footnote 1), who considers the form without *-n* to represent a back formation. LÄGLOS also notes that the *-n* of the Finnic word has been explained as an adjectival suffix.

If the word is borrowed from Indo-Iranian, it is difficult to determine more precisely when the word was borrowed, because the Proto-Indo-Iranian form of this word is identical with later forms in Proto-Iranian, and the word would have been similar even in the unattested Old Iranian languages of the steppe. As is the case with many Indo-Iranian loanwords confined to Finnic, the word could have been borrowed either from Proto-Indo-Iranian or Proto-Iranian or even from a slightly later Iranian language.

? Indo-Iranian etymology (can also be a Germanic loan)

**Fi *ahtera* ‘barren, sterile’ (cognates in Vo, Est, Li, SEst), Md E *ekšt’er*, *jekšt’er*, *jakšt’er* M *jašt’ər***

< ? PU ? *\*äkštärä* ~ *\*akštira*

← PII (or PI) *\*ákšaytra* > OI *áksetra-* ‘destitute of fields, uncultivated’, root PII *\*éšay-* (? *tšay-*) ‘to live’ < PIE *\*tkey-* (Werba 1997: 172–173, No. 36)

(Blažek 1990: 40; Aikio 2015b: 44)

This is another one of Blažek’s etymologies that has been accepted by later researchers. Although the etymology is, in principle, convincing, it is noted by Aikio (personal communication) that the Finnic and Mordvin cognates are irregular, which could point to parallel borrowings into the two branches. However, the borrowing(s) have to be quite old

in any case, reflecting phonological developments from Pre- to Proto-Finnic and Pre- to Proto-Mordvin. There are not many examples of such words pointing to separate existence of Finnic and Mordvin as independent branches at the time of Indo-Iranian contacts. This would point to a relatively late, most probably Iranian borrowing. On phonological arguments it is difficult to say more about the time of borrowing.

A minor obstacle to Blažek's etymology is the second syllable and word-initial consonantism of the Finnic and Mordvin words. Because of the simple inventory of permitted non-initial-syllable vowels, it is not very problematic that the Indo-Iranian diphthong *\*ay* is reflected by a single vowel in the Uralic side, but it is more difficult to understand the metathesis that has occurred.

Because of these problems, it cannot be ruled out for certain that the resemblance of the Finnic, Mordvin and Indic words is only coincidental. Many West Uralic words relating to agriculture display irregular vowel correspondences and these could reflect substrate borrowings from some unknown language (Aikio 2015b: 44–47). Finnic *ahtera* and the Mordvin words would fit into this group of words well.

Munkácsi (1901: 238–289) suggested a different Indo-Iranian origin for the Uralic words, deriving them from PII *\*starī-* (> OI *starī-* 'cow that does not give milk', from PIE *\*sterih<sub>2</sub>-* 'sterile', EWAia II: 757). This idea is supported by Katz (1985: 235–236.) This etymology is much less convincing, as there are no examples of such prothetic vowels before word-initial consonant clusters in Indo-Iranian loans, and the sibilant *\*š* would be difficult to explain from plain *\*s*. Munkácsi also connects Hungarian *eszter* 'infertile, barren' to the Mordvin and Finnic words, but it is obvious that these three words cannot be cognates. Joki (1973: 80) mentions that the Hungarian word is borrowed from Romanian *știr* 'infertile', which is ultimately of Greek origin, borrowed into Romanian via Bulgarian. This Greek source, στειρα 'infertile' is a regular cognate of the Indic word, so ultimately the Hungarian word goes back to the Indo-European origin that Munkácsi had in mind.

Indo-Iranian etymology (probably parallel borrowings)

**Fi *apu* 'help', *auttaa*, *avittaa* 'to help', Est *abi* (has cognates in all Finnic languages) < PFi *\*api***

← PII (or PI) *\*(H)áwas-* 'help', root *\*(H)awH-* > Av *auuah-*, OI *ávas-* 'help' (EWAia I :132, 134; LIV2: 243–244, s.v. *h<sub>1</sub>euH-*)

(Katz 1985: 194; 2003: 158; Koivulehto 1999a: 228; Parpola 2019)

This etymology has been suggested independently by both Katz and Koivulehto and later discussed in detail by Parpola (2019). The Indo-Iranian origin is, in principle, convincing, but also a competing Germanic etymology exists (LÄGLOS I: 31 s.v. *apu*): PG *\*auja-/awi-* > ON *auja* 'luck; divine help, shelter', Gothic *awi-liudōn* 'to thank'; according to LÄGLOS the meaning of Runic *auja* is uncertain, but it is possible that it meant 'help' (Kroonen 2013: 43, s.v. *\*aupu-*). However, Koivulehto finds the Germanic origin much less convincing on semantic grounds (no meaning 'help' is securely attested for this root in Germanic, and also LÄGLOS notes that the meaning of Runic *auja* is uncertain).

In spite of this, LÄGLOS supports the Germanic etymology (Koivulehto’s Indo-Iranian etymology was not yet published when the first part of LÄGLOS appeared, so this possibility is not commented on in the book).

In any case, the nominative *apu* is explained as a result of the “analogical strong grade” in Finnic according to Koivulehto (the phonetically regular form in the modern Finnic languages would be *\*awV*). The second syllable *u* in Finnish *apu* is obviously secondary.

? Indo-Iranian etymology (can also be a Germanic loan)

**Mansi North *âr*, West *oâr* ‘relative from mother’s side’; Khanty North *worti* ‘younger brother of mother’; Hungarian *ara* ‘bride; mother-in-law’**

< ? PUg *\*arV*

← ? Middle Iranian/Pre-Ossetic *\*wra-/\*awra-*, cf. Ossetic *ærwada*, Pamir *wrod* ‘brother’, etc. < PI *\*braHter-* < PIE *\*breh<sub>2</sub>ter-*

(Korenchy 1969; 1972: 47–52 No. 2; Helimski 1997b: 121; Milanova, Holopainen & Bradley [forthcoming])

The assessment of this etymology is complicated because of the irregular vowel correspondences and the difficult semantic developments. Also the relationship of the words in the three Ugric languages is highly uncertain. According to DEWOS (181), the Khanty form can only be related to the Mansi and Hungarian forms if it includes a suffix *-ti*. UEW mentions the Ugric etymology but considers it uncertain. The Ob-Ugric words are not mentioned or commented on by Honti (1982) or Zhivlov (2006). Helimski (1997b: 121) mentions this etymology as problematic.

Problematic is Korenchy’s assumption that the word is a Middle Iranian loanword into the Ugric proto-language. This is an anachronistic idea, as the Middle Iranian period did not start before the last half of the last millennium BC (Windfuhr 2009), and there is no way to date the split of Proto-Ugric, if it ever existed, to later than that (500 BC is a dating that one often finds in earlier literature, but even this has to be way too recent, as is noted by Aikio [2014c: 30]).

The Middle Iranian etymology itself is not impossible but it includes certain difficulties. We should suppose that the three Ugric languages borrowed the word separately from either the same or different Iranian languages. The origin can hardly be from an older Iranian language, as the Iranian words cited here reflect Proto-Iranian *\*braHtar-* (< PIE *\*b<sup>h</sup>reh<sub>2</sub>ter-*). The Ossetic word shows metathesis of *wr* (from *\*br*) to *rw*, but the Ugric words could be borrowed from forms containing either *\*rw* or *\*wr*. Metathesis of such clusters is typical of Ob-Ugric, so either of the forms would probably result in the modern Ugric forms. However, one could expect some trace of the labial glide in the case of *\*rw* (compare the reflexes of Uralic *\*śarwi* ‘horn’), but *\*wr* would have probably been simplified to merely *\*r*. This does not help much in determining the exact donor language, but this means that it does not have to be “Alanic”.

The etymology is also uncertain because of the semantic differences. The attested meanings in the three Ugric languages are very different, and the only common factor is

that all denote some kind of kinship terms. Korenchy attempts to explain the semantic developments in Ob-Ugrian at length. Probably the two Ob-Ugric words are somehow connected, but their relationship to the Hungarian word and the possible Iranian original is everything but certain.

Not an Indo-Iranian etymology

**Fi *arvo* ‘value, price’ (cognates in all Finnic languages); Hu *ár* ‘price’**

< PU *\*arwV* ~ *\*iṛwV* ?

← PII *\*argh<sup>h</sup>á-* > OI *arghá-* ‘price’, Ossetic *ary*, Sogdian *’ry*, Bactrian *apyo* ‘value’ (EWAia I: 114, s.v. *arghá-*; Sims-Williams 2000: 180)

(Munkácsi 1901: 136–137; Jokl 1921; FUV: 129; KESK: 34; Katz 1971: 145; Joki 1973: 251–252; Gamkrelidze & Ivanov 1984: 925, footnote 1; Katz 1985: 287; 2003: 229; SSA s.v. *arvo*; Lushnikova 1990: 31; Parpola 2012: 161; 2015: 66; Holopainen, Junttila & Kuokkala 2017: 116; Aikio 2018)

This etymology was established already by Munkácsi (1901) and it has been widely accepted in the subsequent research literature (Koivulehto 1999a: 218 mentions the etymology among the widespread early borrowings but does not comment on it in further detail). The Finnic and Hungarian words can be regular cognates and there is no obstacle to derive them from a PU form (see below for a more detailed discussion of this). However, there are a couple of phonological issues that have to be discussed. Here the Finnic *-o* does not result from any transparent derivational suffix. Since a cognate is found only in Hungarian, no evidence for the non-initial syllable is found outside Finnic. Here the labial vowel cannot reflect any Indo-Iranian labial vowel, but in theory it could be a substitute for PII *\*a*, as this substitution is found in initial syllables in words such as PU *\*ora* < PII *\*āṛā-*. However, this idea is highly tentative, as there are very few possible parallels to this substitution in non-initial syllables. Aikio also notes that in the Finnic derivative *arvata-* ‘to guess’ the original stem vowel *\*a* is retained.

It has to be stated that theoretically the Iranian word could have been borrowed into Uralic also as *\*iṛwV* – the substitution of *\*a* by *\*i* is found in a number of loans, and the opposition of PU *\*a* and *\*i* becomes neutralized in Finnic, and also in Hungarian in original *\*i-a*-stems.

The word-internal consonant cluster is slightly problematic, as one would expect *\*rk* to be the substitute of PII *\*rg<sup>h</sup>*. However, as there naturally were no aspirated stops in PU, substituting this alien sound *\*rg<sup>h</sup>* with *\*rw* might not be unthinkable. Jokl (1921) has suggested that the *\*w* in the Uralic word might have resulted from a labiovelar stop. While this substitution would make phonological sense, there is hardly any sense in supposing a retained labiovelar stop in an Indo-Iranian loanword, as plain velars and labiovelars were clearly merged very early in Indo-Iranian like in the other satem languages (Edel’man 1986: 39), and other loanwords such as *\*kekrä* (< Pre-II *\*kekro-*) also show this development.

In spite of the unlikeliness of this substitution, it has received widespread attention. Mayrhofer (1952) supports the idea of Jokl. He also notes that the change *\*l* to *\*r* has to be older than the delabialization, but this cannot be true. Other loans like *\*kekrä* point to the direction that both changes were early, but as there are no traces of labiovelars anywhere in Indo-Iranian, and since the change has to be very early as it unites the all *satem* languages, it is improbable that a labiovelar could have been retained at this point. The possibility that Uralic *\*w* reflects a labiovelar is hinted at also by Dolgopolsky (1989: 19–20), although without references to earlier works. Katz (1985, 2003) assumes a similar substitution and it is also mentioned by Gamkrelidze & Ivanov (1984). It is strange that Gamkrelidze & Ivanov can support the idea that a labiovelar was retained in this word at the time of borrowing, when they at the same time consider this word an *Iranian* borrowing. Lushnikova (1990: 31) likewise supports the idea of a labiovelar, but considers this word a very old loan, borrowed from Pre-Indo-Iranian or dialectal Indo-European.

Much more convincing is the idea of Aikio (2018), who considers the Finnic and Hungarian words later parallel borrowings. In later Iranian languages *\*rg<sup>h</sup>* developed to *\*rɣ*, but one could hardly assume a “Middle Iranian” loan to a common proto-language of Finnic and Hungarian. In theory one could suppose a parallel loan from Iranian to Hungarian and Finnic, if the development in Ossetic is early enough to suppose a borrowing from Alanic/Pre-Ossetic to Pre-Finnic/West Uralic. Also the scarce attestation of the word in two distantly related branches could point to parallel borrowing in Aikio’s view), and at least the Finnic word could in his view reflect the Pre-Ossetic form *\*arɣa-*. Aikio also states that the Hungarian word could even reflect a PU/Pre-Hu shape *\*arka* regularly, but it is of course impossible to derive the Finnic word from such a proto-form.

Gamkrelidze & Ivanov (1984: 925, footnote 1) considered this word as a counterexample to the substitution rule that Indo-Iranian *\*a* is substituted by *\*o* in front of *\*r*. If the Finnic and Hungarian words reflect later parallel loans, this problem can be solved by assuming that this rule was no longer productive by the time the later Alanic-type words were borrowed.

As a side note to the Hungarian consonantism, it can be stated that *\*-rk-*, *\*-lk-* clusters often result in geminate *-ll-*, *-rr-* in modern Hungarian (possibly through a common Ugric metathesis *\*-kL-*). It is however disputed whether this is the regular development of such clusters, as A. Molnár (2003) has argued that all geminate reflexes can be argued as secondary, since single consonantal reflexes are attested in Old Hungarian. The Hungarian word *ár* has also variants *árr* and *áru*, of which the latter is probably a derivative but the form *árr* is more enigmatic.

Gamkrelidze & Ivanov (1984: 925, footnote 1) note that the Uralic forms point to a donor form which reflects PII *\*rg<sup>h</sup>* rather than *\*rdž* that is attested in Avestan *arəja-*, cf. Sogdian *ʾrɣ /arɣ/*. A parallel to the retained *g* is provided by Armenian *yargem*, which is borrowed from the same Iranian source. From this it can be assumed that the Uralic word is borrowed from an Eastern Iranian language akin to Sogdian and Alanic.

The Komi words *art* ‘sense, manner, order, series’, *artav-* ‘count, consider, suppose’ has been often connected to the Finnic and Hungarian words in older etymological sources, but the vowel correspondences are totally irregular here, so this “cognate” has to be rejected. According to Aikio, the Komi word is an independent Indo-Iranian loanword from a different source, PII *\*art<sup>h</sup>a-* (> OI *ártha-* ‘aim, purpose, matter, business’). Aikio’s

etymology is convincing. Already Katz (1971: 144) suggested an Iranian etymology for the Komi word, but derived it from PII \**r̥ta-* ‘right, true’ (cf. OI *r̥tam-*, Av *aša-*).

Mordvin E *art’šems*, M *aršəms* ‘to think’ is also listed as a cognate in the UEW, but any etymological connection to the Finnic and Hungarian words is highly unlikely. The cognate is listed as uncertain by Aikio (2018), who further mentions that the Mordvin word can perhaps be segmented as containing a stem \**ar-* and the frequentative-continuative suffix \*-*šə-*, but this cannot be determined for sure.

An Indo-Iranian etymology (parallel borrowings)

**Md E *azor*, *azoro*, M *azor* ‘lord’; Ud *uziṛ*, Ko *ozīṛ* ‘rich’ (< PP \**ozīṛ*); Ms South *ātār*, East, West *ōtār*, North *ōtār*, *ṣtār* ‘lord’ < PMS \**ātār***

< ? PU \**asora* ‘lord’

← PII \**Hásura-* > OI (Vedic) *ásura-* ‘god, master; name for a group of gods; later books of the Rig Veda: demon’, Av *ahura (mazda)*, OP *aura (mazda)* ‘God’, Pahlavi *ōrmazd* id.; < PIE \**h₂m̥su-* (EWAia I: 147–8; Ofitsch 1995: 156; LIV² 269 s.v. \**h₂ens-* ‘zeugen, gebären’; RIVELEX I s.v. *ásura*; Kloekhorst 2008a: 382, s.v. *ḥaššu-*)

(Munkácsi 1901: 26–7; FUV: 129; KESK: 203; Joki 1973: 251–2; Katz 1985: 207–208; 2003: 169; Rédei 1986c: 44; UEW No. 29; Koivulehto 2001b: 247 = 2016: 289; Blažek 2005: 162; J. Häkkinen 2012: 5; Parpolá 2015: 66; Aikio 2015b: 37–39; 2018; Holopainen, Junttila & Kuokkala 2017: 116–117)

This is an old and well-known Indo-Iranian loan (first presented by Munkácsi). The Indo-Iranian word is derived from a root that has cognates in many other Indo-European languages, but a suitable trisyllabic donor form is found only in Indo-Iranian (also the geographical reasons point towards Indo-Iranian as the donor language). The etymology is phonologically and semantically transparent, but the reconstruction of the second syllable of the Uralic word remains disputed (the UEW does not take a clear stance in this, and Koivulehto reconstructs \**asira-*). Aikio (2015b: 37–39) reconstructs it as \**-o*, basing this on the evidence of Mansi vowel development which corresponds to those of other tentative \**-o*-stems, and also on the PII \**-u*; Aikio considers the substitution \**o* ← \**u* to be natural, as he does not consider it possible to reconstruct \**u* for non-initial syllables of Proto-Uralic, and thus \**o* would be the most fitting substitution for Indo-Iranian \**u* in this position. One could of course ask why the vowel has to be reconstructed as \**-o* and not \**-u*, but as the Saami and Finnic words, the “key languages” for this reconstruction show (\**-o*), this is maybe the most fitting choice.

Aikio’s arguments are plausible and this is the only etymology in which Uralic \**o* (as reconstructed by Aikio) in the non-initial syllable corresponds to any labial vowel in an Indo-Iranian loan. However, the problem is that the number of the postulated \**-o*-stems is very low in general, and the other possible \**-o*-stem loans from Indo-Iranian are clearly poorer etymologies (and in these cases the \**o* does not correspond to Indo-Iranian \**o* but has to originate from some other reason).

If one does not accept Aikio's new reconstruction, there is no reason to reconstruct a common proto-form for the Mordvin, Permic and Mansi words. Mansi points to an *\*a-i-* stem, while Mordvin points to *\*a-a*. So while the PII etymology is totally convincing, it remains unclear whether we are dealing with a parallel loan or not. Aikio (2018) also mentions the possibility of parallel borrowing. A cultural word like this is easily borrowed, so an early parallel borrowing is not unthinkable. However, the parallel development would have to be very early because the reflexes of the word have gone through all the sound changes leading from Proto-Uralic to Mansi, Permic, and Mordvin. The situation is similar as with the word *\*ćata* '100' which likewise shows abnormal vocalism, even though the borrowing has to be very early.

Schmid has argued that the loan in Mordvin is late, borrowed in the 6th century BC. This claim is based on the view that the change *\*s > \*h* is late, but this does not automatically prove that the loan has to be borrowed towards the end of Proto-Iranian period. It can equally well be early, and as the Mordvin word has cognates in Permic and Mansi, the loan has to be very early and Schmid's argumentation has to clearly be considered erroneous. Unfortunately, Bartens (1999: 51) mentions Schmid's obsolete dating without commenting on it in any detail.

The retained *\*s* has traditionally been seen as a Proto-Indo-Iranian feature, but Korenchy (1972) has argued that the word is from Proto-Iranian, as *\*-s* was most probably still retained at the earliest stage of Iranian (see Mayrhofer 1989 and Chapter 1 of this dissertation). There is no possibility to decide whether this is a PII or PI borrowing. The wide distribution points to an early origin, and thus PII is a more likely source. Some clearly Iranian loans went through the Ugric/Ugro-Samoyedic sibilant changes, so the word could have been borrowed only from Proto-Iranian, too, especially if the Uralic forms represent parallel borrowings. Also Joki (1973) and Gamkrelidze & Ivanov (1984: 923) assume that this is a Proto-Iranian borrowing, although the view of the latter authors does not have much value since they consider almost every loanword Proto-Iranian because of their peculiar views of Eurasian linguistic prehistory, which do not allow for contacts between Proto-Indo-Iranian and Proto-Uralic/Finno-Ugric.

The Indo-Iranian word goes back to a form with initial laryngeal (this is proved by the Hittite cognate *ḫaššu-* 'king'), but as is the case with most of the loanwords, no trace of the laryngeal is found in the loanword. The laryngeals were certainly still present in some form in Proto-Indo-Iranian, as the laryngeal hiatus of Avestan and Vedic proves that the laryngeals were not lost very long before the codification of the earliest texts. Probably the laryngeal was a 'weak' sound, a glottal stop or the like, already in Proto-Indo-Iranian, at least in this position. Also Katz (2003) assumes that the absence of the laryngeal can be explained like this.

An Indo-Iranian etymology

**Ko *buríś*, *burśi* 'horse's mane'**

< PP *\*buríś* (Csúcs 2005: 323)

< Pre-P ? *\*parśa* ~ *\*barśa*

← Alanian *\*barci-* or PI *\*barša-*; YAv *barəša-* ‘throat, neck of a horse’ < PI *\*barša-*, MP *buš*, Oss *barc*, *barcæ* ‘horse’s mane’ < PI *\*br̥šti-* (Cheung 2002: 55, 173)

(KESK: 42; Joki 1973: 256; Rédei 1986c: 66; Lushnikova 1990: 192; Csúcs 2005: 323)

The Iranian etymology of this Permic word is obvious, even though some details require comment. Although Rédei lists the loans confined to Permic as one group, he still mentions different stages of Iranian as donor forms: this is from “Old Iranian” according to Rédei. He is clearly on the right track here, as because of Komi *u* corresponding to Iranian *\*a* the word has to be quite early, like several other loans that are found only in Permic (see, for example PP *\*gu-*, *\*warn-*). Despite the word’s occurrence only in Komi, the word thus has to go back at least to Pre-Permic level. In Udmurt the word has probably been replaced by some later borrowing.

Also from the Iranian point of view the word has to be relatively early: in most attested Middle Iranian forms the *-rš-* cluster has assimilated to simple *-š-*. However, in Ossetic the reflex is *-rc-*, which could also have produced the Permic forms, meaning that this could be a very early borrowing from an Alanic-type language. Note that the (Old) Iranian reconstruction *\*barša-* that appears in Rédei 1986c and Csúcs 2005 is not the pre-form of the Ossetic word: Cheung explains the Ossetic form from earlier *\*br̥šti-*, and the affricate results from the regular sound change *\*ti > ci*. The Permic sibilant *ś* is ambiguous in this respect: Joki and Rédei (referring to Uotila) note that because of folk etymology, the word has been reanalysed as a compound *bur-śi* ‘good-hair’, and the sibilant is due to the false connection with this word *śi* ‘hair’. This is a plausible explanation. Another possibility would be to assume that the cluster *-rc-* cluster was not possible at the time when the word was borrowed, and *\*-rś-* would simply reflect a sound substitution here. If the word is not borrowed from Alanic but from an Avestan-type language with a form *\*barša*, then it becomes more difficult to explain the substitution. Whatever the case may be, this is no obstacle to the etymology.

The word-initial *b-* appears in some Permic loans that have to be very early (Pre-Permic), including Indo-Iranian loans which have *p-* in the donor form (see *\*pärtä*, *\*parta*), and in at least some inherited Uralic words, which means that the voicing has to be secondary. It is thus not very helpful in dating the etymology, but it does not prevent us from considering this borrowing as rather early.

Semantically the etymology is obvious. Several other loans relating to horse breeding or horse pastoralism were borrowed into Uralic from Indo-Iranian, and there are also words confined to Permic that belong to this group (notably *\*ačwa* ‘stallion’).

An Indo-Iranian etymology

Fi *sata-* ‘yield crop’, *sato* ‘crops’ (cognates in all Finnic languages); SaN *čuoht* ‘to hit (a spot); to snow’ (cognates in all Saami languages, < PSa *\*čuocē*); Md E *čačo-*; Ma *šoć a-*, *šača-* (< PMa *\*šāča-*); Ko *čuž-* ‘be born, grow’, Ud *čži-viži* ‘relative’ (< PP *\*čuž-*); Khanty East *čičim* ‘bear cub’ (< PKh *\*čičim*); Mansi North *sūs-*, *sūns-* ‘grow, increase; have cubs (of a bear)’ (< PMs *\*sūns-*); Selkup *čāži* ‘family, tribe’ < PS *\*caci*

< PU \*čāča- ‘to grow’ (Zhivlov 2014: 127; Aikio 2015b: 14–17; Sammallahti 1988: 552; UEW s.v. čáčǝ-, čāčǝ-) >

← PI \*dzadzanh-a- (root \*dzanh- [Cheung 2007 \*zanH]) < PII \*ǝanh- < PIE \*ǝnh<sub>1</sub>- ‘erzeugen’ > reduplicated present \*ǝi-ǝnh<sub>1</sub>,- ǝnh<sub>1</sub>- > YAv *zīzanənti* ‘sie erzeugen’, present stem \*ǝnh<sub>1</sub>-e- Vedic *janati* ‘(er)zeugt, gebiert’; perfect \*ǝǝnh<sub>1</sub>/ǝnh<sub>1</sub>- > Vedic *jajāna-* ‘wird geboren’ (Werba 1997: 288–289, No. 286; LIV<sup>2</sup>: 163–165, s.v. \*ǝnh<sub>1</sub>-)

(Koivulehto 1999a: 226 = 2016: 228)

This etymology was presented by Koivulehto and it has not been commented on in etymological dictionaries. Although Koivulehto’s etymology was promising in its original form, new Uralic sound laws and the relationship to the Samoyed cognates (Zhivlov 2014; Aikio 2015a) have made this Iranian etymology less credible. Assessing the etymology is difficult because the cognate set of the Uralic words is presented in widely different forms in earlier sources. The most convincing Uralic reconstruction and cognate set presented by Zhivlov and Aikio gives the Proto-Uralic form as \*čāča- ‘to grow’ and presents reflexes in all branches of Uralic except for Hungarian. This is in stark contrast to the earlier distribution that was followed by Koivulehto. Koivulehto followed Sammallahti (1988) in supposing that the word has reflexes only in the central branches Mordvin, Mari and Permic and can thus be derived from the Finno-Permic proto-language.

Sammallahti rightfully rejected many of the cognates that had been presented in earlier literature. The UEW gives a cognate set stretching from Mordvin to Samoyed, but some the cognates are different from the ones used by Aikio. The nasal of the Proto-Uralic form given as one alternative in the UEW is reconstructed because of the Samoyed forms; this is also noted by Abondolo (1996), who writes that the Samoyed forms are not real cognates of the other words that are listed by UEW. Aikio accepts only the Selkup forms that are presented by the UEW, and he considers Nenets *tenc* ‘Art, Sorte; Stamm, Sippe’, Enets Ch *tiđ’o*, B *tiso* and K *tiz* ‘Geschlecht’ and Nganasan *tansa* id. as unrelated. Khanty DN *čáčǝ* ‘hiesig, einheimisch’, O *sāsi* ‘auf seiner eigenen Scholle wohnend, hiesig, an einem Ort wohnend; Pflegling (Haustier; nicht Hund); Wohn-, Aufenthaltsort (auch z. B. von dem Wohnplatz im Sommer)’ and Mansi KU *sasə*, P *šošǝγ*, *sossa* ‘an einem Orte wohnend; heimisch’ cannot be derived from the word with a word-initial affricate that has been reconstructed by Aikio (2015b).

The Finnic and Saami cognates have been earlier considered as a separate Finno-Saamic cognate set (see SSA s.v. *sataa* with references), but because these forms can be phonetically regular cognates to the other forms mentioned by Aikio, there is reason to support the inclusion of these words in this Proto-Uralic cognate set. Also semantically the idea is plausible. The semantic reconstruction of the proto-form is rather similar in all the sources mentioned: Aikio and Sammallahti reconstruct ‘to grow’, the UEW ‘to be born, to grow’.

Finnic *s* and Saami *ć* reflect a regular dissimilation PU \*č : č > Pre-Fi/Pre-Sa \*ć : č > Fi. *s* : *t*, so all of the Uralic forms can be derived from a form with word-initial \*č. The words are regular cognates. The Mordvin and Mari words could also reflect earlier \*š, but there is no obstacle in deriving them from the reconstruction \*čāča- that Aikio suggests.

To return to the topic of the Iranian etymology, semantically there is no problem in the etymology. The small semantic differences between the Iranian and Uralic side could be easily explained away (the Indo-Iranian word and its Indo-European cognates denote primarily ‘being born’), but the phonological arguments and also the wide distribution make the etymology less likely.

The substitution *\*dz* > *\*č* in the anlaut position is phonetically possible, but there are no other convincing examples of this, as *\*š* is the usual substitution for *\*ts* in Proto-Iranian loanwords in the word-initial position. The reasons for this remain, for the time being, unclear. It is of course possible that the affricate *\*č* is the substitute for the voiced affricate, whereas the unvoiced *\*ts* was substituted by *\*š*. In other examples where Iranian words with *\*dz* were borrowed, the substitution seems to be *\*s-* (as noted by Kallio in an unpublished work). One possibility for this problem could be that the forms with *\*s-* are borrowed from post-Proto-Iranian language(s) where *\*dz* had already developed to *\*z* (as happened in Avestan and the majority of the Iranian languages). Uralic *\*s* would be the most expected (“nearest”) substitution of Iranian *\*z-*. There are old ideas of *\*z-* forms being borrowed into the Ugric and Central Uralic languages, and they show precisely this substitution.

The distribution of this word is suspiciously wide in contrast to the idea of its Iranian origin. The Proto-Iranian loanwords are typically limited to a couple of (usually neighbouring) branches of Uralic only (see the entries *\*maksa-*, *\*paksu*, *\*počaw*, *\*serä*), which, together with the obvious fact that they have to be later than the Indo-Iranian loans, points to relatively late, post-PU contacts, and it would be unlikely for a Proto-Iranian borrowing to be found in (almost) all the branches of the family. Chronology makes contact between PI and PU problematic. Early parallel borrowings could perhaps come into question, but because of the difficulties with the sound substitutions, this does not look likely.

In addition to these arguments against the etymology, a further disturbing feature in this word is that the Iranian original suggested by Koivulehto represents the reduplication syllable and initial consonantism of the root *\*dzan(H)-*. It is suspicious that Uralic speakers borrowed only the reduplicated part of the verb – what was the exact source from which the Uralic verb was borrowed, then? Although it is usually very difficult to determine from which form verbs were borrowed, it is risky to suggest that the perfect would be the form that the speakers of Uralic would have found the one that they need to borrow. The impossibility of finding a suitable, precise origin for the borrowing is a further argument against this etymology. It is interesting to note that there are no (convincing) parallel examples of the reduplicated syllable of verbs being borrowed (the noun *\*kekrä* is, of course, a different matter because its supposed Indo-Iranian donor form is a noun that is derived through reduplication); the situation can be compared to the Baltic and Germanic loans in Finnic, where the verbal prefixes usually do not get borrowed, as noted by Kallio (personal communication). Here there is even more reason for caution, as the reduplicated syllable is part of the verbal stem of the perfect forms, and it is dubious to assume that exactly this part of the verbal stem would have been borrowed.

Not an Indo-Iranian etymology

? **Fi. *sarastaa*; SaN *čuoŕri* ‘brownish’, found in all Saami languages except SaKld < PSa \**čuoŕēs* (Álgu, s.v. *čuorgat*)**

← **PII \**čāra-* ? > Old Indic *śāra-* ‘bunt, scheckig, gefleckt’ (EWAia III: 489 s.v. *śāra-*)**

(Sammallahti 1999: 78; 2001: 399; Holopainen 2018b: 157–158)

The Indo-Iranian etymology presented by Sammallahti is both phonologically and semantically credible, but the problems with the unclear etymology of the Indo-Iranian word make the loan etymology suspicious. Sammallahti only assumed an Indo-Iranian etymology for the Saami word, but Aikio (personal communication; see also Holopainen 2018b: 158) has considered the Finnic word to be its cognate. The Permic languages seem to have acquired a separate loanword from the same Indo-Iranian source; the correspondences between the Saami word and the Permic word (Komi *žor*, Udmurt *žar*) are not regular (Sammallahti 1988: 532).

According to Sammallahti (2001), the word’s distribution is limited in the Saami languages, as the noun \**čuoŕēs* appears only in South, Lule, Pite, North and Inari Saami, which could mean that this word was acquired after the split of Proto-Saami. Aikio (2012b: 88, footnote 12), however, refutes this claim, because the derivative (SaN) *čuorgat* ‘grey of hair’ is found in eight Saami languages.

According to EWAia, the background of the Indo-Aryan word is unclear; there is no convincing Indo-European etymology, and the word does not necessarily even go back to Proto-Indo-Iranian (Avestan forms with *-sā-i* attested in personal names have been regarded as possible cognates of the initial element *śā-*, but EWAia considers this equation questionable). This naturally weakens the etymology of the Saami word as well.

As noted by Holopainen (2018b), Aikio (personal communication) connects the Saami word with Finnish *sarastaa* ‘to dawn’; the Skolt Saami cognate of the verb, *čuo’rrjed* ‘valjeta, seestyä (taivas) = to become clear (of the sky)’, suits this semantically, and the connection is also phonologically regular. The meaning ‘brownish’ of the adjective then has to be secondary, probably derived from the colour of dawn. In this case the Indo-Iranian etymology is semantically quite problematic. According to Aikio, a Baltic or Balto-Slavic origin (from a different PIE root) for this word would also be possible: Lithuanian *žarà* ‘Rot (am Himmel), Lichtschein, Dämmerung’ < Proto-Balto-Slavic \**žar-* (\**žor-* according to Derksen 2015: 513, s.v. *žarija*).

? Indo-Iranian etymology

**SaN *čáris* ‘coarse (of wool)’, also in SaS and SaL (< PSa \**čārēs*)**

< **Pre-Sa ? \**šārās* ~ \**šaras***

← **PII ? \**čara* >- Old Indic (compound) *kapúchala-* ‘Haar am Hinterhaupt’ (EWAia I: 302, s.v. *kapúchala-*)**

(Sammallahti 1999: 78–79; 2001: 399; Holopainen 2018b: 154–155)

The Indo-Iranian etymology presented by Sammallahti (2001) is not very convincing. According to EWAia, the word *kapucchala-* does not have a proper Indo-European etymology, which makes Sammallahti's assumption dubious. EWAia (I: 302) states that the explanation of *\*kaput-śala* as a compound word where the latter element means 'hair' is "ganz unwahrscheinlich". On the other hand, there is no other known etymology for this Saami word. Aikio (personal communication) points out that the South (*tjaerries*) and Lule Saami (*tjárres*) words require the reconstruction of a Proto-Saami geminate *\*-rr-*, which makes the etymology even weaker, as this sound is usually not found in inherited, pre-Saami vocabulary.

If the etymology was correct, it would be one of the few examples where Uralic *\*l* corresponds to Indo-Aryan *\*l*. In most loans the Indo-Iranian *\*l > \*r* change is reflected.

Not an Indo-Iranian etymology

## SaN *čearru* ‘fell-top’ (cognates in L, Sk, Kld, T < PSa \**čearō*)

### < ? Pre-Sa \**čero*

← ? PII \**čr̥Has-*, > OI *śíras-*, Av *sarah-* ‘head’ (EWAia II: 638–639; Nussbaum 1986: 139–157; Gotō 2013: 34; Adams 1991: 10)

(Sammallahti 1999: 80; 2001: 399; Koivulehto 2008a: 318, footnote 7 = 2016: 405, footnote 7; Holopainen 2018b: 156–157; Holopainen, Junttila & Kuokkala 2017: 117)

Sammallahti’s relatively recent etymology has not received much comment in the subsequent etymological literature. It is mentioned by Koivulehto and it has been criticized by Holopainen (2018b) and Holopainen, Junttila & Kuokkala (2017). Although the etymology is not impossible, it involves certain problems.

The word’s occurrence in only Saami already raises questions about the plausibility of the etymology, although this is not an obstacle *per se*. Semantically the top of a mountain or fell could be borrowed from this IE root, from which various words with the meaning ‘horn’, ‘head’ etc. are derived. The Proto-Uralic \**šarwi* is a borrowing from Indo-Iranian source ultimately reflecting the same root as the one that is suggested here (see the entry \**šarwi*). The borrowing into Pre-Saami should have happened roughly at the same time as the borrowing of the more widespread ‘horn’ word, as both show palatal \**č*/\**ś* in inlaut position. The Saami word cannot be a later loanword from Iranian, as this would have given Saami *s-* (← Iranian \**s*) or \**č* (Proto-Iranian \**ts*).

Although Sammallahti gives both PIE and PII forms as possible originals, the Indo-Iranian source is less likely: the second syllable labial vowel in Saami could perhaps reflect the Indo-European \**-ew* sequence, and the Pre-Saami \**e* points rather to IE \**e* than the PII sequence \**r̥H* (which was later vocalized to \**ir* in Indo-Aryan and \**ar* in Iranian). The idea that Saami \**er* could reflect the \**r̥H* sequence cannot be ruled out for certainty, but there are no exact parallels to this substitution. The Indo-Iranian word is an *s*-stem \**čr̥Has-* (> Ved *śíras-* N) that continues a PIE \**kr̥h<sub>2</sub>es-/os*), extended with \**-an/n* suffix (Gotō 2013: 34).

As the root for ‘head’ is very widespread in Indo-European, it would be technically possible to find several good candidates for the origin of this loanword. The issue is also made more difficult by the argumentation by Nussbaum (1986) that some of the words continue the simple *aniṭ* root \**k’er-*, whereas some forms contain the suffixal \**h<sub>2</sub>*, which was often thought to be part of the root by earlier researchers. It is thus difficult to determine the exact source for the Saami borrowing – in a way this means that there are multiple possibilities, but in reality this rather weakens the etymology, as it is impossible to prove the Saami form to be borrowed from some particular word in a given Indo-European language.

The Pre-Saami stem-type \**e-o* is not usually found in inherited words, which speaks against an early Indo-European origin of the Saami word (the suggested early borrowings with non-initial syllable labial vowels are suspicious etymologies in general, as argued by Holopainen, Junttila & Kuokkala 2017). The etymology thus remains unclear. It is a rather boring but inevitable conclusion that the etymology of this Saami word remains uncertain, together with a large amount of other Saami words. As Aikio (2004) has listed many Saami

words of obscure “Palaeo-European” origin denoting arctic geography, it may well be that this word belongs to the same layer of substrate borrowings.

? Indo-Iranian etymology

### SaN *čearda*, found also in SaL and SaS

< PSa \**čearte* ‘tribe’

← Pre-II \**sćerdʰo-* > PII \**śardʰa-* > Old Indic *śárdha-* ‘Stärke, Schar’, Av *sarədanā-*, root \**čardʰ-* ‘trotzen, seine Stärke zeigen, 2. furzen’ (EWAia II: 619–620, *ŚARDH*; Werba 1997: s.v. *srdh*; Lubotsky 2001a: 50–51)

(Sammallahti 1999: 78; 2001: 399; Aikio 2012b: 88, footnote 12; Koivulehto 2001b: 49 = 2016: 290; Holopainen 2018b: 155–156)

This is another of Sammallahti’s Indo-Iranian etymologies for Saami words. The etymology has also been discussed by Koivulehto and mentioned also by Aikio (2012b) in his criticism about Sammallahti’s distribution criterion in the dating of the loanwords in Saami.

The Indo-European origin of this Saami word is promising, but there is a number of problems in the way this etymology has been presented in the literature. Sammallahti uses outdated Proto-Indo-European and Indo-Iranian reconstructions (PII \**śardʰa*, PIE \**k̑erdʰo-*, \**k̑erdʰā*) for this word, largely referring to data found in Pokorny’s IEW. According to Lubotsky (2001a: 50–51), the Proto-Indo-European shape of the word is \**skerdʰo-* (with *s*-mobile), and in Proto-Indo-Iranian this becomes \**sćardʰa-*. If Lubotsky is correct in assuming that the word-initial consonantism reflects PIE \**sk* that became PII \**śc* in palatal environments, it seems more probable that this is a (Pre-)Indo-Iranian loanword than a Proto-Indo-European one; the Pre-Indo-Iranian form \**sćerdʰo-* would serve as a good origin for the Saami word, but \**skerdʰo* probably would not, although one could of course argue that already at phonetic level the PIE \**sk* could have been palatalized and thus substituted by Pre-Saami \**ć*, but this is speculation that is impossible to prove.

This word is also more widely represented in Saami languages than Sammallahti thought (Aikio 2012b: 88, footnote 12), as the word has also a cognate in South Saami (Sammallahti mentioned only the Lule and North Saami forms). As noted by Aikio, this contradicts Sammallahti’s views that due to the distribution of early loanwords within Saami, the Indo-Iranian and other early Indo-European loanwords in the Saami languages were acquired at the time when Saami was already an independent branch and was spoken near the area where the Saami languages are currently distributed. This is not a very well-founded idea. Aikio (2012b) situates the Saami homeland in southern Finland (see also Kallio 2009, Saarikivi 2011), but he has argued that even the Baltic loanwords in Saami are acquired via Finnic which was situated between the Saami and Baltic language areas. The situation with the Indo-Iranian loanwords is probably similar.

Álgu (s. v. *čearda*) mentions also that cognates for this word have been postulated from many Uralic languages in earlier sources (such as FUV: 148): Mari *šār* ‘Gemütsart, Charakter, Art und Weise, Gewohnheit, Beschaffenheit, Bau, Konstruktion’, Komi *šer* ‘Gewohnheit, Sitte, Mode, Weise’, Khanty *J šēr* ‘Art, Beschaffenheit’, Mansi (South) *sēr* ‘Faj; Art; mód, szer; Art u. Weise’, Hungarian *szer, szēr* ‘Mittel Gerät; Zeug; dial.mód; Art,

Weise', Nenets *śer?* 'Sache, Angelegenheit; schlechte Sache; auf Weise, wie' (+ cognates in Enets and Nganasan). Although the cognates are not commented on in Álgú, it is obvious that they are irregular, and it is understandable that Sammallahti (2001) does not mention them. The cognate set of the other words that Álgú lists is also mentioned in the UEW (s.v. *\*śerʒ*), but the Saami word is not included because of the *\*-rt-* cluster that is not reflected in the other languages. For the possibility of deriving these words separately from Indo-Iranian, see the entry *\*śerV ~ \*śertV*.

However, Aikio (personal communication) notes that because the word has the vowel combination *\*ea–e*, it cannot be older than Proto-Saami. This makes the Indo-Iranian etymology of *\*čearṭe* more suspicious, as Pre-Indo-Iranian and Proto-Saami could hardly be considered simultaneous, since Proto-Indo-Iranian and Proto-Uralic both probably split up around 2000 BC (Kallio 2006: 12, 17–19). Interestingly, relatively many other assumed Indo-Iranian loanwords confined to Saami show similar problems. Already both Sammallahti and Koivulehto (2001b: 248) noticed the discrepancy with the Saami vowels, but both assumed that the Saami stem vowel has to be secondary. If this vowel combination could be explained as a secondary development through established rules of word formation, the word could go back to Pre-Saami, but this is not the case, and one cannot assume ad hoc secondary developments to save loan etymologies. According to Sammallahti this vowel combination could have developed in the environment of palatal consonants: he mentions the (Pre-)Germanic etymology of Saami *geahčča-* 'to look' (from Germanic *\*gātja-*) as a parallel.

? Indo-Iranian etymology

**Mansi East *śōk*, North *śāk* 'hammer' (< PMs *\*ćāk*); Khanty North *śak*, *śax* 'Hammer', South *t'ak* (< PKh *\*ćāk*)**

< POUG *\*ćākV* 'Keule' (Zhivlov 2006: 142)

← PI (?) *\*ćakuš-*, *\*ćaku-* > YAv *ćakuš-*, *ćaku-* 'Wurfhammer, Wurfaxt' (AiWb: 575, s.v. *ćakuš-*)

(Korenchy 1972: 78–79, No. 37; Lushnikova 1990: 195; Sipos 2003: 256–257)

The main problem with this etymology is the palatal affricate in Ob-Ugric. It is very difficult to understand why the Iranian affricate *\*č-* was substituted by *\*ć* in Ob-Ugric. Korenchy argues that Ob-Ugric speakers could have heard Iranian *\*č* as "weicher empfunden" than Uralic/Ob-Ugric *\*č* and that would be the reason for this substitution, but the explanation seems ad hoc. Katz (1985; 2003) has argued that this is possibly a loan rather from Uralic to Indo-Iranian, because the Iranian words do not have a convincing Indo-European etymology. Katz's idea is supported by Kümmel (2018a).

However, Bielmeier (2001: 5–7) argues that the Iranian word can rather be borrowed from Turkic: in Turkic there are forms like Chagatai *čakan* 'Streitaxt' and Kyrgyz *čakan* id., which can be derived from the verb *čaq/čak-* 'schlagen, prägen'. As the Turkic nouns have a transparent etymology, this explanation is promising. Perhaps both the Iranian and the Ob-Ugric words can be considered as Turkic borrowings, although this does not solve the problem with the Ob-Ugric affricate. As a side note it has to be said that there are not many

generally accepted Turkic borrowings in Old Iranian, and it is unclear where very early contacts with these two families could have taken place, but a technical term like ‘axe, hammer’ can be easily borrowed even in less intensive contacts. More specific details of the Turkic etymology of Avestan *čakuš-* have to be left to Turkologists to discuss.

Sipos lists this word among etymologies that might belong to the inherited word stock of Proto-Ob-Ugric. This is a hasty conclusion: although the UEW lists the Ob-Ugric/Khanty words as uncertain reflexes of a PU form *\*čukkV*, there is no reason to support this ad hoc explanation, because the vocalism of the forms that the UEW lists shows chaotic correspondences and the vocalism of the Ob-Ugric words cannot be derived from earlier *\*u*. If the Khanty and Mansi words reflect an inherited word that goes beyond Ob-Ugric, the word was probably an *\*ä-i*-stem, as this vowel combination regularly yields *\*ā* in both Proto-Khanty and Proto-Mansi.

? Indo-Iranian etymology

### **Ko *dom-* ‘to bind’, Ud *dum-* ‘to bind’**

< PP *\*d̥mi̯* (Csúcs 2005: 330)

← PI *\*dáHman-*, > OI *dāman-* ‘Band, Seil, Fessel’, NP *dām*; root *\*daH* < PIE *\*deh<sub>3</sub>-* (EWAia I: 716–717, s.v. *DĀ<sup>3</sup>*)

(Joki 1973: 257; Rédei 1986c: 67–68; Lushnikova 1990: 197; Csúcs 2005: 330)

The word is probably a relatively early borrowing into Pre-Permic because of the vowel developments. The vowel correspondence *o ~ u* within Permic is similar as in the case of many other loans, such as *ozir ~ uzir* and *śo ~ śu*, although this loan is probably later because of the word-initial *d* that does not regularly reflect any PU phoneme. Rédei argues here an umlaut-like development where *\*-a* in the second syllable has influenced the initial syllable vowel which should be PP *\*o* but has become *o* instead. It is probable that the combination *o ~ u*, whatever be its Pre-Permic predecessor, was typical of borrowed Indo-Iranian *\*a-a*-stems, and Rédei’s idea can be rejected.

Semantically the etymology is convincing. Joki and Rédei cite only nominal forms of Indo-Iranian, but the verbal root *\*daH* exists. The Permic verb with *\*-m* looks more like a borrowing from the noun that is formed with the suffix *\*-man-*, but this is not an obstacle to the etymology.

An Indo-Iranian etymology

**Ko *d̥iṛ* ‘(long) time’; Ud *d̥iṛ* id.**

< **PP \**d̥iṛ*** (Csúcs 2005: 330)

< **Pre-P \**dura* /*tura* ?**

← **I \**dūra*- ‘far’** < **PII \**duHra-***, > **Av *dūra*- ‘far’, OI *dūŕá*- id.** (EWAia II: 739, s.v. *dūŕá*-)

(Joki 1973: 257; Rédei 1986c: 67; Csúcs 2005: 330)

In spite of its occurrence in only one branch of Uralic, the word is probably an early borrowing because of the vocalism. The *d-* here has to be secondary. It is possible that the word was borrowed into Pre-Permian at a time when *d-* was already possible, or *d-* has developed secondarily from earlier \**t-*.

Rédei also mentions that Iranian \**darga-* ‘long-lasting’ (< PII \**dr̥Hga-*) could be an alternative etymology, but he considers the borrowing from \**dūra-* more convincing semantically. Note that Rédei’s Proto-Indo-Iranian reconstruction \**darg̃a-* (\**dargya-*) is impossible. Both options are also mentioned by Csúcs (2005), who likewise follows Rédei’s reconstructions.

An Indo-Iranian etymology

**SaN *earti* ‘side-piece of a carcass, the side of a person or quadruped’ (cognates in all Saami languages); Mo E *ird’es*, *ird’is*, M *ird’es* ‘ribs’; Ma *erda* ‘thigh’; Ud *urd-li* ‘ribs’, *urdes* ‘side’, Ko *ord-li* ‘ribs’ (< **PP \**ɹrd***)**

< **PU \**ertä*** (UEW s.v. *ertä*; Sammallahti 1988: 552)

← **Pre-II \*(H)*érdʰo-*** > **OI *árdha-* M ‘Seite, Teil, Hälfte, Ord’, *ardhá-* M, N ‘Hälfte, Partei; halb, die Hälfte ausmachend’; YAv *arəða-* M ‘Seite, Hälfte’, *arəḏah-* N ‘Seite’, Oss Digor *ærdæg*** (EWAia I: 119, s.v. *árdha-*; NIL: 106, footnote 2, s.v. *dʰeh₁-*)

OR ← **PII \**rdha-*** > **OI adverb *rdhak* ‘besonders, abgesondert, abseits, beiseite’** (EWAia I: 258–9 s.v. *rdhak*)

(Munkácsi 1903: 378; Jacobsohn 1922: 206; Joki 1973: 254–255; Katz 1985: 114–115; 2003: 98; Rédei 1986c: 50; Helimski 2000: 192; Koivulehto 1988; 1991: 105–06; 2001b: 291; Sammallahti 1998: 126)

The Indo-Iranian origin of the Uralic word has not been doubted since Munkácsi (1903), although Koivulehto (1991) has suggested that the word could be derived equally well from Proto-Indo-European. The word is probably borrowed from a Pre-Indo-Iranian word that still reflected the PIE vowel \**e*, and thus belongs to the earliest layers of Indo-Iranian loanwords in Uralic. This word belongs to the same stem-type as some other well-known borrowings such as \**mertä*, and because the zero-grade of the root is attested in the adverb OI *rdhak*, the word could also be borrowed from a form with PII \**r̥* with the sequence \**er*

as its substitute. However, because the Uralic noun is semantically closer to the Indo-Iranian noun, and because it is unlikely that a Uralic noun would be borrowed from an Indo-Iranian adverb, it is much more probable that the Uralic word is indeed borrowed from Pre-Indo-Iranian *\*(H)érd<sup>h</sup>o-*. Joki also mentions an older opinion by Jacobsohn that the word is borrowed from Proto-Iranian. This is impossible because of the *\*e*, and there are absolutely no other signs of iranisms in the word.

Koivulehto (1991) briefly discusses this word as one of the examples of early Indo-European borrowings which show no reflex of a laryngeal. Koivulehto considers the word PIE (Pre-II) and does not comment on the donor language more precisely, simply noting that because of the *\*e* the borrowing has to be “recht früh”. In his 1991 work Koivulehto supported the idea that the words without a reflex of the laryngeals can be borrowed from an eastern IE dialect which had lost the laryngeals earlier or from words which had never contained laryngeals (this is relevant especially for the words with vocalic anlaut – as is well-known, Indo-Europeanists disagree whether one should follow the idea of an obligatory CVC(C)- root structure, forcing one to reconstruct a laryngeal for the anlaut of all the roots that appear vowel-initial, cf. Kapović 2017: 43–44, 52). Here Koivulehto mentions that the word is derived from a root *\*er(H)-*; however, Mayrhofer reconstructs a laryngeal-initial root, which is not commented on by Koivulehto (1991) although he does refer to Mayrhofer’s EWAia.

Connected to this problem is obviously the Indo-European background of the Indo-Iranian word. Although Koivulehto is convinced of its Indo-European etymology, RIVELEX is more doubtful and considers the origin of the Indo-Iranian word to be uncertain. EWAia considers the derivations *\*árdha-* (< *\*Hérd<sup>h</sup>o-*) and OI *r<sup>h</sup>dhak* connected to the root *\*Hard<sup>h</sup>-* which clearly reflects a Proto-Indo-European root, but the noun *\*árdha-* or the adverb *r<sup>h</sup>dhak* do not *per se* have formal cognates outside of Indo-Iranian. This means that the Uralic word has to be borrowed from an early form of Indo-Iranian, and the word indeed belongs to the layer of Pre-Indo-Iranian borrowings as has often been assumed.

According to NIL (106, footnote 20, s.v. *d<sup>h</sup>eh<sub>1</sub>-*), a close formal parallel is possibly found in Latin *orbis* ‘Radkranz, Felge’ that could reflect PIE *\*h<sub>2</sub>(o)r-d<sup>h</sup>h<sub>1</sub>i-*. The Latin word obviously reflects the PIE *\*o*-grade, and formally the Indo-Iranian word could, of course, go back to earlier *\*Hord<sup>h</sup>o-*. Here the Uralic word shows that the Indo-Iranian word has to reflect old *\*e*-grade.

Mari *örδəž* (< PMa *\*ürδəž*), which was considered a reflex of this word in many earlier works, is borrowed from Permic (Aikio 2014c). Katz (1985, 2003) also noted that the Mari word cannot be regularly derived from the same Uralic form as the other words mentioned here, and he considered it as a possible parallel borrowing from Indo-Iranian,<sup>13</sup> but this is more unlikely than a later borrowing from Permic. Berezki (2013 s.v. *erde*, *ördəž*) connects the two forms but does not comment on their relationship. The obvious Indo-Iranian etymology of the Uralic word-family is not mentioned by Berezki.

The other Uralic words are regular cognates, so the Indo-Iranian word has to be borrowed into Proto-Uralic already. However, the Saami word points to an earlier *\*-rtt-*; a Pre-Saami cluster *\*-rtt-* corresponding to Indo-European *\*-rt-* is attested in several loans,

<sup>13</sup> Katz also mentions the possibility that the Mari word is a “neugebildete hintere Variante” of the Uralic word, but this idea is based on his peculiar vowel reconstruction and has to be rejected.

but it is unclear whether this reflects an outright substitution or whether the development into a geminate is a later, independent development in Saami. It is more likely that Saami has regularly developed \*-*rtt*- from Uralic \*-*rt*- and this is argued by Kallio (2012b: 231, footnote 10), who also notes that Permic shows secondary \*-*rtt*- in some loanwords (see also the entry \**mertä*). According to Rédei (1986c) the Saami cluster is a result of “sporadic gemination”, but the development seems rather to be regular. It seems that both Pre-Saami and Pre-Permic independently had secondary gemination of \*-*rt*- clusters (and perhaps \*-*rk*- clusters as well, cf. \**wärkkä* ‘kidney’).

Helimski (1997b; 2000: 192) has noted that some Uralic forms seem to point to a PU reconstruction \**ertäs(V)*. The -*s* here theoretically could be the substitute for the Indo-Iranian masculine ending -*s*; such a substitution is found in some loans (cf. the entry *taivas*). The Mordvin word could be regularly derived from such a form, but it is dubious whether the Permic -*s* could reflect the same -*s*. Rédei had assumed that the words with vocalic and consonantal auslaut reflect chronologically different borrowings from Indo-Iranian, but it is not plausible to assume that already at the time when the Indo-Iranian language was still phonologically almost identical to Proto-Indo-European, the Uralic languages could have already exhibited such dialectal diversity. The coexistence of -*s* forms and forms without *s* in Permic also speaks against the very early origin of *s*. All the forms can be derived from \**ertä* if we assume that the -*s* in some words is secondary.

An Indo-Iranian etymology

**Fi *ehiti*- ‘to make it, to arrive, reach on time’ (cognates in all Finnic languages); SaN *astu* ‘leisure’ (+ cognates in L, Kld, T) < PSa \**eštō*; Mari *ašte-*, *äšte-* ‘to do, make, work’ < PMa \**ištā-*; ? Komi *ešti-*, *ješti-* ‘(mit etw.) fertig werden fertig od. vollendet werden (Arbeit);(zu etw.) Zeit haben; (etw. tun) können, rechtzeitig hinkommen; reifen, reif werden’ (< PP \**ešti-* Csúcs 2005: 331)**

< PU \**eštä-* (UEW s.v. \**ešte-*)

← PII \*(*H*)*išti-*, cf. Av *išti-* ‘possession’, < PIE root \**Heyk-* ‘to have in one’s power’ (EWAia I: 270–271 s.v., *Eṣ*’; Werba 1997: 162 No. 12; Cheung 2007: 158, s.v. \**Haiš*; LIV2: 260, s.v. *h<sub>2</sub>eṣ-*)

(Hyllested 2014: 14)

See also \**icä*, \**isäntä*

Hyllested was the first to suggest an Indo-Iranian etymology for this “Finno-Permic” word in his analysis of Koivulehto’s laryngeal etymologies. Koivulehto (1991: 75–78) derived this “Finno-Permic” verb from Proto-Indo-European \**yéh<sub>2</sub>-g<sup>w</sup>-ie/o-*, which is reflected in Baltic \**yēgti* (> Lith *jēgti* ‘vermögen’, *pa-jēga* ‘ability’), but Hyllested quite convincingly refutes this etymology in his survey of the laryngeal loans that show Uralic \**š* as the substitute for the PIE laryngeals. His reasons for rejecting Koivulehto’s etymology are the problems with the scant attestation of the IE formation (only in Baltic) and the problems with the initial \**j-* (which cannot be reconstructed on the basis of the FP internal evidence). Instead,

Hyllested supports a loan from some *satem* language, citing the Indo-Iranian origin as the most plausible.

Hyllested's Indo-Iranian explanation is a promising alternative to Koivulehto's etymology (which, indeed, has to be rejected), but this Indo-Iranian origin for the Uralic word is not completely suitable either because of the vocalism. The attested Indo-Iranian forms reflect zero-grade vocalism of the root *\*Heyk*´-, and it is unlikely that the PII long *\*ī* (or earlier *\*Hi*) could result in Uralic *e*; here Uralic *\*i* would be the phonetically nearest and expected substitution. One would certainly expect a retained diphthong here if the Uralic word were borrowed from an Indo-Iranian word with full-grade *\*ey*: it would result in PU *\*\*ej*, so that would not be a better candidate either. So while the etymology is semantically good, it cannot be accepted because of the phonological problems.

Koivulehto has derived the words *isä* and *isäntä* from the same Indo-Iranian root (see the respective entries). These etymologies could serve as counterevidence to the substitutions that Hyllested assumes, but these two etymologies are also problematic due to other reasons.

The relationship between Komi *eštĭ-*, *ještĭ-* and the Finnic word is irregular, and the Komi word can be borrowed from Finnic according to Saarikivi (2018: 308–309).

Not an Indo-Iranian etymology

**Mansi South *ätər* East *ātər*, North *atər*, West *etər* ‘bright, clear (of weather)’ (< PMS *\*ätər* (~ -ə)); Khanty East *etər*, South *etər*, North *etər* id. (< PKh *\*ätər*)**

< **POUg (?) *\*eträ*** (Zhivlov 2006: 108; Honti 1982: 73) < ? **Pre-Ob-Ugric/PU *\*iträ***

← ? **PI *\*wīdra*/*\*idra-* ‘clear’ > Oss *ird* ‘bright, clean’, OI (Vedic) Loc. *vīdhré* ‘bei heiterem Himmel’ (< PII *\*wi-idhra-*)** (Abaev 1958–1989 I: 547; KEWA I: 97; Cheung 2007: 157; 2002: 193)

(Morgenstierne 1938: 224; Paasonen 1926: 20; Joki 1973: 260, No. 34; Katz 1985: 85–86; 2003: 78; Holopainen 2019)

Although this is an old etymology (first suggested by Paasonen and supported by Morgenstierne 1938: 224, s.v. *lāróyo*),<sup>14</sup> the etymology is not mentioned by either Korenchy (1972) or DEWOS at all. Joki accepts the etymology, and so does Katz. I agree with them: the etymology is acceptable both on semantic and phonological grounds. In Zhivlov's modern reconstruction of Proto-Ob-Ugric, the Khanty and Mansi words clearly reflect a word with a pre-Ob-Ugric *\*i* vowel (cf. PU *\*miñä* ‘miniä’ > POUg *\*meñä*), so the Iranian word is phonetically completely suitable. Also semantically the Khanty and Mansi words are very near to the Ossetic word and its attested cognates in the Indo-Iranian languages. However, the Ob-Ugric word also has a tentative Uralic etymology which I will comment on below after reviewing the Iranian etymology.

---

<sup>14</sup> Morgenstierne refers also to a personal communication with Björn Collinder, who had derived both the Ob-Ugric words and Saami *Nfier'to* ‘fine weather’, *Tvier'ta* from PU *\*vetra*. As Joki remarks, this explanation cannot be correct, and the Saami word is a different borrowing from Germanic *\*wedra*.

Katz (2003: 78) mentions that the Ob-Ugric word supports the idea that the Ossetic word *ird* is derived from Proto-Iranian *\*idra-* and not from prefixed *\*wi-idra-*. Katz's assumption is based on the notion that if the Ob-Ugric words were borrowed from *\*widra-*, *\*w-* should be visible in Khanty (there is no reason that the word would be borrowed without a substitution for the word-initial *\*w-*, and this *\*w-* would not have been lost in Khanty, only in Mansi). Katz is only correct if we assume that the Ob-Ugric word is borrowed precisely from a form that is ancestral to the modern Ossetic word. Many of the Iranian borrowings in Ob-Ugric are clearly derived from an Alanic-type language (Helimski 1997b), so it is a plausible assumption that this word belongs to the same layer. However, not all of the loans in Ob-Ugric are clearly Alanic, and it is possible that the unattested Iranian contact languages of the steppe were more different from the forms that are ancestral to Ossetic. It is any case obvious that the Iranian languages were spoken across a huge region, and Ossetic represents only the survival of a part of this huge Iranian community.

It seems though that in any case the form with the *\*wi-* prefix which is attested in Indo-Aryan does not have an exact cognate in Iranian. Cheung (2002) is uncertain whether the Yidgha word *lārōyo* 'clear sky' reflects a Proto-Iranian prefixed form or not. Cheung (2007) notes that Ossetic and Yidgha words are archaic formations and "exact correspondences of Sanskrit *vīdhra-* (AV) 'bright, clear (of sky)', Greek *ἰθαρός* 'bright, clean, pure'.

This word is one of the Iranian loanwords that can be assumed to have been borrowed into Proto-Ob-Ugric, as the Khanty and Mansi words are regular cognates. If it is indeed a common Proto-Ob-Ugric word, the borrowing has to be early (earlier than "Middle Iranian", which is a term that often characterizes the Iranian loanwords of Ob-Ugric in Uralicist sources; see the Introduction). In any case the borrowing has to be quite old, as the metathesis is a Pre-Ossetic development according to Cheung (2002: 39), as it is reflected in Greek glosses of Scytho-Sarmatian names, such *Πουρθαίος* which probably contains a reflex of PI *\*puθra-* 'son'. Palunčić (2013: 32–39) dates the metathesis to the first centuries CE.

Because some of the loanwords that have been considered Proto-Ob-Ugric are irregular and point to parallel borrowings, all of these Ob-Ugric words need special attention. It cannot be ruled out that the word has been borrowed independently into Khanty and Mansi, but it has to be admitted that there is no compelling reason to assume that these two words are not true cognates. The evidence of this etymology and the other Ob-Ugric words of Iranian origin will be commented on in the Results.

As mentioned above, the Ob-Ugric words also have a possible cognate set in other Uralic languages. Aikio (2002: 24) has supported the view that the Mansi word is a derivation that reflects PU *\*itā-*, which can be reconstructed on the basis of PS *\*ātā-* 'sehen; sichtbar sein' (SW 16), the Khanty verb *et-* and possibly the Finnic verb *itää* 'to sprout, germinate'. The equation of the Samoyed, Khanty and Finnic words has been suggested by Helimski (1999), who does not however mention the Mansi word. Aikio does not mention the Khanty word *etər* in this context, although it is obvious that the same problem involves both the Khanty and the Mansi words: both would be derived from the reflexes of PU *\*itā-*, although the reflex of the verb itself is not attested in Mansi. In this view, the resemblance of *etər* and *ätər* to the Iranian words would be purely coincidental. Aikio further mentions Tundra Nenets *ngødyer* 'clearance (raivio)' which he considers a parallel formation derived from the Samoyed verb. The idea that the Ob-Ugric words are

derivatives (with the Ob-Ugric denominal suffix *\*-r*) is also presented by Riese (2001: 52), although Riese does not mention any underived base word for this (no reference to PU *\*itä-* is made).

The Tundra Nenets word is probably derived from the Nenets reflex of this Uralic verb, but its resemblance to the Khanty and Mansi words might be coincidental. There is no productive derivational suffix *-r* in either of the Ob-Ugric languages, and it is further notable that in Khanty there is also a verb *etərəmt-* ‘klar werden, sich aufklären, schönes wetter werden’ which is clearly derived from the noun *etər*. Also the fact that in Mansi the reflex of the verbal root *\*itä-* is not even attested speaks against the idea that the Khanty and Mansi nouns are derived from this verb.

As is noted by Holopainen (2019: 76), it can be tentatively suggested that the Uralic verbal stem *\*itä-* ‘to become visible, to appear’ that is probably reflected by the Finnic and the Samoyed forms is borrowed from this same Indo-Iranian verbal root *\*Hid<sup>h</sup>-*, although here the semantic difference is notable. The idea of Koivulehto (1988b: 3–5; 1991: 51, footnote 23) to derive the Uralic word from a Pre-Germanic form *\*g<sup>h</sup>i-* (> Goth *keinan* ‘keimen’) is unconvincing for phonological reasons (Holopainen 2019: 74–76).

An Indo-Iranian etymology

**Mansi South *ērəw*, East *ērī*, West *jeriy*, North *ērīy* (< PM *\*īrəy* (~ -ə)) ‘Lied’; Khanty East *ärəy* ‘Heldenlied; Kirchlicher Betgesang’, South *arə*, North *ar* ‘Lied’ (< PKh *\*ārəy*)**

< **POug** *\*ēräy* (Zhivlov 2006: 154)

← ? **PI** *\*ā-garH* ‘to greet, call’ > **YAv** *āyar-* ‘to greet’, **Buddhist Sogd** *nyr’y* ‘to praise, honour’, **Oss.** *arğaw* ‘story’, **Iron** *arğawyn/arğud*, **Digor** *arğawun/arğud* (denomin.) ‘to perform a church service’ (Abaev 1958–1989 I: 65–66; Cheung 2007: 107, s.v. *\*garH*)

(Munkácsi 1901: 237, 524–526; Korenchy 1972: 52–53, No. 3; Lushnikova 1990: 180; Sipos 2002: 13)

As Korenchy notes, the initial etymology by Munkácsi conflated two Indo-Iranian forms: OI *arká-* ‘Strahl, Sonne, Feuer, Lied’ is unrelated to Ossetic *arğaw*, and Korenchy considers the latter Iranian form as the source of the Ob-Ugric words (*arká-* is from PII *\*arč-* ‘to sing’ [EWAia I: 114–115, s.v. *ARC*; Cheung 2007: 163, s.v., *Harč*]). Korenchy reconstructed the Proto-Ob-Ugric form as *\*ārγV* (~ *\*īrγV*), but Zhivlov reconstructs a form with *\*ē* instead. Also Proto-Ob-Ugric *\*wērä* ‘Fischzaun’ shows similar vocalism (see the entry). As Zhivlov (2006) does not comment on the Indo-Iranian origins of the Ob-Ugric words, it remains uncertain whether he assumes that the Iranian vowel is substituted by *\*ē* in cases like this, or whether the word was borrowed earlier into a Pre-Ob-Ugric form that developed to *\*ēräy* later. The Khanty and Mansi words manifest similar vowel correspondences as well-known Uralic words such as *\*ñälmä* ~ *\*ñelmä* ‘tongue’ and *\*sänV* ~ *\*senV* ‘mushroom’, so the cognateship between Khanty and Mansi seems to be regular here, and thus the words can at least theoretically be derived from an Ob-Ugric proto-

language. Because the vowel combination can point to earlier Uralic \**ä*, it is possible that the words were borrowed before \**ä* developed to \**ē*, and the substitution was thus Uralic \**ä* ← Iranian \**a*.

The metathesis of \*-*rK*- clusters is typical of both Ossetic and the Ob-Ugric languages, and Korenchy argues that in this case the metathesis had probably already happened on the Iranian side, although it remains unclear why this should be the case.

Semantically the etymology is convincing, although the meanings of the Iranian and Ob-Ugric words are not exactly the same. In both language families meanings related to ‘praising’ and religious services are attested. Words related to religion are easily borrowed (Saarikivi 2014: 199). See also the entries \**juma*, \**tajwas*.

An Indo-Iranian etymology

**Ko *gu-* ‘steal’, deriv. S V *guśal-*, P *guśav-* < PP \**gu-* (Csúcs 2005: 334)**

< ? Pre-Permic \**kada-*

← **PI \**gada-*, YAv *gaða-* ‘thief’, ? Alanic name Γαδας** (or from Γαδατας ‘von Urstier gegeben’) (AiWb 488; Zgusta 1955: 87–88)

(Joki 1973: 262; Gamkrelidze & Ivanov 1984: 928; Rédei 1986c: 69; 2004; Katz 2003: 230; Csúcs 2005: 334; Kümmel 2018a)

Joki states that the Komi word reflects earlier \**a*. This is correct also according to modern views of Uralic historical phonology (Sammallahti 1988: 545; many other loans, such as Ko *kundi-* ‘dig’ < \**kana-* show a similar development). Earlier (PU) \**δ* has become lost regularly.

Semantically the etymology is convincing. However, Rédei (1986c) correctly notes that Joki has glossed the meanings of the Iranian words incorrectly, as Avestan *gaða-* is a noun and means ‘thief’, not ‘to steal’. In Iranian the word is attested only as a noun, whereas the Permic word is verb. Rédei argues that due to the nomen-verbum system of Uralic, the Iranian noun could have been borrowed as a verb into Uralic.

The truth might be simpler than this. Even though there is no existing verb in Iranian, EWAia (II: 460–461, s.v. *gadh*) connects the Avestan word to the Indic root *gadh* ‘ergreifen’, attested in Vedic several times in the gerund *gádhya-* and in some other forms (see Werba No. 401, s.v. \**gandh* for information on further attestations). Werba assumes a PII root \**g<sup>h</sup>and<sup>h</sup>* (with *gadh* reflecting the zero grade of *n*) and does not mention the Iranian forms at all. If this etymology is correct, we can assume that the Permic word has been borrowed from a reflex of this verbal root.

Lushnikova (1990: 208) suggests another etymology, deriving Permic *guśal-* from Iranian \**gauz* ‘to conceal’ (> Av *gaoz-*). For the Iranian root, see Cheung (2007: 117, s.v. *Ägauz*). Usually the Permic forms *guśal-*, *guśav-* with -*ś* are considered derivatives of *gu-*, but if this etymology is accepted, they could represent a borrowing from a different Iranian root. However, the palatal sibilant in these forms is unexpected, and it is better to treat the *gu-*, *guśal-* words as belonging to a single word family. Permic *gu-* cannot be borrowed

from *\*gauz*, as the sibilant could not have disappeared in Permic. Lushnikova's etymology is probably not correct.

Katz and Kümmel accept the connection of the Permic and Iranian words but rather argue for a Uralic borrowing into Iranian. This is unlikely in the view of the distribution (as is noted by Rédei 2004b: 324) and Komi voiced *g-*, which very often appears in loans. If the word is an old borrowing, the Permic *g-* has to be secondary, but it cannot be ruled out that the word has been borrowed into Pre-Permic later, when *g-* already existed. Because of the Permic vocalism, the word cannot belong to the latest layer of Iranian loans in Permic but has to be relatively old, but there is still no reason to assume that it would go back to Proto-Uralic and could have been borrowed into Proto-Iranian.

An Indo-Iranian etymology

### Ud *gurt* 'house, village', Ko *gort* 'house' < PP *gort*

< Pre-P *\*gert(t)ä* ~ *\*kert(t)ä*

← PI *\*gr̥da-* or PII *\*gʰr̥dʰa-* > OI *gr̥há-* 'house, dwelling-place', Av *garδaa* 'dwelling-place of demonic beings'; < PIE *\*gʰr̥dʰo-*<sup>15</sup> (EWAia II: 495, s.v. *gr̥há-*)

(Munkácsi 1901: 648; Joki 1973: 262; Gamkrelidze & Ivanov 1984: 931; Katz 1985: 295–296; 2003: 236–237; Rédei 1986c; Lushnikova 1990: 206–207; Koivulehto 1999a: 229 = 2016: 231; Csúcs 2005: 334)

The Permic word for 'house' and 'village' is a typical cultural borrowing, and this etymology was noted already by Munkácsi (1901: 648) and it is noted and accepted by most sources of Iranian loanwords and Permic etymology. The word is attested only in Permic, but it has to be an early borrowing because of the vocalism: the correspondence Udmurt *\*u* ~ Komi *\*o* points to Pre-Permic *\*e* (: *ä*). However, a similar vowel correspondence is attested also in some other loans which are found only in Permic, such as *\*domV-* 'to bind'. The Indo-Iranian word has syllabic *\*r̥*, and the substitution *\*er* ← *\*r̥* is found in some etymologies, although usually the substitution is different (one possible example is *\*mertä* 'human', although it is disputed whether this is actually borrowed from the zero-grade form, see the respective entry). Furthermore, the word is not attested in later Iranian languages but only in Avestan. Koivulehto has criticized the attempts of Joki (1973) and Rédei (1986c) to derive the etymology from an otherwise unattested Proto-Iranian form *\*gerdo-* or *\*gerda-*. Postulation of such forms with *\*e*-grade only on the basis of loanwords is methodologically dubious. Also Csúcs lists "Old Iranian" *\*gr̥da* as the donor form, so he is clearly not following the idea of *\*e*-grade that is postulated by Joki (1973) and Rédei (1986c).

The voicing of initial *\*k* and other stops in Permic is a frequent development, found in a number of Permic words, so it is not an obstacle for the etymology (it is tempting to think that Permic *g* reflects the voiced Indo-Iranian *\*g* here, but it is unlikely that voiced stops

---

<sup>15</sup> Beekes (2010: 1644–1645, s.v. *χόρτος*) notes that Indo-Iranian and a number of other IE forms reflect PIE *\*gʰordʰ-*, which is however a phonetically unsuitable pre-form for Greek *χόρτος*. The relationships between the various IE words require further study.

were present at the early stage of Pre-Permic into which this word was borrowed). The other option is that the word is borrowed later from Iranian, and the Permic vowels reflect the substitution of some later Iranian vowel.

Because of the limited distribution one could argue for a relatively late (Proto-Iranian) borrowing, but on the basis of linguistic arguments only one can say little on the exact date of the borrowing. The word-internal consonantism (the retained stop) points to PII or PI, not PIA. The word does not have to be very early, as we do not know when the syllabic \*r was vocalized in the Iranian languages of the steppe, and even a later form with \*ər could have produced Pre-Permic \*kertä.

Helimski (1984: 198; 1991: 221) has criticized the Iranian etymology, arguing that Permic -rt- is difficult to explain from Iranian \*-rd-. However, as Kallio (2012b) has noted, here the -rt- probably reflects Pre-Permic \*-rtt-: this secondary gemination of \*-rt- is attested also in the Permic reflexes of \*mertä. Helimski argues that the Permic word could be borrowed from a hypothetical Turkic form \*gurt with Permic g as the substitution for g'. The borrowing would have taken place at a much later time, with the Proto-Permic vowel substituting Turkic u. This is a rather problematic explanation. The Turkic word is jurt, Chuvash šurt, and there is no real evidence of a Turkic dialect that has \*g' from earlier \*j. Helimski commits the same methodological error as Joki and Rédei with their assumption of a full-grade form \*gerda- mentioned above. Permic g naturally cannot be a substitute for Turkic j. As Helimski's objections regarding the Permic word-internal consonantism are not very strong, there is no reason to support his hypothetical Turkic etymology in favour of the traditional Indo-Iranian explanation. Also Rédei (1986c) criticizes Helimski's explanation because of its hypothetical nature.

The Iranian word is also borrowed into Tocharian B: pl. tantum OBL *ker(c)iyem* 'paleis'; from the form \*gardiya- 'dienaar', corresponding to OI *grhya-*; the TB word cannot be inherited from PIE \*g<sup>h</sup>ord<sup>h</sup>(i)yo- 'domesticus' according to Isebaert (1980: 88–89), but Tremblay (2005a: 426–427) considers both the Iranian loan etymology and an inherited Indo-European formation possible.

An Indo-Iranian etymology

**Fi *ihana* 'wonderful; (dial.) healthy, blooming', dial. and Ka *ihala* 'dear', *ihastua* 'to fall in love; (dial.) to bloom, to be refreshed' < ? PFI \**iha* 'life force, joy' (SSA I s.v. *iha*)**

**< ? Pre-Fi \**iša***

**←PI \*(H)*iš-* > Av *iš-* 'force', OI *iš-* 'Getränk; Lebenskraft' < PIE root \**h<sub>2</sub>eys-* 'in Bewegung setzen' (EWAia I: 198 s.v. *iš-*, 271–272 s.v. *Eš<sup>2</sup>*; Schindler 1972: 10–11; RIVELEX s.v. *eš<sup>1</sup>*)**

(Tunkelo 1913: 99–100; FUV: 130; Koivulehto 2001a: 367–368; Rintala 2003: 306–308; Katz 1985: 167–168, 179–180; 2003: 139, 147–148)

The group of various *iha* words in Finnic form a bunch of complicated etymologies. The idea that Finnic *iha* stems from Indo-Iranian was initially suggested by Tunkelo (1913) and

the Indo-Iranian etymology was later expanded and modified by Koivulehto (2001). The etymologies are also commented on in detail by Rintala (2003). Tunkelo's etymology is also mentioned by FUV, but Joki (1973: 116, 225) only briefly remarks that *iha* is related to *iho* 'skin' and its Indo-Iranian etymology is not discussed any further. A relationship to *iho* is not very convincing semantically (see Rintala 2003: 294–296), and it is probable that Koivulehto is on the right track, even if some of the details of the etymology remain uncertain.

In Koivulehto's view the two *\*iha* words, *\*ihhta* and *\*iso* are all borrowed separately into Finnic from different Indo-Iranian words, most of them belonging to the root *\*Hayš-* 'to wish'. According to the same scholar, *\*iha* 'wonderful' is borrowed from a different but homonymous Indo-Iranian word *\*Hiš-*.<sup>16</sup> The traditional view (SSA s.v. *iha*; FUV: 130) is that the *iha* group of words belong to the same word family in Finnic, but according to Koivulehto they are parallel borrowings. SMS does not distinguish the two word families, giving only one headword *iha* 'lust; wish'. EES accepts the Indo-Iranian etymology but likewise does not distinguish the two *iha* words.

Rintala (2003: 311) accepts Koivulehto's conclusions of two different Indo-Iranian loanwords but mentions that the two words are not semantically very far from each other. The problem with the etymologies is that in some cases it is rather difficult to differentiate parallel borrowings from words that are derivatives. It is suspicious that there should be several words found only in the Finnic branch which are borrowed from the same Indo-Iranian word group separately but roughly at the same time.

The Indo-Iranian origin of the word is credible, although polysemy and abstract meanings make it difficult to reconstruct the original meaning of the Finnic word. Semantically the Finnic word is close to the meanings that are attested in Avestan and Vedic. According to Schindler (1972: 10–11) the original meaning of the PII root noun was 'Impuls, Stärkung'. Tocharian B *aise* 'Kraft' is a parallel formation with similar meaning from the same Indo-European root. If we assume that the meanings relating to 'life force' in Vedic were present in the donor language of this loanword, the etymology is unproblematic. As is the case with the other *iha* word, the dating of the borrowing is difficult, because *\*š* (> *Fi h*) substitutes for the Indo-Iranian RUKI sibilant *\*š*. The limited distribution could hint to a later Iranian borrowing.

Rintala has also assumed that the Finnic adjective *ihana* could be a separate borrowing from an unattested Indo-Iranian verbal adjective *\*išana-* from the root *iš-*, because it would be difficult to explain the exact derivational process of the Finnic adjective from *\*iha*. As there is no trace of an Indo-Iranian adjective of the type *\*išana-*, this explanation has to be rejected as too speculative. It has to be admitted that Rintala is right in noting that the derivational process behind *ihana* is obscure.

## An Indo-Iranian etymology

<sup>16</sup> Interestingly, there is one more Indo-Iranian verbal root that bears a strong resemblance to the Finnic *iha* words. *\*ih-* 'begehren, (er)streben'; Av *iz* < PII *\*Hayšh*; (Cheung 2007: 159–160, LIV<sup>2</sup>: 222; Werba 163 No. 13). This root is not mentioned by Koivulehto 2001a. Here similar problems with substitution (*\*č* or *\*š*?) as with the etymology of *iha* 'yearning, passion' (see below) would occur if we assumed that the word is borrowed from this Indo-Iranian verb.

Est *iha* ‘yearning, passion (Vorlangen, Begierde)’, Finnish *ihastu-* ‘take a fancy to something, be overjoyed with something’

← ? PI *\*(H)istsā-* < PII *\*Hisčā-*, cf. OI *icchā-*; id; root *eš-* ‘suchen; wünschen, begehren’, present stem *icchá-* < PII *\*Hayš* or *\*HayšH*<sup>17</sup> < PIE *\*h<sub>2</sub>eyš(H)*; OI root-noun *iš-* ‘Labung, Kraft, Opfergruss’ < PII *\*Hiš* (EWAia I: 270–271 s.v. *EŠ*<sup>1</sup>; Werba 1997: 162 No. 12 s.v. *iš-*; Cheung 2007: 158; LIV<sup>2</sup>: 260; Schindler 1972: 10–11; RIVELEX s.v. *eš*<sup>2</sup>)

(Tunkelo 1913: 99–100; Koivulehto 2001a: 365–366; Rintala 2003: 306–308; SSA; EES s.v. *iha*)

According to Koivulehto this is a different word from the synonymous *iha* word that is attested in northern Finnic only (see above). As noted above, Koivulehto’s ideas are accepted by Rintala (2003). Semantically the two words are rather different, so from this point of view it is easy to understand Koivulehto’s suggestion that they represent two different borrowings. However, there are various problems with Koivulehto’s etymology for this *iha* word.

According to Koivulehto, the noun *\*iha* is borrowed from unattested PI *\*istsā* (= OI *icchā*) as Pre-Finnic *\*iča*. However, *iha* can only result from earlier *\*iša*, not *\*iča*. On the other hand, it is unlikely that the PI cluster *\*sts* in word-internal position would result in Pre-Finnic *\*š* and not *\*č*, as at least the substitution of PI *\*ts* is *\*č*, cf. *\*mača*, *\*počaw*. Pystynen (2019: 45–46) argues that the substitution was *\*ičča*; there are two possible examples in the list of Sammallahti (1988) where the geminate *-čč-* seems to be reflected as Finnic *-h-*, *\*pičča* > *piha* and *\*keččä* > *kehä*. However, the reconstruction of both Uralic forms involves various difficulties (for *\*pičča* see Aikio 2014: 39), and the development *\*-čč-* > *\*-š-* is thus highly disputed. Another problem with this idea is that it is not at all clear that the Proto-Iranian reflex of PII *\*isčā-* was *\*istsa-*; this could have developed to *\*isa-* already, and from this word we would not expect Finnic *\*ičča*. Lubotsky (2001a: 49) mentions that *\*sč* had probably become *\*s* in Proto-Iranian already.<sup>18</sup>

Because of the uncertainty of the reconstruction of the geminate affricate *\*čč*, the uncertain sound substitution and the hypothetical nature of the Iranian origin, it is better to reject this etymology and assume that the *iha* words mentioned here belong etymologically to the other family of *iha* words discussed above. The words are ultimately of Indo-Iranian origin, but probably not borrowed from an unattested *\*istsā-*.

Although the Finnic word is semantically close to the Indo-Iranian words, it seems that it is possible to find a fitting original, especially if we assume that the Finnic word was borrowed as a noun. The noun *iha* seems to be primary, and the verb *ihastua* is derived from it. Koivulehto remarks that there is a noun *iš-* in Avestan, which has earlier been glossed as ‘wish’, but there are no arguments in favour of this meaning of the Avestan word

<sup>17</sup> About the possible *set* root, see RIVELEX (s.v. *eš*<sup>2</sup>) with references.

<sup>18</sup> Lubotsky writes: “In OP, PIIr. *\*sč* yields *-s-* word-internally and *θ-* word-initially. The most plausible explanation for the double treatment in OP was proposed by Nyberg (1931) 45, who assumed that PIIr. *\*sč* developed into *\*ss* already in Proto-Iranian. In Old Persian, this *-ss-* was preserved as *-s-*, but in the anlaut position it was simplified to *s-*, which became OP *θ* together with PIr. *\*s* < PIE *\*k*<sup>1</sup>. This means that OP *θ* did not develop directly from PIIr. *\*č*, but went through the stage of PIr. *\*s*.”

(see also Schindler 1972: 10–11). Schindler argues that in PII there was only one root noun *\*Hiš-* with the basic meaning ‘Impuls, Stärkung’.

Schindler’s remarks are also cited by Katz (2003: 148, footnote 76). The idea that the noun *iha* ‘yearning’ is borrowed from an Indo-Iranian root noun of the same meaning is thus problematic. Katz (2003: 148) is of the opinion that *\*-iš-* in Vedic *gav-iṣ-* ‘Rinder begehrend’ is the source of the borrowing. This possibility is also mentioned by Koivulehto, although he considers the noun *\*istsā-* to be a more probable origin. It is methodologically dubious to take a Vedic root-compound and then argue that the root in this nominal compound could have existed as an independent noun. This makes the etymology look uncertain.

Also forms in related languages have been considered cognates of Finnic *iha*. Rintala (2003: 296–297) also mentions that Moksha Mordvin *ežālgâdâ-*, *ožālgâdâ-* ‘to rejoice’ has been connected etymologically with the Finnic word in earlier research (in SSA it is mentioned as an uncertain cognate), but the relationship is uncertain because of the phonological irregularity (the variation of the initial-syllable vowels *e* and *o* in Mordvin is unexpected). Moksha *-lgâ* is a derivational suffix. The Mordvin and Komi words were not commented on by Koivulehto.

Komi *ištî-* ‘be charmed, become delighted; wish, desire, long; be very gay, frolicsome’; *išmî-* ‘be animated; be frolicsome, licentious; be desirous, lustful (basic stem: *iš-*)’ is mentioned as a cognate by FUV (130) and this explanation is also mentioned by Rintala. Neither the Komi nor the Mordvin word is mentioned by UEW, where no idea of reconstructing a common proto-form for these Uralic words is found. Saarikivi (2006: 36) considers the Komi word to be a loan from Finnic. Komi *i* often reflects Finnic *i* in early loanwords from Finnic. This view has been supported by J. Häkkinen (2019: 36–37). Pystynen (2019: 45–46) argues that the Komi word might be a separate borrowing from Iranian.

It is obvious that neither the Komi nor the Mordvin words are cognates of the Finnic word, but both of them could be separate borrowings from Iranian. Words with affective meanings are easily borrowed, and both Mordvin and the Permic languages have several layers of Iranian loanwords, so there is nothing troubling in this explanation.

The possible relationship of Finnic *iha* to the Indo-Iranian root *\*Hiš-* remains uncertain. It may well be the case that all the *iha* forms belong into one and same Finnic word family, but its precise Indo-Iranian origin is difficult to determine.

? Indo-Iranian etymology

**Md E *ežālgâdâ-*, M *ožālgâdâ-* ‘to rejoice’ < ? *\*iša***

See *\*iša*

Fi *iho* ‘skin’ (+ Est, Vo, Izh, Ka, Lu); SaN *assi* ‘back of skins or furs; the skin as distinct from the hair on it; the outer layer of flesh’, L *jissie* ‘kjøttside av skinn; det blå i himmelen (delvis klar himmel)’, U *jissee* ‘kjøttside av skinn; det blå i himmelen (delvis klar himmel)’ < PSa \**ešē*; Mordvin E *jož*, M *jožo* ‘die äußere Haut, Oberfläche’; Mari *jüž-* in the compound *južwüt* ‘Lymphe; das Wasser, welches sich in der Schwele sammelt’ < PMa \**juž-wət*; ? Komi *ež* ‘Hautteil des Felles’ (< PP \**ež* Csúcs 2005: 331); ? Khanty East *el* ‘Körper’, North *el* ‘Haut’ (< PKh \**äl*)

< ? PU \*(*j*)*iša-* ~ ? \*(*j*)*išo-* (UEW s.v. *jiša* (*iša*); Lehtiranta 2001: 12–13; SSA s.v. *iho*; Abondolo 1996: 60; Bereczki (2013 s.v. *južwət*) EES s.v. *ihu*)

← ? PI \**idza-* or ← later Iranian \**iza-*, cf. Av *izaēna-* ‘made of leather’; < PIE ? \**h<sub>2</sub>aik-*, > Greek *αἴξ* ‘goat’, Armenian *այց* (i-stem) (Schindler 1972: 76; Bailey 1979: 484; Clackson 1994: 90, footnotes 15, 16; Kümmel 2014: 2)

(Koivulehto 2001a: 369–371 = 2016: 267–269; Holopainen, Junttila & Kuokkala 2017: 117–118)

Koivulehto’s Proto-Iranian etymology suffers from both irregularity of the Uralic cognates and the uncertain etymology of the scarcely attested Iranian (Avestan) word. The etymology has not been much commented on in the subsequent etymological literature. Bereczki (2013 s.v. *južwət*) does not mention the possible Indo-Iranian origin. Also, EES does not mention the possible Indo-Iranian origin (interestingly, EES refers to some other etymologies that Koivulehto suggested in the same 2001 article). EES mentions that the Uralic words can possibly be derived from the word-family of Finnic *iha* (< \**iša*), but this makes semantically no sense, and because *iha* does not have any cognates outside of Finnic, this idea should be rejected.

The Uralic cognates are irregular, which makes it problematic to assume that they represent an early Iranian borrowing into some proto-language stage. There are various views about the relationship of the Khanty words with the western forms: UEW has considered the Khanty word a reflex of PU \*(*w*)*ülä* ‘up’, but Abondolo (1996: 60) supports the view that the Khanty word is related to the western forms, although he also mentions UEW’s point of view. Abondolo’s reconstruction of a Proto-Uralic form \**iša* with \**i* in the first syllable is, of course, impossible, as this could not be reflected by Finnic *i*, Saami \**e* or Mordvin *o*. The Khanty word is probably unrelated to the other words discussed here, and the explanation found in the UEW is probably correct.

The relationship of the other words is less clear. The UEW notes that Mari and Mordvin point to word-initial \**j-*, but the Saami and Komi words cannot be derived from a reconstruction \**jišV* (note that the Lule and Ume Saami forms with anlaut *j-* show regular, late prothetic *j-*, but the North Saami word shows that the word originally had no \**j-*). The UEW nevertheless assumes that the Mari, Mordvin, Saami and Finnic words can be cognates, and mentions the possibility that the *j-* in Mari and Mordvin is secondary. This is an ad hoc explanation, as these branches seldom manifest a prothetic *j-*, and the etymology cannot be saved by simply declaring that something irregular is just secondary. Also Koivulehto mentions the problematic *j-* of the Mordvin word, but he does not see this as an obstacle to the Iranian etymology. Bereczki (2013 s.v. *južwət*) also notes that the *j-* in Mari

and Mordvin is irregular and criticizes the UEW's idea that Mari \**j*- could be secondary. According to Bereczki, Mari sometimes has prothetic *j*- before *i*, *ö*, *ü* but not before *u*. Mari *u* would also be an irregular reflex of Uralic \**i* (Aikio [2014b: 157] lists the word among the unexpected instances of Proto-Mari full vowel \**u*; a regular reflex of Uralic \**i* in disharmonic *-a*-stems would be \**ü*), so the Mari word can in no way be regularly derived from \**iša* or \**jiša*.

Furthermore, it is difficult to reconstruct the stem vowel of the Proto-Uralic word, because the Mordvin, Saami and Finnic forms all point to different vowel combinations. The Saami forms unambiguously point to an \**-a*-stem. As noted by (Aikio 2015b: 38) and Holopainen, Junttila & Kuokkala (2017: 117), the Proto-Finnic \**-o* in this word cannot be explained through Finnic word-formation. As the Mordvin forms *jož*, *jožo* also display irregular stem-vocalism, one could argue that this belongs to the category of old (Proto-Uralic) \**-o* words, that Aikio (2015b) has recently discussed. However, as Finnic *-o*-stems have been very frequent at some point of Proto-Finnic, the *-o*-vocalism can still have secondary vocalism, although its precise origin cannot be explained and requires further research. Finnic *iho* and Saami \**esē* could then theoretically both go back to a PU form \**iša*. The Mordvin and Mari words cannot be derived from such a proto-form. Komi *ež* could actually reflect a similar proto-form, as Komi *e* is the regular reflex of Uralic \**i-a* vocalism (parallel cases include the Indo-Iranian loans \**wiša* 'venom, green' > Komi *vež*).

Another problem is that the substitution \**š* ← \**dz* is unexpected word-internally, as one would expect \**č* (> Finnic *t*) or \**ks* here. There are a couple of examples of the substitution \**ks* ← \**dz* (\**maksa-*, \**paksu*), and \**č* would be the phonetically expected substitution. As noted above (see \**iša*), Pystynen (2019: 45) supports the old view that Finnic *h* can reflect PU \**čč* (the single \**č* develops to \**t*) and has assumed that in the loanword *iha* the Iranian cluster \**sts* was substituted by \**čč*, which then regularly develops into Finnic *h*. Here one cannot argue that *čč* could have been the substitute for Iranian \**dz*, as this would explain only the Finnic words, and not the consonantism of the other Uralic forms, which point regularly to single \**š* (cf. \**pičča* > Mari *piče*, Udmurt *puč*, Komi *poč*).

The etymology of the Avestan word is also not clear. Following Bailey (1979: 484), Clackson (1994: 90, footnotes 15, 16) is supportive of the connection of the Avestan to Greek αἴξ 'goat' and Armenian *ayc* id. According to Clackson, all these forms reflect PIE \**h<sub>2</sub>eig* 'goat' (the Armenian *i*-stem possibly reflecting a PIE \**h<sub>2</sub>ig*-*ih<sub>2</sub>*- with the feminine-collective suffix). Bailey (1979: 484) derives the Avestan word from Iranian \**iza-* 'hide, skin', but this may be a later semantic development from 'goat'. The Avestan word includes the suffix *-aēna-* that is usually used to characterize adjectives of material, although it can have a wider meaning. Clackson furthermore connects phonologically irregular Sanskrit *edaka-* 'sheep, wild goat', as well as Albanian *edh* 'kid, goat' here as possible cognates. Schindler (1972: 76) presents the same cognate set and argues that *edaka-* includes the suffix *-aka-* which is added to *ed-* that is abstracted from a falsely analysed plural instrumental form \**h<sub>2</sub>eig<sup>h</sup>iyos-* > \*\**ed<sup>h</sup>iyas-* where *d* reflects a regular sandhi development (see Thieme 1953; the same explanation is also mentioned by EWAia I s.v. *eda-*). Clackson also ambiguously mentions that Old Indic *ajā-* 'goat' (< ? PIE \**aǵa-* 'goat') might be etymologically related to \**h<sub>2</sub>ig*-, but it remains unclear to me how these two Indo-European words with different vocalism could be related. Beekes (2010: 41 s.v. αἴξ) is more critical of the connection of these Indo-European words and assumes that

the Greek and Avestan words, as well as OI *ajá-* ‘goat’ can be parallel borrowings from some (substrate) language.

As the Indo-Iranian etymology for the Uralic words turns out to be very unlikely, one could ask what is the origin of this group of words with irregular sound correspondences. This kind of irregularity is typical of parallel borrowings, and it would be tempting to consider the Uralic words as loanwords from somewhere. Although a precise donor language cannot be established, it is interesting to note that many words which have West Uralic or Finno-Permic distribution manifest irregular correspondences and often include the phoneme \*š. A number of such cases have been listed by Aikio (2015b: 44–47; see also Zhivlov 2015), who assumes that the words in question are substrate words that have been borrowed independently into the already disintegrated western branches of Uralic. Although most of these words deal with flora, fauna and agriculture, it is not unthinkable to assume that this word for ‘skin, hide’ could belong to the same layer of substrate words, especially if the word originally referred to the skin of animals.

Not an Indo-Iranian etymology

**Fi *ihta* (dial., obsolete) ‘lust, eagerness (Lust, Begriede, Eifer); (archaic, in early written Finnish) chance, possibility, liberty’; eagerness, will, hurry, (meaning attested in Virolahti, south-eastern dialect); quite (= aivan, vallah; in some western dialects of Finnish)’; ? conjugated from *ihan*, dial. *ihran*, *ihlan* (< \**ihtan*) ‘just, quite (ganz und gar)’; other conjugated forms: (early written Finnish) *idhoijlansa* ‘loose, free (valloillaan, vapaana)’, (South-Western and Satakunta dial.) *ihdallansa* id., (in some western dialects) *ihdaltaan*, *ihdoillaan*, *ihdoiltaan* ‘spontaneously (itsestään (esim. tulla kipeäksi))’ < PFi *ihta* (SSA s.v. *ihan*; SMS)**

< Pre-Fi \**išta*

← PII/PI \*(H)*išta-*. cf. OI *išta-* ‘wished, desired; wish, desire’, (Old) Avestan *išta-* ‘Gewünschtes’ (Y 40,4) PIE root PIE \**h<sub>2</sub>eys-* (~ \**Hays-*) (EWAia I: 270–271 s.v. *Eš*’; Werba 1997: 162 No. 12 s.v. *iš-*; Cheung 2007: 158; LIV<sup>2</sup>: 260)

(Koivulehto 2001a: 366–367; Rintala 2003: 304)

See also \**iša*, \**iso*

Koivulehto assumes that *ihta* is borrowed from the verbal adjective \**Hišta-* that belongs to the same root as the possible donor form of *iha* (see above). This is a convincing etymology both phonetically and semantically, but it is difficult to date the borrowing. Koivulehto notes that the Indic word appears already in the Rigveda. The Finnic *h* (< \*š) reflects here the Indo-Iranian sibilant \*š (the result of RUKI). The Finnic word could theoretically also reflect a pre-form \**ješa*: many of the previously assumed cases of \**i-* can be reconstructed as \**je-* or \**ji-* (Pystynen 2015; Aikio 2015a: 8–10; see the entry \*(j)*išo*), but there also remain some cases where PU initial \**i-* still has to be reconstructed, such as \**ipsi* ‘smell’. The limited distribution of the word could point to a (relatively) late borrowing, probably from Proto-Iranian, as noted already by Koivulehto. The word appears already in the

oldest Finnish texts of Mikael Agricola, but because these texts are mostly based on the Southwestern dialects, where the word appears anyway, this does not mean that the word has to be old or that its dialectal distribution has been wider in the past.

Indo-Iranian *\*(H)išta-* has possible Greek cognates that are formally quite similar. According to Chantraine (2009: 448–449) Greek ἰότης ‘will, decision’ (attested only in dative ἰότητι, ἰότητι, ἰότητι and accusative ἰότητα) is possibly from the same PIE root as Indo-Aryan *iṣ-*, *icchāti* ‘souhaiter’: the Greek form could reflect earlier *\*(H)isto-tāt-*, from the verbal adjective *\*(H)isto-* which would be formally identical to PII *\*(H)išta-*. However, Chantraine also mentions that the Greek word might instead be related to *Ἰεμαι* (*\*Fistos* ‘voulant’). Also Beekes 2010: 595–596 mentions the two possibilities (*\*h<sub>2</sub>iso-teh<sub>2</sub>-t* or *\*h<sub>2</sub>isto-teh<sub>2</sub>-t* in Beekes’s reconstruction). If the Greek word is from *\*Fistos*, then according to Beekes we have to assume haplological shortening. Beekes notes that the Boeotian personal name *θειοφιτος* points to a form beginning with *\*F-*, which is an argument in favour of the latter etymology for the Greek word.

Semantically the etymology is relatively unproblematic. The Indic, Avestan and Finnic words are close to each other, but because of the polysemy found in Finnish and Vedic it has to be admitted that the exact meaning of the borrowed form is difficult to determine. All the meanings in Finnic can probably be derived from ‘desire’, which is also a meaning attested in Vedic. Words relating to affection are easily borrowed (Saarikivi 2014: 200–201).

It is impossible to connect *ihta* to *iha* etymologically through Finnic internal derivational processes. Because of the semantic and phonetic plausibility and lack of competing etymologies, it is probable that the etymology is correct.

An Indo-Iranian etymology

**Est *isu*, Sest *iso* ‘appetite’; in Ingrian and Finnish attested in the verb *isota* ‘to have appetite’ (< PFI *\*iso*)**

← (Old) Iranian *\*isa-* ‘to seek, to pursue, to want’, cf. Av *isəmna-* ‘seeking (suchend, begehrend)’, *isaite* ‘he/she pursues’, cf. OI *icchāti* ‘seeks, wishes’, < PII *\*(H)iš-śc-a*, root *\*Hays* (Cheung 2007: 158; LIV<sup>2</sup>: 260; EWAia I: 270–271)

(Koivulehto 2001a; Katz 2003: 148; EES s.v. *isu*; Holopainen, Junttila & Kuokkala 2017: 117)

See also *\*iša*, *\*išta*, *\*ešti-*

Koivulehto’s recent etymology has not received much attention in the literature. Katz (2003) has independently suggested the same Indo-Iranian etymology, although his arguments differ from those of Koivulehto. EES s.v. *isu* mentions the etymology, although it is not discussed very informatively (EES simply notes that the donor can be identical to Old Indic *icchāti* ‘wishes’, which does not reflect Koivulehto’s careful argumentation very well). EES also considers the etymology as uncertain because it lacks parallels in other Uralic languages.

This is one of the etymologies where the Finnic *-o* is not a transparent result of a derivational process, and thus could be projected back to some earlier proto-language. However, the Iranian word, which Koivulehto considers the original of the Finnic word, does not have any labial element in its non-initial syllable, so the Finnic *\*o* can be a

substitution here (this is hinted at by Holopainen, Junttila and Kuokkala [2017: 117]), reflecting the realization of the Iranian phoneme *a* and its reception by Uralic speakers.

However, it has been noted (Lehtimäki 1972: 149–150; Nikkilä 1988: 138) that there are several *o*-stems among the Germanic and Baltic loans that were borrowed into this stem type without apparent reason or they are such early derivatives that no traces of underived labial vowels are found in any varieties of Finnic. The underived *-o* here thus does not automatically imply that Iranian *-a* was substituted by Pre-Finnic *\*o*.

The Iranian etymology itself is credible. The Finnic word has to be a relatively late loan (later than Proto-Iranian) because *\*s* in the Avestan word reflects PIE *\*sk* or *\*sk̑* (> PII *\*śc̑*, see Kobayashi 2004: 67–74; Lubotsky 2001a). In Proto-Iranian the word was probably *\*(H)istsa-*, and this would have likely resulted in Pre-Finnic *\*iča*, not *\*isa*. Katz (2003: 148) argues that *iso* could be borrowed from from PII *\*isčā-* (“FUA *\*iščā-*”) as *\*iša* or *\*iča* and then developed regularly to Finnic *\*iso*. Although this kind of development is in principle possible, the problem is again the PII cluster *\*śc̑* or *\*śc̑̑*, as we do not know the exact phonetic nature of this cluster and how it would have been substituted. Many of the Indo-Iranian loanwords that are found only in Finnic and other West Uralic languages are probably later Iranian loans, and there are no convincing examples of extremely archaic loans that would be confined to Finnic. As the Finnic word would be a regular outcome of a loan from Iranian *\*isa-*, the word probably belongs to the same loanword layer as other relatively late loans like *\*wasa* and *\*warsa* (see Koivulehto 1999a: 226–227).

Semantically the etymology is plausible, although the Finnic word seems to point exactly to ‘appetite’, whereas the meaning of the Iranian word is more general, at least based on the Avestan attestations. This kind of semantic relationship (the restriction of meaning) is typical of loanwords (see the Introduction). The only problem is that no precise donor form can be shown, as there is no noun *\*\*isa-* known from Avestan or any other Iranian language. This is a similar problem as with the *\*iša* word discussed above. However, in this case the substitution with Uralic *s* from Iranian *s* is unproblematic, and there is less reason to doubt the etymology. Because of the relatively small size of the Old Iranian corpus it can be speculated that a corresponding noun once existed in Iranian. In any case it seems that Finnic *isota* is derived from a noun *\*iso* and not the other way around.

An Indo-Iranian etymology

**Fi *isä* ‘father’ (cognates in Ka, Lu, Ve, Vo, Est, Izh, Li), SEst *esä*, *edsä* id.; SaN *áhčči* (irregular), SaKld *âž’ž’* < PSa *\*ečē* (Lehtiranta 2001: No. 2.); Md *očü* ‘father’s older brother’; Mari M *iza*, H *əzä* ‘older brother; father’s younger brother’ (< PMa *\*ičä*); Hu *ős* ‘ancestor’ (< ? OHu *ifemucut* ‘our father’); Mansi East, West *ās* ‘mother’s father’ (< Proto-Mansi *\*ās*)**

< ? PU *\*ičä* ~ *\*ečü* (Sammallahti 1988: 541: *\*išä*; UEW s.v. *ičä*; SSA s.v. *isä*; Bereczki 2013: s.v. *iza*; EES s.v. *isa*; Aikio 2014c: 154)

← PII *\*(H)i(H)ič-* ‘master, lord’ > *īs-* M ‘Herr’, *īsä-* M ‘Herr, Gebieter’; root PII *\*Haič* < PIE *\*h<sub>2</sub>eyk-*, cf. OI *is-* ‘beherrschen, Herr über etwas sein, gebieten,

**Gewalt haben’, OAv *isē* ‘ich verfüge’, *isuan-* ‘Gebietler, Mächtiger’** (EWAia I: 207, s.v. *Īś*; Werba 1997: 424 No. 541 s.v. *\*iś*; RIVELEX 2: 240ff. s.v. *īś*)

(Tunkelo 1913: 118–119; Koivulehto 2001a: 371–375; = 2016: 269–273; Milanova, Holopainen & Bradley [forthcoming])

The Indo-Iranian etymology suggested by Tunkelo and supported by Koivulehto involves several problems, mostly in the Uralic side. Although the etymology could in principle be correct, the irregularities in the Uralic cognate set prevent us from reconstructing a Proto-Uralic word, which means that it is unlikely that the etymology is correct in the way it was suggested by Koivulehto. Parallel borrowings into already differentiated branches would have to be extremely early. The word cannot be borrowed from Iranian, as the Uralic forms clearly point to an earlier *\*ć* (if the word was borrowed from Iranian, we would have *\*\*ičä* or *\*\*isä* in Uralic). The Indo-Iranian root begins with a laryngeal which causes the later long-vowel forms in Avestan and Sanskrit. As with most loanwords into Uralic, the laryngeal is left without a substitution here.

The Finnic words could reflect *\*ičä* that is the reconstructed form found in most sources. Aikio (2014c: 57) mentions a South Estonian form *jedsä*, which according to him is “old Võro”, meaning the old South Estonian literary language, but in reality this form is attested only in the Seto dialects of South Estonian, and there only as a Lallwort (EMS s.v. *esä*). There is thus no reason to reconstruct an earlier affricate for the Finnic forms on the basis of this word only. The normal South Estonian word is *esä*. Koivulehto (2001a: 271, footnote 273) mentions the South Estonian word as an irregular form, and SSA notes that *esä* corresponds regularly to the irregular forms of North Saami and some other Saami languages (see below). Kallio (2018b: 131–132) has shown that, based on early South Estonian written sources, the sound change *\*i > e* has happened in the modern era, and the South Estonian form reflects earlier *isä*. It has to be purely accidental that some Saami forms could reflect an apparently similar pre-form as the South Estonian word.

The Saami words point to different reconstructions. Lehtiranta notes that the western Saami forms point to Proto-Saami *\*āććē*, whereas Inari and Ter Saami point to *\*ēćē*. The Skolt and Kildin Saami forms regularly reflect Proto-Saami *\*ēćē* which is a regular cognate to the Finnic word and reflects the possible Uralic reconstruction *\*ičä*. Koivulehto (referring to Sammallahti 1998: 227) remarks that the Saami forms which point to *\*ā* are secondary, and the vocalism is influenced by the palatal affricate *\*ć*.

The Mordvin word is considered uncertain by EES and UEW, and SSA mentions that it does not belong here at all. The word is irregular, as *o-* points to earlier *\*i-a*-stem, and the stem vowel is totally irregular. One can of course argue that affection has played a role in the development of this kinship term, but it is impossible to prove this. No alternative etymology for the Mordvin word has been represented as far as I know.

The Mari word can regularly reflect earlier *\*ičä*, with the regular development of *\*ć > z* intervocalically (see Aikio 2014c: 87). The Indo-Iranian etymology is not mentioned by Bereczki (2013 s.v. *iza*), who supports the traditional Uralic etymology of the Mari word.

The Hungarian word *ős* with *ő-* is problematic, as this is not the usual reflex of PU *\*i*. Also the consonantism is irregular, as *\*ć* is reflected by Hu *s* (i.e. [š]) in a couple of cases only, most of which are uncertain. It is possible that the Hungarian word does not belong here at all, as the meaning is also divergent from the rest of the Uralic forms. It is true that

Old Hungarian *isemuc* means ‘our father’, but it is not even certain that the modern Hungarian form could have developed from this, as Old Hungarian *i* is usually not reflected by modern Hungarian *ő*.

The Samoyed word forms a special case here. According to Koivulehto, Proto-Samoyed *\*ejsä* (> Tundra Nenets *nīše*, SW: 22) is a parallel borrowing from an Indo-Iranian form with the full-grade diphthong *\*ai-*, (*\*Hayć-*). The Proto-Samoyed form could reflect Proto-Uralic/Pre-Samoyed *\*äjčä*. Uralic *\*äj* would be a phonetically plausible substitution for the diphthong *\*ey* in the full-grade form of the root (the word could be a Pre-II borrowing from *\*Heyć-?* Or *\*äj* from *\*Hay-?*). Borrowing from another branch of Indo-European is unlikely, as the root is attested only in Indo-Iranian and Germanic (cf. Gothic *aih* ‘ich habe, besitze’, OHG *hūs-eigo* ‘Hausherr’).

Furthermore, the Nganasan form (*jase*) is irregular and points to PS *\*jesä*, which makes the reconstruction of even the Proto-Samoyed word very difficult. It is possible that the Samoyed and “Finno-Ugric” words are related, but the reconstruction is prevented by unknown conditioned developments. Aikio (2014c: 57) reconstructs the Uralic word uncertainly as *\*ičä/\*ecä* and notes that the Samoyed cluster *-js-* is irregular.

Aikio (2014c: 57) mentions that Yukaghir *ečē* ‘father’ is borrowed from Samoyed, but at the time when the Samoyed affricate was still retained. This means that for the Uralic word certainly an affricate has to be reconstructed. Also the fact that the word was borrowed into Yukaghir very early from Samoyed means that the possible borrowing from Indo-Iranian into Uralic/Samoyed should be dated very early.

Semantically the etymology would be unproblematic. As Koivulehto notes, some other kinship terms, like the Finnic word for mother (← Proto-Germanic *\*aiþīn-*) are loanwords in the Uralic languages, so it is no obstacle to consider the word for ‘father’ a borrowing. Other kinship terms borrowed from specifically Indo-Iranian include at least *\*sasar(V)*.

Semantically the loan is not obvious, but the development from ‘lord’ to ‘father’ is possible. Koivulehto notes that here the speakers of Uralic did not borrow the normal Indo-Iranian word for ‘father’ but they borrowed the word in the meaning of ‘the head of the family’ (“Lehnnemenden das IE Wort als ‘Herr, Besitzer einer Grossfamilie’ kennengelernt haben”). This is semantically plausible. It is well known that even the “normal” Indo-European word for ‘father’, *\*ph<sub>2</sub>tér* originally denoted something else than the pure biological father – the word’s primary meaning, still found in the Latin expression *pater familias*, is the protector, head of the family. Against this background it might be easier to understand why the term *\*ph<sub>2</sub>tér* or its Indo-Iranian reflex *\*pHtār* was not borrowed, as the semantics of these terms were possibly still under development at the time of the early contacts between Indo-Iranian and Uralic contacts (see also Milanova [forthcoming]). Interestingly, the Indic word means in later texts also ‘husband (Gatte, Ehemann)’.

An interesting semantic parallel to Koivulehto’s suggestion is provided by some later Iranian languages, which use forms borrowed from Turkic *ağa* ‘lord’ in the meaning of ‘father’; for example, New Persian *aqa* ‘mister; father’, in Dari also ‘elder brother’ (Veronika Milanova: personal communication).

Even though the semantics pose no problem for the Indo-Iranian etymology, the phonological problems discussed above force us to abandon the etymology.

Not an Indo-Iranian etymology

TN *nīse*, E *ese*, Nga *jase* ‘father’, Slk *esi*

PS \**äjsä* < Pre-S ? \**äjčä* (SW: 22)

See \**ičä* ~ \**ecä*

? Indo-Iranian etymology

**Fi *isäntä* ‘master’ (cognates in all Finnic languages)**

< ? Pre-Fi \**ičäntä* or \**isäntä*

← PII (or PI) \*(H)i(H)*ičāna-* > OI *īsāna-* ‘ruling, dominating’, Av *isāna* ‘ruling over something’ (present medium participle from the verb *īś-*; EWAia I: 207, s.v. *īś-*; for the II root, see the above entry on \**ičä*)

(Tunkelo 1913: 115–118; Koivulehto 2001a: 371–375; = 2016: 269–273)

The Indo-Iranian etymology is an interesting possibility, but not something one can accept without criticism. Koivulehto argues that *isäntä* is not a derivation from the word *isä* ‘father’, though this is the view of most linguists as well as native speakers of Finnic languages (the word *isä* also has a possible Indo-Iranian etymology, see above), just as the word *emäntä* ‘mistress’ looks like a transparent derivation from the old Uralic word for ‘mother’, *emä*. Koivulehto notes that the word should reflect earlier \**isänä*, and the *nt* is irregular. Koivulehto claims that there are parallels for this development, such as doublets *sarana* ~ *saranta*, *sarvena* ~ *sarventa*, but it is suspicious that in this case we do not have doublets but only the form with *-nt-*.

Koivulehto mentions that Indo-Iranian etymology for the word has been proposed earlier too by Tunkelo (1913: 115–118) from the active participle \**isant-* ‘besitzend’ < PII \**ičant-* of the same verb, but this participle is not attested in any Indo-Iranian language, which makes the etymology unlikely.

Semantically the Indo-Iranian etymology is problematic, as the Indo-Iranian root denotes ‘ruling, mastering’, and the meanings of the noun *isäntä* could easily be derived from this. An argument against the native derivational origin of the word *isäntä* (as well as *emäntä*), also mentioned by Koivulehto, is that the derivational processes behind these words are not properly known. The *isä-* and *emä-* elements are transparent (or look like that), but the derivational suffix is not; this has been considered a petrified, unproductive derivational suffix (Raun 1988: 565), but it is problematic to find it only in these two words.

The geminate *-ss-* in Estonian *issand* ‘lord’ is secondary according to EES (s.v. *isa*).

Not an Indo-Iranian etymology

Fi *vuosi* ‘year’; Ko *vo* ‘year’, Ud *wa-pum* ‘time, moment’ < PP *\*wo*; Kh E *al* ‘year; NT *ŋeŋo* + cognates in Slk, Kam, Mat (< PS *\*erö*)

< PU *\*iði* ‘year, autumn’ (Aikio 2012a: 233–234)

← PII *\*watas-* ‘year’ > OI *vatás-*, < PIE *\*wet-os-* (see *\*wasa* for references)

(Koivulehto 1999a: 218–219 = 2016: 220–221; SSA s.v. *vuosi*)

Koivulehto’s etymology is convincingly rejected by Aikio (2012a). Although the etymology is semantically plausible, Aikio’s improvements to the phonological reconstruction of the Uralic word make the etymology impossible: word-initial *\*w-* cannot be reconstructed for the Proto-Uralic form, although one would expect it if the word was borrowed from this Indo-Iranian form, and the consonants PU *\*δ* and PII *\*t* do not match (there are no examples of such a sound substitution).

Borrowing from related Indo-Iranian forms has been suggested for *\*wasa*, *\*watsa* and Mansi *\*wēšay* (see the entries).

Not an Indo-Iranian etymology

Komi *ĩšti-* ‘be charmed, become delighted; wish, desire, long; be very gay, frolicsome’; *ĩšmi-* ‘be animated; be frolicsome, licentious; be desirous, lustful’ < PP *\*iř-* ?

See *\*iřa*

Fi *jama* ‘Fuge, Naht zwischen den Teilnetzen des (Zug)netzes; Grenze; Vereinigungspunkt; (passender) Zeitpunkt; (schlechter) Zustand’, verb *jamata* ‘to join together (seine nets)’ (cognates in Karelian, Ludic, Votic and Estonian); SaN *juopmi* ‘twin’ < PSa *\*juomē* ‘twin’ (missing from Kld and T); ? Erzya- Mordvin *jav* (in *kijaks-jav* ‘floor joint’ (SSA s.v. *jama*; Álgu s.v. *juopmi*; SMS s.v. *jama*)

< ? West Uralic *\*jama* or *\*jīma*

← PII *\*yamá-* > OI *yamá-* ‘twin’, *yamí-* ‘Zwillingschwester’ (? > *yamá-* ‘Name einer mythischen urchenischen Gestalt, des ersten Sterblichen und Königs im Totenreiche’); root *yam-* ‘halten, festhalten, strecken, zügeln, lenken, ausstrecken, darreichen’, OP/Av *yam* < PII *\*(H)yam* < PIE *\*(H)yem-* (EWAia II: 400–401 s.v. *YAM*; LIV<sup>2</sup> s.v. *iem*; Werba 1997: 221–222, No. 150, s.v. *yam*)

(T. I. Itkonen 1928: 25; SKES s.v. *jama*; Katz 1985: 374; 2003: 296; SSA s.v. *jama*; Sammallahti 2001: 408; Koivulehto ms.; Blažek 2005: 163; Holopainen 2018b: 147–148)

Although the Indo-Iranian etymology by T. I. Itkonen is semantically and phonologically convincing, the etymology is not mentioned by Joki (1973), Rédei (1986c) or the UEW (s.v. *jama*). The Finnic and Saami words are regular cognates and can reflect either *\*jama*

or *\*jīma*. Although the meaning of the Finnic word is rather different from that of the Saami and the Indo-Iranian original, the semantic development can be explained and it is probable that Saami has retained the original meaning here. SSA is dubious about the semantic relationship between the Finnic and Saami words. However, EWAia notes that the Indo-European root *\*(H)yem* from which the noun ‘twin’ is derived might have originally denoted ‘to keep together’. From this the meaning of ‘twin’ can be derived, and the semantic relationship between ‘twin’ and ‘keep together’ in Indo-European is not far from that of the Finnic and Saami words, and I see no reason to reject the etymological connection between Finnic *jama* and Saami *\*juomē*.

However, it is not certain whether *\*yamá-* is derived from this root; LIV<sup>2</sup> (s.v. *jem*) also remarks that it is uncertain whether the noun *yama-* belongs to the root *\*(H)yem-*. It is also uncertain whether the name of the mythical being *yama-*, Av *yima-* is derived from the same Indo-Iranian root, but EWAia assumes that this is probably the case. If the original meaning of the Uralic word is ‘twin’, the further connections of PII *\*yamá-* within Indo-European have no relevance for the etymology of the Uralic word.

As the noun *jama* can be reconstructed for a common proto-language of Finnic and Saami, it is probable that the word was borrowed as a noun. The Finnic verb *jamata* is probably derived from the noun *jama* and not the other way around.

The Mordvin word probably does not belong here, as *v* is not a regular reflex of PU intervocalic *\*m* (compare PU *\*d’īmi* > MoE *l’om* ‘bird-cherry’, Aikio 2015b: 57). Mordvin *-jav* is connected to Finnic *jakaa* in SSA (s.v. *jama*), and it is probable that it is a cognate of this Finnic word, as this would be phonologically totally regular. In Mordvin there is also a verb E *java-*, M *java-* ‘to divide’ which is derived from Uralic *\*jaka-* (UEW s.v. *\*jaka-*; not mentioned by Aikio 2015b), but this is not an obstacle to derive the noun *-jav* also ultimately from the same Uralic stem. Semantically the connection is not obvious. Interestingly, there is also a verbal root *\*yaw-* ‘to (re)bind, join’ in Indo-Iranian (Cheung 2007), and the Mordvin noun could perhaps be borrowed from this (although precise donor form cannot be shown).

Phonologically, the Indo-Iranian etymology is faultless. Although nearly similar formations with the meaning of ‘twin’ are found in Germanic (*\*jumja-* ‘divine twin’ > ON *Ymir*, Kroonen 2013 s.v. *jumja*) and Lithuanian (*jumis* ‘paired couple, double fruit’), these would not be suitable originals for the Finno-Saamic word because of the vocalism, and the Indo-Iranian origin is most probable.

Vainik (2014) has considered the word *\*juma* ‘god’ a borrowing from a reflex of this same Indo-Iranian root *\*Hyam-* (see the entry for *\*juma*), but in the light of vocalism it is more likely that *juma* is borrowed from PII *\*dyuman-*.

Katz has suggested a different etymology, assuming that the Uralic word is borrowed from Pre-II *yugma-* ‘paarig’ (> OI *yugma-*). While this would be semantically certainly convincing, the vocalism is problematic in this case and Katz’s etymology cannot be supported.

An Indo-Iranian etymology

**Fi *jyvä* ‘grain’, SEst *iva id.*; ? Md M *juv* ‘Acheln, Spreu’; Ko *id (deriv.)* ‘Gerste’, *jī-ki* ‘Granne an Ähren’, Ud *ju* ‘Getreide’, *jīdī (deriv.)* ‘Gerste’ (< PP \**ju*)**

< Pre-Fi/PU \**jewä* (UEW s.v. \**jewä*; SSA s.v. *jyvä*)

← Pre-II \**yewo-*, > Vedic *yáva-* ‘barley, wheat, grain’, Av *yauua-* *id.*; < PIE \**yewo-* ‘cattle fodder (?)’, > Tocharian B *yap* ‘barley’, Lithuanian *jāvas* ‘corn’, Hittite /*ea*/ ‘eine Feldfrucht’ (EWAia II s.v. *yáva-*; 2017: 292–293; Mallory 2015: 19; Isebaert 1980: 98–99; Derksen 2015: 210)

(Toivonen 1928: 229–30; FUV: 130; KEWA s.v. *yáva-*; Joki 1973: 265; Katz 1985: 265; 2003: 212–213; UEW s.v. *jewä*; Dolgopolsky 1989: 20; Koivulehto 1983; ms.; Pystynen 2015; Kallio ms.)

This is a convincing etymology which is accepted in most sources, but there are various views about the exact Indo-European donor language and the relationship of the Uralic words. Moór (1956: 343–344) assumes that the donor language was some archaic Indo-European language, but he does not specifically mention an Indo-Iranian origin. Moór still lists the Samoyedic words like Tundra Nenets *ja* ‘flour’ as cognates to the Finno-Ugric words (it is also not observed by him that the Finnic, Mordvin and Permic words are not regular cognates). Moór considers it possible that the Samoyedic word is borrowed from Pre-Hungarian, so not directly from Indo-Iranian. This idea is related to Moór’s outdated views about Hungarian linguistic prehistory: Moór notes that although the word has no attested reflex in Hungarian, this could be explained by assuming that the word had become lost because it resembles two other Hungarian words, *jó* ‘good’ and the now obsolete *jó* ‘river’. This is not a plausible idea, because the word \**jewä* would have a front-vocalic reflex in Hungarian, and because the forms in Finnic, Mordvin, Permic are probably parallel borrowings (see below). The Samoyedic word for ‘flour’ is now considered to be a parallel borrowing from the same source, as noted by Janhunen (1983) and K. Häkkinen & Lempiäinen (1996); see below.

KEWA accepts the Indo-European etymology of the Uralic word but does not comment on the donor language more specifically. Joki (1973) argues that the words in Finnic, Mordvin and Permic could perhaps be later borrowings from an Iranian language with *æ*-vocalism, like Ossetic *jæw*. This is not probable, as Iranian \**ä* would have rather resulted in Uralic *ä*-vocalism, but none of the Uralic forms point to old \**ä*. Finnic points to \**e*, and Mordvin and Permic words reflect earlier back vowels. Rédei (1986c: 50–51) states a confusing conclusion that the word is borrowed from early Proto-Indo-Iranian or early Proto-Iranian, but Rédei reconstructs \**e* for both forms. Also Dolgopolsky (1989: 20) gives a confusing reconstruction of “Middle pA [= Proto-Aryan] \**jewa* or \**jewo*” (here \**j* = \**y*), which does not reflect the modern understanding of Proto-Indo-Iranian phonological reconstruction.

The UEW assumes that \**e* is possibly caused by the influence of \**j-* and the word could have been borrowed from Proto-Indo-Iranian \**jawa*, but there are no parallel examples to this development and this explanation cannot be supported. The non-initial syllable suits the substitution pattern \**-ä* ← \**-o* (in front vowel words) found regularly in Pre-II loanwords. The Finnic *ü* might be a secondary development due to the labializing element *w*, as there are other examples of such development too (see Pystynen 2015). It is also

possible that the word was borrowed from some other Indo-European language. Garnier et al. (2017: 292–293) point out that the original meaning of the word might have been ‘cattle fodder’, and the meaning ‘grain’ is a later development. This would mean that an Indo-Iranian etymology is likely, not a Proto-Indo-European one. However, a similar semantic development has also led to the Baltic words, such as Lithuanian *jāvas* ‘corn, grain’ (usually encountered as plural *javai*; Derksen 2015: 210), and there is no reason to consider a borrowing from Indo-Iranian more plausible than an early borrowing from Balto-Slavic.

Mordvin *u* is explained in the UEW as a result of influence from the word-initial *v*. Even if this was a possible explanation, one can also consider the Mordvin word a loan from Proto-Indo-Iranian (or Proto-Iranian) *\*yawa-*; this would explain the *u*-vocalism easily, as one can suppose that the Mordvin word was borrowed as Pre-Mordvin *\*jowa* (*\*o* as the substitution of II *\*a* because of the labial environment), and from this Mordvin *u* would regularly develop. One has to admit that in *\*a*-stems Mordvin has usually retained a reflex of the word-final vowel (Bartens 1999: 60–66; Kallio 2012a: 163), and a form like *\*\*juvo* would thus be expected, but here one can note that in the derivation *juvodo-* ‘worfeln, schwingen’ the vowel has been retained.

The Permic words show anomalous vocalism, and here UEW is probably right in assuming that *\*w* has had an influence on the vowel developments. In no case can the Komi and Udmurt forms reflect *\*jewä* like the Finnic words, so they are also most clearly parallel borrowings from (Indo-)Iranian. This has been noted also by Koivulehto (2006: 160).

K. Häkkinen & Lempiäinen (1996: 154–156) support the idea that the word is borrowed from Indo-European separately into the various branches of Uralic. They also consider Koivulehto’s idea of a Pre-Baltic (Balto-Slavic) origin of the Finnic word possible but note that it is problematic to assume that the word is borrowed from Baltic to Permic or first into Finnic and from Finnic to Permic. Recently the etymology has been discussed briefly by Blažek (2016: 55), who does not note the phonological irregularities but assumes that Finno-Permic *\*jewä* is borrowed probably from Indo-Iranian. Pystynen (2015) supports the idea that the Uralic words are parallel loans.

Kallio (ms.) argues that the Finnic word is not borrowed from Indo-Iranian but rather from Proto-Balto-Slavic. According to Kallio, Proto-Balto-Slavic *\*ew > \*ow* is a later change than Indo-Iranian *\*e, o, a > \*a*. Kallio still assumes that the Mordvin, Permic and Samoyed words are borrowed from Indo-Iranian. This is a convincing idea, although theoretically the Mordvin and Permic words could also reflect later borrowings from Balto-Slavic or Baltic.

In any case the words have to be parallel borrowings from Indo-European, and it is obvious that the Finnic, Mordvin and Permic words cannot go back into Proto-Uralic or “Proto-Finno-Permic”. The early idea that the word is borrowed from Proto-Indo-European is not very likely for any of the Uralic words in the light of the substitution and semantics, even though phonetically Finnic *\*jüwä* could equally well reflect PIE *\*yewo-*.

The Samoyedic word for ‘flour’ is a parallel borrowing from the same source, as noted by Janhunen (1983) and Häkkinen (1996). In some earlier sources, such as FUV, it is considered possible that the Samoyed word is related to the words in the Finno-Permic languages to the west (so the borrowing would have happened before the split of the Samoyed languages), but it is impossible to derive the Samoyed word regularly from

\**jewä*. The semantic development from ‘grain’ to ‘flour’ is easy to explain, especially as the speakers of (Pre-)Proto-Samoyed probably did not practice agriculture. Blažek (2016: 55) assumes that the Samoyed word is borrowed from Komi.

? Indo-Iranian etymology (the Finnic word probably from Balto-Slavic)

**TN *ja*; Nga *ja*; FE *žāsa* id. ‘flour’**

**< PS \**jäā* ‘flour’ < ? Pre-S \**jäwi***

**← (P)I \**yawa-* or later (Alanic) \**yæw-*, > Oss *jæw* ‘grain’** (see \**jewä* for references)

(Janhunen 1983: 123–124; Joki 1973: 265; Häkkinen & Lempiäinen 1996: 154)

The Samoyedic word’s relationship to Finnic *jyvä* and the Mordvin and Permic words with similar meanings has a long history, but it is now obvious that they cannot be regular cognates. Probably the Samoyed word is borrowed separately from Indo-Iranian, like Joki (1973: 265) and Janhunen (1983: 123–124) suggest. Their arguments are accepted by Häkkinen & Lempiäinen (1996: 154). Rédei (1988a: 659) argues that the Samoyedic word is borrowed from Proto-Permic or an early form of Komi, but there are no arguments to support this instead of the direct borrowing from Iranian.

Samoyed \**jäā* is clearly an Indo-Iranian loanword, but the exact stage from which it was acquired is difficult to determine. The Samoyed word could probably reflect earlier Pre-Samoyed \**ä-i*-stem with \**ä* as the substitution of Iranian \**a*). The word has to be an old \**i*-stem because \**w* has been lost (see Aikio 2002: 49) which is unexpected in the light of the Iranian \**a*-stem. The word belongs to the small layer of borrowings that are found only in Samoyedic. However, not all of these borrowings belong necessarily to the same layer. The word \**pulā* ‘bridge’ looks like a rather young borrowing, whereas this word and the words \**wārkā* ‘bear’ and \**tājka* ‘spear’ can be much earlier borrowings.

The Indo-Iranian origin is the same as for the West Uralic word for ‘grain’ (see \**jewä*). In Samoyedic the meaning has switched to ‘flour’ probably because of cultural reasons (the speakers of Proto-Samoyedic did not practice agriculture).

Janhunen notes that the word is attested only in North Samoyedic but may have been borrowed into Proto-Samoyed; as North Samoyed probably does not present a real node but is simply an “areal-genetic unit” (see Helimski 1982; Salminen 2002), the regular cognates of the word in the northern languages are enough to prove its Proto-Samoyed status. As the taxonomic validity of the North Samoyed versus South Samoyed branching has frequently been questioned, even by Janhunen (1998) himself, the borrowing into Proto-Samoyed is the most reasonable option. In Selkup and the Sayan Samoyed languages newer loanwords have probably replaced the old Samoyed word for ‘flour’: in Selkup, *muka* ‘flour’ is borrowed from Russian (Alatalo 2004: 122, No. 825), and Mator *talkan* ‘flour’ is from Russian, Mator and Kamas *un* from Turkic respectively (Joki 1952: 386–387; Helimski 1997a: 350, 374).

An Indo-Iranian etymology

**Fi *juoni* ‘row, line, plot, story’; PSa *\*juone* ‘trail’ > SaS *juone* ‘road’, cognates in SaKl and SaSk (Álgu s.v. *juone*); Mordvin *jan* ‘path’ (SSA s.v. *juoni*; SMS s.v. *juoni*)**

< PU *\*jini* (Aikio 2012b: 233)

← PII *\*HyáHna-* ‘path, way’, > OI *yāna-* ‘fahrt’, root *yā*, Av *yā-* < PII *HyaH-* ‘fahren, (dahin)ziehen’ (EWAia II: 407, s.v. *yā1*; Werba 1997: 407–408 s.v. *yā*; LIV<sup>2</sup> s.v. *ieh<sub>2</sub>*)

(FUV: 130; Sammallahti 2001: 408; Koivulehto ms.; 2009: 81–82 SSA s.v. *juoni*)

This is an early etymology, mentioned by FUV (who does not mention the Finnic *juoni*, only the Saami and Mordvin words), but the word is not discussed by Joki or Rédei. In SSA the word is for some reason considered to be of “descriptive” origin. This assumption is heavily criticized by Koivulehto (ms.). The Indo-Iranian etymology is plausible, even though certain details are obscure.

It is unclear how the word should be properly reconstructed for Proto-Uralic or Proto-West Uralic. On the basis of the Saami and Mordvin words *\*jani* can be reconstructed, but in Finnic there are multiple related words (*juoni, jana, juna, jono*) which do not stand in a clear derivational relationship with each other. However, Finnic *juoni* is a regular cognate of the Saami and Mordvin words, and in the new vowel reconstruction of Aikio (2015b) they have to reflect earlier *\*jini*, not *\*jani*. Aikio does not mention the word in his list of Uralic *\*i-i*-stems.

The word *jana* can be a late borrowing from some Germanic language: Kluge (1975, s.v. *Jahn*)<sup>19</sup> mentions Middle High German *jān* ‘line’ as an “obsolete” source of the Finnish word. Deriving the form *jana* from here would leave us with the cognate set Fi *juoni*, Sa *juone*, Mo *jan*, which is more easier to explain. The Germanic etymology has been criticized by SSA, who considers also *jana* descriptive.

LÄGLOS (I: 139 s.v. *jana*) considers the Germanic etymology of Finnic *jana* as possible but uncertain (“? Germ. LW”). Proto-Scandinavian *\*(j)ānaR* would be the most probable source of *jana*. The only reservation that LÄGLOS expresses is the unclear variation of different forms within Finnic. LÄGLOS does not comment on the relationship of *juoni* and the assumed cognates in Saami and Mordvin. Also Koivulehto (ms.) supports the Germanic origin of *jana*. The Germanic word is possibly from the same Indo-European root as Indo-Iranian *\*HyaH* (LIV<sup>2</sup> s.v. *ieh<sub>2</sub>*). Koivulehto assumes that *jono* and *juna* might be secondary (affective) variants of *juoni*, here agreeing with SSA. It is difficult to say more about their relationship. One possibility would be to assume that some of the Finnic words are borrowed from a reflex of Saami *\*juone* but neither *jono* or *juna* would be an expected outcome of such a borrowing.

The *\*i*-stem of the Uralic word is unexpected and difficult to explain. Koivulehto (2009) states that early Indo-European loans were often acquired as *\*i*-stems, but this is certainly not true of the Indo-Iranian loans where *\*a*-stems clearly prevail regardless of the stem type of the donor language, and one would certainly expect an *\*a*-stem here because of the Indo-Iranian *\*a*-stem. Uralic *\*jani* is one of the few convincing examples of an Indo-Iranian *a*-stem acquired as an *\*i*-stem, as other potential examples such as *\*sali* and *\*sarwi* (see the respective entries) can be explained otherwise.

---

<sup>19</sup> In newer editions of Kluge the etymology is not mentioned.

It is probable that at least part of the words treated here are borrowed from Indo-Iranian; for the Mordvin word Indo-Iranian is the only possibility.

An Indo-Iranian etymology

**Fi *jumala* ‘God’ (cognates in all Finnic languages); Mo E *jon-dol* ‘lightning’; Mari *jâmo*, *jumâ* ‘God’ < PMA \**jũm***

< PU \**juma* ‘God’ (UEW s.v. *juma*; SSA s.v. *jumala*; Bereczki 2013 s.v. *jumo*)

← PII/PI \**dyuman-*, > OI *dyumná-* ‘Herrlichkeit, Himmelsherrlichkeit’, OAv *diuamnəm* ‘Himmelsherrlichkeit, Himmlischkeit’, from the PII root \**dyaw-* < PIE \**dyew-* (EWAia I s.v. *dyáv-*, *dyumná-*)

(Paasonen 1907: 17–21; SKES s.v. *jumala*; Joki 1973: 90–91; Katz 1985: 223; 2003: 180; Koivulehto 1999: 227, 228, 232 = 2016: 229, 230, 234; SSA s.v. *jumala*; Parpola 2012: 161; 2015: 66; Saarikivi 2014: 207)

Paasonen’s etymology has been doubted by Joki (1973: 90–91), who comments on the etymology only in the research history part but not in the Wörterverzeichnis. Joki is mainly concerned about the second syllable of the word, arguing that \**dyumna-* would not be borrowed as \**juma*. Also SKES is doubtful of the etymology. KEWA (s.v. *dyáuš*) follows Joki in considering the etymology doubtful. Paasonen’s etymology is supported by Koivulehto but he does not comment on Joki’s main point of concern, the second syllable of the Uralic word.

Joki’s arguments are not very strong, as also among the Baltic loanwords of Finnic there are *n*-stems that are borrowed into Finnic as vocalic stems. It is also disputed whether there actually were consonantal stems in Proto-Uralic, meaning that borrowing \**dyumant* as \**juma* is expected. We also do not know what the exact source of the Uralic word was. In general, the non-initial stems have only limited value in Uralic loanword-research (Kallio 2008: 269). Otherwise the etymology makes sense phonetically, although the same substitution P(W)U \**j* ← PI(I) \**dy* is not found elsewhere among the Indo-Iranian loans. However, a near-parallel is provided by the Indo-European etymology of Finnic \**jäwtä-* ‘to chew’ ← PIE/Pre-Balto-Slavic \**gyew-*, suggested by Koivulehto (2016: 349–350). In both cases the anlaut cluster of a stop and \**j* is substituted by Uralic \**j*. Koivulehto (ms.) claims that the substitution \**j* for the cluster \**dy-* is expected in the light of the substitutions of other word-initial consonant clusters in the early loanwords from Indo-European into Uralic.

The word was probably borrowed into a common proto-language of Finnic, Mordvin and Mari; the forms found in Finnic, Mordvin and Mari are regular cognates, but the Finnic word includes a suffix *-la* which has an unknown origin. According to Saarikivi (2014: 207, footnote 11), the Finnic suffix could be connected to the Uralic word \*(*w*)*ülä* ‘up, upper part’. The *-la* would be a reduced form which has become adapted to the vowel harmony. Saarikivi argues that as the names for ‘god’ in Mordvin (*niške-pas*; see the entry \**pakas*) and Khanty (*Torem-Ilim*) contain the term ‘upper, heaven’ serve as semantic parallels to a compound that would mean ‘the upper god’. Saarikivi’s idea is interesting but

it is very difficult to prove. The precise origin of the Finnic suffix remains unexplained. Rédei (1991: 97–98) has argued that PU *\*-la*, *\*-lä* is a denominal derivational suffix, but it is highly uncertain whether this can be considered the same suffix that is found in the Finnic word. The suffix looks similar as in the Finno-Ugric (?) noun *ńoma-la* ‘rabbit’ from PU *\*ńoma* id. (TN *nyuw* id.), but since the function of the latter suffix is uncertain, it remains unclear whether these two suffixes have something to do with each other.

T. Sköld (1983: 63–64) states that the Saami words for ‘god’ like SaN *ipmil*, Kld *jimmel*, Lu *jupmēl* are independent parallel borrowings from Finnic *\*jumala*, as the Saami forms cannot be regular reflexes of *\*juma*, nor are they regular reflexes of a single Proto-Saami form.

There are also other words relating to mythology borrowed from the Indo-Iranians. Although the Indo-Iranian word does not specifically denote ‘God’, the word is a god’s epithet, and the semantic development is easy to explain. Katz (1985: 223, 2003: 180) supports the etymology, but he agrees with the old connection of *\*juma* with the Ugric words meaning ‘good’, such as Hu *jó*, arguing that the Uralic word is a calque based on PII *\*bʰaga-* and *\*vásu-* ‘good, God’ (Setälä 1900: 42). The idea that the Ugric forms could continue PU *\*juma* is not probable, as the Ob-Ugric forms rather point to an old *\*o*, and also the vowel correspondences within the Ugric languages show various irregularities. Berezki (2013, s.v. *jumo*) does not mention the Indo-Iranian etymology at all.

Vainik (2014) has found Koivulehto’s etymology unconvincing because of semantic reasons and instead considers the Finno-Volgaic word a loan from a reflex of the PIE root *\*(H)yem-* ‘twin’. While this is an interesting possibility, Vainik does little to analyse the vocalism of the borrowing, and she does not suggest a specific origin donor form, which weakens her etymology. From the point of view of vowel substitution Vainik’s etymology is not very good. Jürine et al. (2016) mention both Koivulehto’s and Vainik’s etymologies but don’t take a stance on the issue.

#### An Indo-Iranian etymology

**Fi *juo-* ‘to drink’ (cognates in all Finnic languages); SaN *juhka-* id. (< PSa *\*juke-*; cognates in all Sa languages); Ma *jüa-* id. (< PMa *\*jüa-*); Ko *ju-*, Ud *ju-* (< PP *\*ju-*); Hu *i-* (*iszik*, *iv-*) id.; TN *nger-*, Slk *ör-*, Mtr *er-* ‘to become intoxicated’ (< PS (der.) *\*e-r-*), TN *ngexəl-* ‘to drink with one gulp’, E *ihora-* ‘to drink excessively’ (< PS (der.) *\*e-kəl-*) (SW: 21–22)**

< ? PU *\*juyi-* or *\*jiyi-* (Aikio 2002: 38–40; 2015b: 65; Zhivlov 2014: 116, 134; UEW s.v. *juye-* (*ju-e*); Sammallahti 1988: 543; Berezki 2013 s.v. *jüam*)

← Pre-II/PIE *\*gʰighew-*, root *\*gʰew-* ‘to pour’, PII *\*ǵhaw-* > OI *hu* ‘spenden, libieren, opfern’, perfect *juhvé* ‘hat geoffert’, Av *\*zu* (attested in *zaoθrā-* ‘Opfertrank’, *ā-zuiti-* ‘Opferbutter’) (EWAia II s.v. *HAV*; Werba 1997: 270, No. 251 s.v. *hu*)

(Koivulehto 1991: 17, footnote 7; 1999a: 216 = 2016: 218; Aikio 2002: 40, footnote 7)

Koivulehto has argued for the Indo-European origin of this Uralic verb in already in 1991, but in (1999a: 216) he mentions specifically that the word is an Indo-Iranian borrowing. This would naturally mean a very early form of Indo-Iranian, as Koivulehto argues that PIE \*ǵ was retained at the time of borrowing. Aikio (2002) added the Samoyed verb to this comparison, assuming that the word-initial \*j- has become lost in Proto-Samoyed. This distribution would make the Uralic word one of the most widely-spread Indo-Iranian loanwords in Uralic.

Both the etymology itself and the idea of its Indo-Iranian character are problematic. First of all, if the etymology was corrected, there are actually no phonological, derivational or semantic characteristics that would make the word Indo-Iranian. The examples of the substitution of “Pre-II” \*ǵ by Uralic \*j are not very convincing (cf. the entry \*jäyi-). The reconstruction reflects simply the PIE stage. But also the PIE reconstruction of the verb has been criticized; Aikio (2002) notes that the reconstruction of a reduplicated stem \*ǵ<sup>h</sup>ug<sup>h</sup>ew- is erroneous, and \*i should be the reduplication vowel here.

Furthermore, the reconstruction of the vocalism of the Uralic word is uncertain, and contradictory reconstructions can be found in the recent works of Aikio (2015b) and Zhivlov (2014). The uncertainty of the reconstructed proto-form naturally hinders any effort to find a phonologically suitable origin for the word. Zhivlov is probably on the right track in assuming that the PU vowel in the initial syllable was \*i, but the development of the word in at least Finnic and Saami is not yet convincingly known. According to Aikio, some languages point to \*u, some to \*i. None of the Uralic words point to the \*u which is reconstructed by the UEW. Mari *jüa*- regularly points to earlier \*i; the UEW and Bereczki (2013 s.v.) assume that the front vowel in Mari is due to influence from \*j-, but this is an ad hoc explanation (most of the cases where the UEW reconstructs a back vowel and Mari shows *ü* can be regularly explained through the development \*i > Mari *ü*). Reconstructing \*i for Proto-Uralic would not solve the problems with the Indo-European original, as there are no known cases of the substitution \*i ← \*e.

Kümmel has considered the Uralic word possibly related to PIE \*y<sup>u</sup>H-s- ‘Suppe, Brühe’ (Kümmel 2009) or \*Hy<sup>u</sup>H- ‘to eat’ (Kümmel 2015; for a more detailed discussion of the Indo-European word, see Nikolaev 2014). Neither option is semantically very convincing, and as long as the vowel reconstruction of the Uralic word remains unclear, so will all the ideas of its external relations remain uncertain. Even less convincing is the Indo-Uralic comparison with PIE \*h<sub>1</sub>eg<sup>wh</sup>- ‘to drink’ (Kassian, Zhivlov & Starostin 2015: 320, 328; the IE root attested in Hittite *ekuzi* and Tocharian A B *yok*, with remnants in derivations elsewhere, see Weiss 2010: 373). Their reconstructions which omit the j- of the Uralic word and the laryngeal of the Indo-European root have been criticized by Kallio (2015a: 370).

Not an Indo-Iranian loanword

**Fi *jää*- 'to stay, remain' (cognates in all Finnic languages)**  
**< Pre-Fi < ? \**jäyi*-**

← ? Pre-II \**g<sup>h</sup>eg<sup>h</sup>oh<sub>1</sub>-* > OI (Vedic) *jāhā-ti* 'leaves, rejects, wins', YAv *zazāiti-lābt zurüick; gewinnt'*, Oss I *zajyn*, D *zajun* 'to stay', PII root \**ǵ<sup>h</sup>ah-* (EWAia II s.v. *hā*; Werba 1997: 331, No. 374 s.v. *hā*; LIV<sup>2</sup> s.v. \**ǵ<sup>h</sup>eh<sub>1</sub>*)

(Koivulehto 1999a: 216–217 = 2016: 218–219)

This is one of Koivulehto's examples of verbs borrowed from Indo-Iranian. Although it is certainly possible to borrow verbs, most of these examples are unconvincing. This etymology is unconvincing due to several phonological problems. The vowel substitution is unlikely: Koivulehto argued that since there was no word-initial \**je-* at the stage where this word was acquired, \**jä-* would have been a natural substitution. However, now we know better that there are at least some instances of Proto-Uralic \**je-* (Aikio 2015a: 8–10).

Also the consonantism is problematic: in the word-initial position the Pre-Indo-Iranian palatovelar is substituted by \**j*, but word-medially it is reflected by the velar spirant \**ɣ*. There is no substance in postulating these two different substitutions, as these developments lack (good) parallel examples. The substitution \**ǵ<sup>h</sup>* > \**j* lacks convincing parallels. Although Koivulehto has presented several etymologies where Indo-European words with retained palatovelars were borrowed into Uralic, it seems that the only convincing case is the well-known PU \**aja-* 'to drive' (> Fi *ajaa* etc.), usually considered a loan from PIE \**h<sub>2</sub>ag<sup>h</sup>-o-* 'to drive'.

Considering these observations, there is good reason to reject the etymology. One has to note also that there has never been particularly a strong reason to consider this word an *Indo-Iranian borrowing*; even if Koivulehto was right, the word could be borrowed from Proto-Indo-European already, as the origin postulated by Koivulehto shows no Indo-Iranian development. Actually, the only Indo-Iranism here seems to be Grassmann's law: in Koivulehto's original, the reduplicated syllable has lost its aspiration, as famously happens in both Old Indo-Aryan and Greek. However, we know that this sound law is a late, separate development in both branches, and there is no reason to postulate this for Proto-Indo-Iranian or an early stage such as Pre-Indo-Iranian! But this would not affect the loan etymology anyway, as one cannot find differences for the substitution of \**ǵ<sup>h</sup>* and \**ǵ<sup>h</sup>*.

Semantically the etymology is also problematic. Koivulehto argues that the meaning of the Finnic word is close to the one attested in Ossetic. However, the original meaning of the word seems to be 'to leave, reject', at least based on the Old Indic and Avestan attestations. LIV<sup>2</sup> notes that the meanings referring to 'winning' have developed from 'to leave someone behind in a race'.

As a side note it is important to point out that many of Koivulehto's etymologies which allegedly reflect an archaic Indo-European form with retained palatovelars can usually be explained otherwise. Some of them are simply wrong etymologies (such as this one), often because of sound laws that have been discovered after Koivulehto's time. However, many others are good etymologies but can be explained as loans from some satem language, i.e. Indo-Iranian or Balto-Slavic.

Not an Indo-Iranian etymology

Ko *kundi-* ‘eingraben, vergraben, verscharren; aufschütten, umschütten’, OP *kundi-* ‘bestreuen’ < PP \**kon-*; Hu *hány* ‘werfen, schleudern’; Ms North *χün-* ‘Schöpfen’, East *χün-*, South *kōn-*, West *kūn-*, (< PMs \**kūn-*); Kh North *χan-* ‘aus dem Kessel schöpfen’, East *kīn-* ‘graben; aushöhlen, schaufeln, schippen (Schnee)’, South *χen* ‘graben; zernagen (das Wasser das Ufer)’ (< PKh \**kān-* ~ \**kīn-*)

< PU \**kana-* (UEW s.v. *kana-*; Sammallahti 1988: 545; Zhivlov 2014: 119)

← PII \**khanH-* (\**kHanH-*) > OI *khan<sup>i</sup>-* ‘graben’, *khánāmi* ‘grabe aus’, OP *kantaniy-* ‘zu graben’ (EWAia I: 445–7 s.v. *KHAN*; Cheung 2007: 229–230, 232–233)

(Munkácsi 1901: 384–388; Korenchy 1972: 73–75; Joki 1973: 275; MSzFE II: 264; UEW s.v. \**kanV-*; Katz 1985: 312; 2003: 249–250)

Munkácsi’s etymology is accepted by Korenchy, Joki and Katz, but it has not been commented in the more recent works by Koivulehto. Joki considers the Permic and Ugric words parallel borrowings from Iranian but his reasons for this assumption remain unclear, as he presents no phonological (or other) criteria. Zhivlov reconstructs the stem vowel of the Uralic word as \**a<sub>1</sub>*, because in his reconstruction Proto-Khanty \**ā* and Hungarian *á* reflect PU \**a* only in \**a–a<sub>1</sub>* stems (Zhivlov assumes that the Khanty forms with \**ĩ* are ablaut variants). The etymology is not mentioned by Aikio (2015b). One can note that it seems that there is still work to do with the exact reconstruction of this PU word, but it does not affect the Indo-Iranian etymology which is quite plausible. The Permic word points to Uralic \**a*. In any case it seems that the PU vowel in the initial syllable is rather \**a* than \**ĩ*, contrary to many other borrowings where it often is difficult to decide which phoneme was the substitution of PII short \**a* (or where the evidence rather points to PU \**ĩ*).

Mayrhofer is uncertain of the Indo-European background of the Indo-Iranian original, citing a possible Phrygian cognate *keneman* ‘probably... a monument or a part of it’. The possible proto-form would be \**kenh<sub>2-</sub>*. Mayrhofer also comments on the problems of Indo-Aryan \**kh-* and Iranian \**k-* (one would expect \**x-* in Iranian), but this need not concern us here, as the word is in any case attested in both main branches of the Indo-Iranian family, and the laryngeal/aspiration would not influence the Uralic sound substitution much.

According to Cheung, the Iranian verbal roots \**kan-* ‘to throw’, \**kanH-* ‘to dig’ and \**kān-* have become mixed through folk etymology in several Iranian languages. The divergent meanings that the UEW reconstructs for the Uralic word (‘streuen, schütten, werfen; graben’) could point to the conclusion that the words in fact represent parallel borrowings from different Iranian verbs. At least the Hungarian word can rather be considered a borrowing from a representative of \**kan-* ‘to throw’ on semantic grounds. Phonologically this would also be completely plausible. However, since the word has palatal *ny-* in Hungarian and retroflex *ŋ* in Khanty, it seems that the word was inherited into both branches from Proto-Ugric, where PU \**n* developed (in certain environments) to \**ŋ*, which is then reflected by Hu *ny* and Khanty *ŋ* (according to Zhivlov 2016a).

The Permic word with a *-nd-* cluster could perhaps be considered a later borrowing from an Iranian word with *-nt-* or *-nd-*, such as Old Persian *kantanaiy-*, Pashto *kan-/kand/*, Oss. *Ikænd* ‘building’, but the borrowing should be quite late because no denasalization has happened in Permic. However, *-dĭ-* is also a verbal suffix in Permic, meaning that the form *kundĭ-* could be a derivative, even though it is unclear what the underived form is here.

An Indo-Iranian etymology

**Md E *kara-*** ‘graben; pflügen’; **Ma *kare-*, *kore-*** (< PMa *\*kâra-*); **? Ud *kĭrĭ-*, Ko *kĭr-*** ‘to dig; to wash’ (< PP ? *\*kĭrĭ-*); **Khanty North *χĭr-*** (< PKh *\*kĭr*)

< PU *\*kara-* (UEW s.v. *kara-*; Aikio 2015b: 55; Zhivlov 2014: 128)

← PII or PI *\*karH-* ‘to spread out, scatter (esp. seed), sow; to make furrows, plough’, YAv *kar-*, *kāraĭieiti*, Parthian *k’r-i* ‘to sow, plant’ Khotanese *kār-* ‘to plant’, Sogdian *kyr* ‘to plant, sow’, Chwarezmian *k’ry* ‘to till, sow, plant’, Oss *Ikælyn/kald*, *D kælun/kald* ‘to flow; to be pulled down; to stumble; to pour, spill; to throw on the ground’; OI *karĭ* ‘to strew, scatter’ (Cheung 2007: 239–240 s.v. *\*karH<sup>2</sup>*; EWAia I: 311, s.v. *KAR<sup>1</sup>*)

(Joki 1973: 266; Rédei 1986c: 51–52)

This old Uralic cognate set is not mentioned by Sammallahti (1988), but it has been rehabilitated in recent works on Uralic vocalism. UEW lists the Permic and Khanty cognates tentatively under two different proto-forms, *\*kurV* and *\*kurV* (? *\*karV*), both with the meaning of ‘dig’, but the matter has been clarified by Aikio and Zhivlov who only reconstruct one verb *\*kara-* ‘dig’. Aikio (2015b) reconstructs the PU stem as *\*kara-* but notes that the vowels in the Permic words are unclear (the Permic words are not listed by Zhivlov 2014). Zhivlov reconstructs the stem vowel as *\*a<sub>2</sub>*. Joki considers the Khanty cognate as uncertain, but Zhivlov connects the Khanty word here and reconstructs Proto-Khanty *\*kĭr-*. Rédei argued that the word is ‘Finno-Volgaic’ and that the Permic and Khanty words continue a different proto-form, *\*kurV* ‘to dig’. Rédei enigmatically mentions that the Iranian word was however not borrowed into Finno-Volgaic but already into Finno-Permic. Because of the uncertain vocalism, the Uralic words could also be parallel borrowings from Iranian.

It is possible that the Permic word is a parallel borrowing from the Iranian verb. However, it has also been suggested recently by Metsäranta (2017) that the Permic word is derived from a PU stem *\*kuri* ‘a furrow scoured out by water’. This is a more probable solution than assuming an Iranian origin for the Permic verb, as there are no known cases of Iranian *\*a* being substituted by Proto-Proto-Permic *\*ĭ* and it is also a semantically plausible explanation.

Mari *kurala-* ‘to plough’ has been connected to the Permic and Khanty words by UEW, but it is not mentioned by Zhivlov or Aikio; Bereczki (2013 s.v. *korem*) argues that the Mari word is not the cognate of the other words listed by UEW (Bereczki supports UEW’s reconstruction of *\*kurV*, see above). It is obviously not their cognate, but it can well be a

parallel loan from Iranian. Bereczki (2013 s.v. *kuralam*) further argues that the Mari word is borrowed from these Permic forms. This can be correct, but there are no semantic or phonological difficulties that would prevent us from considering the Permic and Mari words as cognates.<sup>20</sup>

I see no reason to doubt the Indo-Iranian etymology of Uralic *\*kara-* in the form it is reconstructed by Zhivlov, as phonetically and semantically the etymology is convincing. Here the Iranian forms are semantically closer to the Uralic words: the Old Indic verb denotes ‘to strew, scatter’. The Ossetic words are not phonetically or semantically near the Uralic words, so even if the Uralic words are parallel borrowings, the words cannot be borrowed from Alanic.

An Indo-Iranian etymology

**Ko *kureg*, *kureg* ‘chicken, hen’; Ud *kureg*, *kurek* id. < PP *\*kureg***

**< Pre-P *\*karka***

**← I *\*karka-*, Av *kahrka-* ‘calling word for chicken; chicken’, Oss *kark* ‘chicken, hen’, NP *karg*; Pashto *čirg* (< *\*karkya-*) (Cheung 2002: 196–197)**

(Munkácsi 1901: 300–302; KESK: 147; Joki 1973: 276; Gamkrelidze & Ivanov 1984: 922; Rédei 1986c: 71–72; Lushnikova 1990: 216)

This word is probably a relatively old borrowing, acquired into Pre-Permic already, as the Permic *u* vocalism cannot be explained on the basis of the later Iranian words. In many later borrowings Iranian *a* (or *æ*) is reflected by Permic *a*, but there are several loans confined to Permic that show *u* which has developed from Pre-Permic *\*a* or *\*i* (see *\*ačwa*). It is difficult to say anything more precise about the dating or the donor form of the word: this can be an old borrowing from an early form of Alanic, but there are no compelling reasons to assume so, purely formally the word could be borrowed even from Old Iranian. Rédei assumes that the borrowing is from Old Iranian but this seems more like guesswork, as there are no phonological means to date the Iranian donor form to some particular stage. As the Iranian word has a reflex in Ossetic *kark*, the Permic word could equally well be an Alanic borrowing.

Various developments of the cluster *-rk* can be found in loans: Komi *verk* seems to point to a Pre-Permic geminate cluster *\*-rkk-* (see the entry *\*wärkä*), but here *-rk* has developed into *rVg* with an epenthetic vowel and the voicing of the stop. In the inherited Uralic vocabulary, *\*rk* > *\*r* is the regular development. Voiceless *k* can be found in some Udmurt forms, and Rédei assumes that these reflect the original state of affairs, and the voicing is due to analogy from words that end in *Vg*. As the relations between the Udmurt and Komi words are otherwise regular, it is clear that the forms with word-final *-g* and *-k* cannot reflect parallel borrowings from different Iranian words with *-rg* and *-rk*. Interestingly, also the word-final *k* in Ossetic *kark* (instead of *karg*) is irregular according

---

<sup>20</sup> I am grateful to Niklas Metsäranta for pointing this out to me.

to Cheung (2002: 196–197), but this phenomenon is probably unrelated to the problem of the word-final consonant of the Permic word.

An Indo-Iranian etymology

**\*karši**

See **\*kärsä**

? Indo-Iranian etymology

**Fi karhu ‘bear’, Est karu id. (cognates in Izh, Ka, SEst, Vo)**

**< ? Pre-Fi \*karšV**

← **Pre-II/PII \*Hr̥kša-s (? \*Hr̥tša-s) > OI r̥kša- ‘bear’, YAv arša- id., Khot arra id., Chwarezmian hrs, MP xirs, NP xir, Kurd hirč id.; < PIE \*h<sub>2</sub>rtkó-s** (EWAia I s.v. r̥kša-; Kümmel 2014: 2; Lipp 2009)

(Katz 2003: 114; Delamarre 1992: 151–154; 2003: 56; EES 130 s.v. *karu*)

This rather recent etymology is not found in most standard etymological works. EES supports it but offers no further arguments for its credibility. Katz’s and Delamarre’s etymology for the Finnic word is complicated. In short, one could say that it is too good to be true, as the Indo-Iranian origin involves complicated clusters and its development from Proto-Indo-European is also not straightforward, and thus it is too easy to explain the Finnic form through simplifying substitutions of the Indo-Iranian form. While this etymology cannot be totally ruled out, there is simply not enough argumentation to support it.

The reflex of the Indo-Iranian laryngeal is one problem. While it is not phonetically implausible for the Indo-Iranian laryngeal to be substituted by \*k, there is actually not a single parallel example for this among the more transparent and convincing Indo-Iranian loan etymologies. Although there are various, more or less convincing examples of the Indo-European laryngeals being substituted by \*k in early loanwords (Koivulehto 1991; Hyllested 2014: 11–12), none of them happens to reflect an Indo-Iranian form.<sup>21</sup> Thus this substitution remains an interesting possibility, but something that is not yet proven convincingly.

---

<sup>21</sup> The lack of laryngeal reflexes in Indo-Iranian loans (opposed to other IE loans) can point to a “weak” realization of this phoneme in Proto-Indo-Iranian, as has been argued by Koivulehto (1991). Kallio (2008) says the same about the Proto-Balto-Slavic reflex of the laryngeals, as this also lacks any substitution in the loans. Kümmel (2014) has argued that the laryngeal was retained in some Iranian languages even until modern times, making this assumption of the “weakness” of the laryngeal unlikely. It is also clear due to the laryngeal hiatus in Gathic and Vedic hymns that the laryngeal was retained in Indo-Iranian until shortly before the compilation of these oldest Old Iranian and Indo-Aryan texts, but it might have had different realizations in different positions.

The shape of the Finnic word (with non-initial syllable *-u*) does not point to an old word, as the labial vowels in non-initial syllables are usually secondary and there are no convincing parallel examples (Holopainen, Junttila & Kuokkala 2017: 133–134).

Blažek (2017a: 177, 178, 180) considers Delamarre’s original etymology problematic (Blažek finds the substitution of laryngeal by *k* unconvincing) and offers a modified version of it. According to Blažek, the Finnic word is borrowed from Proto-Indo-Aryan *\*karkṣī-* ‘she-bear’ (< *\*kartṣī-* < *\*kartśī-*). The word *\*karkṣī-* is not attested in Old Indic, but according to Blažek it can be reconstructed on the basis of later Indo-Aryan forms, Bashkaryk *kāch* and Torwali *kēṣ* id.

Also Blažek’s etymology has to be seriously doubted. There is no proof that the ‘she-bear’ form with the prefix *\*ka-* existed already Proto-Indo-Aryan or in Proto-Indo-Iranian, and it is impossible to prove that the word existed in a language which was in contact with (Pre-)Finnic. Phonologically the word is less problematic, as Finnic *h* (< *\*š*) could easily be explained from a “thorn cluster”, regardless of whether that was *\*kš* or *\*tš* in Proto-Indo-Iranian (see Lipp 2009).

According to Saarikivi (personal communication), the Finnic word might be related to North Saami *guoržžu* ‘monster’. In this case Finnic *h* should reflect earlier *\*č* (this might be possible in a consonant cluster). The relationship between the Finnic and Saami words remains open, but it does not have much relevance for the Indo-Iranian loan etymology.

? Indo-Iranian etymology

**Hu *keszeg*, *kesze*, *keszi* ‘Weissfisch, *Leuciscus abramis*’; Ma North *kāsew*, East *kṣṣəŋ*, South *kāseŋ*, West *kāṣəŋ* ‘Plötze, Barbe, *Leuciscus rutilus*’ (PMA *\*kāsəŋ*); Kh South *kōsə* ‘Plätze, Barbe, kleiner junger Fisch’ (< ? PKh *\*kūsə*)**

< ? PUg *\*kasiw* (Sammallahti 1988: 516)

← ? Iranian, cf. Ossetic *kasag* ‘Fisch’ < PI (?) *\*kasāka-*; OI *kaśaku-* ‘Coix barbata’ (EWAia III: 78)

(Joki 1973a: 270; Katz 2003: 117, footnote 137)

This is not a very convincing etymology. Although the etymology is commented on by Joki (1973), it is not even mentioned by Korenchy (1972). According to Mayrhofer (EWAia III: 78), *kaśaku-* is not a proper Indic word at all but a non-Indo-Aryan gloss in a Sanskrit text. Thus there are no reasons to suppose a common proto-form for the alleged Old Indic word and Ossetic *kasag*. The etymology of the latter word remains unclear.

The latter assumption is made unlikely by the fact that the Ugric words are not regular cognates because of the irregular vowel correspondences. Even though Sammallahti (1988) reconstructs the word to Proto-Ugric, the assumed reflexes in Hungarian, Khanty and Mansi do not point to *\*a*, and it is impossible to derive these forms from any other unitary reconstruction.

The exact connection of the Ugric words and the Ossetic word thus remains unclear; the word can be borrowed separately from Alanic to the different Ugric languages (if we suppose that the Ossetic form goes back to such early times) or the Ossetic and Ugric forms

can be parallels from same, third source. But the word is certainly not an early Indo-Iranian borrowing to Uralic.

Not an Indo-Iranian etymology

**SaN *guohtut* ‘to graze’ (cognates in all western Sa languages and SaI (< PSa \**kuotō-*) (Álgu s.v. *guohtut*)**

< Pre-Sa \**kata-*, \**kato-*, \**kīta-*, \**kīto-*

← PII \**k(h)Hād-* or PI \**xād-*, OI \**khād-* ‘(zer)kauen, (fr)essen’, perfect *cakhāda*, YAv *vī-χaδa-* ‘auseinanderquetschen’ (EWAia s.v. *khād*; Werba 457, No. 624 s.v. *khād*)

(Koivulehto 2007: 254; Holopainen 2018b: 159–160)

This relatively recent etymology is one of the Indo-Iranian loans that are found only in Saami. For the now obsolete, phonologically irregular comparison to Samoyed forms like TN *χado* (UEW s.v. \**katz*; FUV: 35), see Helinski (1996: 66) and Holopainen (2018b: 159–160). Koivulehto’s etymology is convincing: semantically the Saami and Indo-Iranian words are very close, and phonetically the substitution is expected. This term fits well with other terminology connected to herding and pastoralism borrowed into Uralic languages from Indo-Iranian. The exact donor language (PII or PI) cannot be determined: the Saami word could also be from Proto-Iranian \**xād-* with \**k* as the substitute of \**x*. The word is not attested in Ossetic, so it is probable that this is not an Alanic borrowing. Theoretically the Saami word could be borrowed from Proto-Indo-European or some other archaic branch of Indo-Iranian, as \**k(h)Hād-* possibly reflects PIE \**kh<sub>2</sub>a(h<sub>1</sub>)d-* with a cognate in Armenian *xacane-* ‘bites’ (EWAia s.v.; Mayrhofer 2005: 114; Holopainen 2018b: 159–160). As the word is not attested in other branches of the family that were in contact with Saami (Germanic or Balto-Slavic), the Indo-Iranian origin remains the most convincing possibility.

The vowel substitution is probably \**a* here, although \**i* cannot be ruled out. If the Saami word is an original \**a*-stem, the initial syllable can point to either \**a* or \**k(h)Hād-*. The Proto-Saami \**-o* in this word is of unclear origin.

An Indo-Iranian etymology

**Fi *kehrä* ‘spindle’, (der.) *kehrätä* ‘to spin’, Ka *kesrä*, *kezrä*, Lu, Ve *kezr*, Est *keder* (: *kedra*), Li *kied’der* (< PFi \**kecrä*); SaN *gearsi*, IpR *kärsi* ‘spindle’, Lu *kier’sē*, *kär’sē* ‘wheel of the spinning wheel’ (< PSa \**kearsē*, Lehtiranta 2001: 48–49); Md E *šteré*, *štšere*, M *kštir* ‘spindle’, E *šterd’e-*, M *kštard’e-* ‘to spin’**

< PU \**kečrā* ‘spindle’ (UEW s.v. *kečrā*; SSA s.v. *kehrä*)

← Pre-II *\*ketstro-* > PII *\*čatstra-m* > OI *cātra-m, cātra-m, cattra-m* ‘spindle’, root *kr̥t* ‘to spin’, Pashto *cāṣai* id. (PI *\*častra-*), < PII root *kart* (: *čart*) (EWAia I: 539, s.v. *cātra-*; LIV<sup>2</sup> s.v. ; Werba 1997: 170–171, No. 30 s.v. *kr̥t*) (Koivulehto 1979: 71–78; Katz 1985: 348, 2003: 277; Rédei 1986c; Parpola 1999: 194–195; 2017: 268–270; Kallio 2012b: 231, footnote 9)

Koivulehto’s etymology is convincing, and there is no doubt that this Western Uralic word is an Indo-European borrowing. However, determine the exact donor language is open to different arguments, as the word-medial consonantism of the Uralic word is difficult to reconstruct: the Finnic words point to a cluster of affricate and *\*r*, but the reconstruction of these kinds of clusters is difficult, as the examples are scant and the development of clusters is divergent in different Finnic languages. Kallio (2012b: 231, footnote 9) argues that *\*kečrā* is the most plausible reconstruction, and that the Mordvin words which presume a pre-form *\*keštrā* can be explained through a metathesis of *\*č* to *št*. The disappearance of the first-syllable vowel in the Mordvin words and the subsequent consonant-assimilation in Erzya can be explained through vowel reduction which was caused by the stress in the final syllable. Paasonen (1922: 66) notes that August Ahlqvist in the 19th century still wrote down a disyllabic Moksha form *kištir*, which means that at least in Moksha the loss of the vowel was quite recent.

Mari *šüδar, šüδür* ‘spindle’ (< PMA *\*šüδar*) has been connected to the Finnic, Saami and Mordvin words in early research (see SSA with references), but it is not considered as the cognate of these words in SSA. Rédei (1986c) still lists the Mari word as an uncertain cognate. Parpola (2017) has tentatively suggested that the Mari word could be borrowed from Mordvin. Although this is not impossible, the lack of generally acknowledged borrowings from Mordvin to Mari makes this assumption highly uncertain, and more evidence would be needed before the Mordvin origin of Mari *\*šüδar* could be accepted. It is in any case impossible to derive the Mari word regularly from *\*kečrā*. The Mari word could perhaps be a later borrowing from some reflex of Iranian *\*častra-*. The *š* could regularly reflect PU *\*č*. However, it is difficult to explain the Mari vocalism and the consonant *δ*, so the Indo-Iranian origin of the word remains purely hypothetical.

Both Koivulehto and Parpola have considered this word a Pre-Iranian borrowing, as the Iranian development of *\*tstr* would explain the Uralic form well in their view. Koivulehto (1983) considered the word an early Indo-European borrowing, but later (Koivulehto 1999) argued that the Uralic word is borrowed specifically from Pre-Iranian *\*kesro-*. It is certainly a better candidate than an Indo-Aryan form, but there is no reason not to derive the Uralic word from Pre-Indo-Iranian already (the same layer as *\*meti*, *\*mekši* and a number of other convincing early borrowings), as *\*-tstr-* could probably result in a cluster *\*-čr-*. A “Pre-Iranian” form is also problematic because of the retained Indo-European vowel *\*e*: Koivulehto and Parpola have suggested a number of such etymologies, but practically none of the examples is compelling, as the etymologies can usually be explained otherwise.

Parpola notes that phonetically *\*ketstro-* would be an equally good source as Pre-Iranian *\*kestro-*, but according to him the idea of an early Indo-Iranian borrowing would not fit the archaeological context very well. Because linguistic arguments should be given priority in the study of loanwords, I see it as more likely that the word is indeed Pre-Indo-

Iranian (with retained \*k and \*e vocalism) than an Iranian form showing the development \*tstr > sr.

Actually, there is even no problem to consider this word a borrowing from Proto-Indo-European, as here the Pre-Indo-Iranian word is practically identical to its Proto-Indo-European predecessor. The Indo-European instrumental suffix \*-tlo has two allomorphs, \*-tlo and \*-tro, and as this formation is attested only in Indo-Iranian, we do not really know whether the word reflects an Indo-European shape \*ket-tlo- or \*ket-tro- (the environments for the different allomorphs are unknown, see Olsen 1988). Of course, there is no certainty that this word goes back to Proto-Indo-European, but both the root and the derivational suffix are inherited from Proto-Indo-European, so postulating a PIE form is unproblematic. In any case the borrowing is much earlier than the later Iranian (Proto-Iranian or Alanic) loanwords which are shared by the western Uralic languages.

An Indo-Iranian etymology

**Fi *kekri*, *köyri* ‘pagan new year feast’ (only in Finnish and Karelian) (SSA s.v. *kekri*; SMS s.v. *kekri*); SaN *geavri* ‘circular thing’ (cognates found in all Saami languages except U and L, < PSa \**keakrē*) < PWU \**kekrä***

← Pre-II \**kekro-* (< PIE \**k<sup>w</sup>ék<sup>w</sup>l-o-s* (m.), \**k<sup>w</sup>(e)k<sup>w</sup>l-éh<sub>2</sub>-* (n.), root \**k<sup>w</sup>el-* >) PII root \**čarH*, noun \**čakrá-m* > OI *cakrá-m* ‘wheel’ (EWAia I: 521–522; LIV<sup>2</sup>: 386–388; Parpola 2007: 4–5; Werba 1997: 286–287, No. 284)

(Katz 1985: 110; 2003: 95; Koivulehto 2000b: 241–250; 2001b: 249 = 2016: 239–248, 291; Parpola 2015: 63; Holopainen 2018b: 148–149)

North Saami *geavri* ‘a circular thing’ can be directly derived from a Pre-Saami form \**kekrä*. The Finnic and Saami words can be regular cognates, if the Finnic *i* is the result of a derivational process as Koivulehto argues (the *i* here cannot be old, as this is an unaltering *i*-stem). The word has to be a very early borrowing, as it precedes the second palatalization (PIE \**k* > PII \**č* before front vowels). It has been assumed that the Uralic \**e* vowel points to the delabialization of the \**k<sup>w</sup>*, an Indo-Iranian development, shared with other satem branches. An even stronger indicator of an Indo-Iranian origin is the development \**l* > \**r*; this etymology shows that the development has to be early, and most of the loanwords support this idea. The loans with retained \**l* are probably not true Indo-Iranian borrowings (see the entry \**śali*).

It is, of course, suspicious for a very early (Pre-Indo-Iranian) borrowing to have so limited distribution. The word has to belong to the same layer of words as such as borrowings as \**mekši* and \**meti*. It is thus suspicious not to find cognates in other branches of the family. Finnic has borrowed various words for ‘wheel’ from different Indo-European languages (Parpola 2010: 9–10), and as these words are easily borrowed, a cognate could have been lost elsewhere in the family. Parpola mentions the word \**kekrä* as a borrowing into Proto-West Uralic. This can be the case. As many of the Proto-Indo-Iranian loanwords show various irregularities that can point to parallel borrowings, it is possible that Proto-Uralic was diverging at the time of the earliest Indo-Iranian contacts

already. Most of the Pre-Indo-Iranian \*e-loans have a wider distribution, but also the well-known example \*kečrā ‘spindle’ has a West Uralic (Finno-Saami-Mordvin) distribution only.

SaN *geavli*, which would reflect a pre-form \*keklä, is an independent borrowing from the same Indo-European word according to Koivulehto (2000a: 248). Koivulehto considers this an early Germanic loan (the Proto-Germanic cognate is \*h<sup>w</sup>ehla-, Kroonen 2013 s.v.), but Kümmel (2018a) has argued that this could be an Indo-Iranian loan too. Due to the \*l the latter option seems unlikely (it is shown in this work that no convincing cases of retained Proto-Indo-Iranian \*l can be found among the loanwords). One could try to argue that the vowel \*e shows that the velar of the donor form was not labiovelar, which could point to a satem language; among Germanic loanwords there are examples where labial consonant influences the vowel quality in substitution, but there are very few examples of this substitution among earlier Indo-European loans. The exact origin of *geavli* remains open, but it is possible that it is borrowed from some branch of Indo-European.

An Indo-Iranian etymology

**SaI korttađ ‘to tie’ < PSa \*kerte-; SaI korttas < PSa \*kerttes ‘leather strap’ (cognates in all Saami languages except in Ter Saami; Aikio 2012b: 88, footnote 12)**

**< Pre-Sa \*kert(t)ü**

← PII \*krt- ← Pre-II \*kert- > OI crtáti ‘he/she ties’, < PIE \*k(w)ert- (EWAia I: 536–7, s.v. CART; LIV<sup>2</sup> 356–57, s.v. \*k(u)ert-)

(Koivulehto 1988a: 41–43; 2001a: 249; Holopainen 2018b: 160–161)

The Saami verb \*kerte- is reflected only by Inari Saami *kortta-*, but Aikio (2012b) has argued that the verb serves as the basis of the derived noun \*kerttes ‘leather strap’ which has a wide distribution in all Saami languages except in Ter Saami. Sammallahti (2001: 407) considered the Saami word as an early Indo-European loan, but not specifically Indo-Iranian. The Indo-Iranian origin of the Saami word is in principle convincing, but as the IE root is attested also in Germanic, the word must have been present in some Indo-European proto-language, and it remains unclear whether the word really is borrowed from *Indo-Iranian* to (Pre-)Saami. Koivulehto has noted that Germanic *\*hurpi-* > Gothic *haúrds* ‘door’, Old High German *hurd* ‘Flechtwerk’ also reflect PIE \*kwr̥ti-, and the word can thus have had wider distribution originally. It is difficult to solve the question of the exact origin of the Saami word on phonological or semantic criteria, as in principle the word could be either an early loan from Germanic, an Indo-Iranian loan or also a borrowing from some other archaic branch of Indo-European, although the latter option is certainly the least convincing.

This is one of the Saami words which point to Pre-Saami \*rtt (see also \*ertü, \*kertü and \*porta). As the Indo-Iranian/Indo-European origin had single *t* in all the cases, Saami (as well as Permic) seems to have secondary germination of *t* in *rt* clusters, a has

been noted by Kallio (2012b: footnote 10). The vowel substitution here is ambiguous, as Pre-Saami \**e* here could point to \**e* in the donor language, but it could also reflect the syllabic *r*.

The Permic verb ‘to tie’ (PP \**kerti-* > Komi *keřt-* and Udmurt *kertti-* ‘to bind’) is derived from PU \**käri-* ‘wrap, tie’ (KESK: 142; UEW s.v., Aikio 2002: 18) and thus do not belong here. Koivulehto has assumed that the Permic words are parallel loans from the same Indo-European source as the Saami word. As the Permic vowels can regularly reflect PU \**ä*, it looks more likely to derive them from this Uralic word rather than support Koivulehto’s loan etymology for them.<sup>22</sup>

An Indo-Iranian etymology

**Fi *kanto* ‘tree stump’ (cognates in all Finnic languages); SaN *guottu* id.; Md E *kando* ‘windthrow’; Ms West *kēnt*, North *χānta* ‘storehouse pillar’ (< PMS \**kīntā*); Kh North *χont*, East *kant*, West *χont* ‘tree stump’ (< PKh \**kānt*)**

< PU \**kīntaw* (UEW s.v. *kanta*; Sammallahti 1988: 543; Aikio 2015b: 59)

← PII \**skandhá-* > OI (Vedic) *skandhá-* ‘Schulterknochen’, \**skándhas-* > OI (Vedic) *skándhas-* ‘Baumast, Zweig’, root \**skandh-* < ? PIE \**skendh-* (EWAia II s.v. *skandha-*)

(Katz 2003: 243–245; Pystynen 2017; Aikio ms.)

This is a convincing etymology. Katz’s etymology has been supported by Pystynen (2017) and Aikio (ms.). The Uralic word’s meaning fits well with the meaning ‘tree trunk’ that is attested for the Vedic *s*-stem. The meaning attested in Mansi has to be secondary.

Here the Uralic vowel clearly points to the substitution of PII \**a* by Uralic *i*, but it is difficult to determine whether this substitution is due to unaccented \**a*. The Indo-Iranian \**-as*-stem \**skándhas-* had an accented *a* in the nominative, but the Indo-Iranian word is also attested as an *a*-stem \**skandhá-* with different accentuation, and the Uralic word could also be borrowed from that form.

Katz assumes that the Uralic words represent two parallel borrowings, with one reflected by Mordvin and the Ob-Ugric languages and the other by Saami and Finnic. As all the Uralic forms can be regularly derived from a Proto-Uralic reconstruction \**kīnta*, there is no need to assume parallel loans, and instead the word could have been borrowed into Proto-Uralic already.

An Indo-Iranian etymology

<sup>22</sup> Permic Ko *gered*, Ud *gerd* ‘knot’ are also mentioned by the UEW and Rédei and 2004: 322) but it is uncertain if these belong even to the same word family as the verb *keřt-*, *kertti-* because of the voiced *g-*.

**Md E kšńe, kšńi, M kšńi ‘iron’; Ma kūrtnö, kərtni**

< ? \**kūrtnV*

← PII \**krtí-* > OI *krtí-* ‘dagger, knife’, PII \**kártana-* > OI *kártana-* ‘a cut’ (EWAia I: 390, s.v. *krtí-*)

(Munkácsi 1901: 390–394; Jacobsohn 1922: 203; Joki 1973: 273, No. 62; Katz 1985: 325–327; 2003: 260–261; UEW s.v. *kürtz*; Bereczki 2013 s.v. *kürtnö*)

See also \**kārtV*, \**kīrV*

This is a well-known Indo-Iranian etymology, but the cognate set presented by UEW which includes the Permic words *kert*, *kort* cannot be correct, nor can be the reconstruction that the UEW gives. Bereczki considers only the Mari and Mordvin words as cognates. However, he does not mention the Iranian loan etymology or the similar words in Ugric and Permic. Bereczki notes that the element *-ńV-* in the Mordvin and Mari words is a derivative suffix, more precisely a Uralic denominal suffix \**ńV* (referring to Bereczki 2002: 180–181).

Koivulehto (1990: 90) argues that the *ń* in the Mordvin and Mari forms can be explained from the suffix \**-ana-* (< PIE\**-eno-*) in the donor language: Koivulehto cites the Vedic verbal noun *kártana-* ‘a cut’ as the reflex of the probable source of these words. Koivulehto cites Rédei (1986c) who assumes that the Mordvin and Mari forms are parallel borrowings.

Bereczki (2013) reconstructs the pre-form of the Mari word as \**kūrtnV*. The Mordvin word could probably be derived from the same pre-form. Mordvin *š* is derived from earlier \**r*: in some words *r* was spirantized to facilitate the pronunciation of consonant clusters that had developed when reduced vowels were lost (Keresztes 2011: 54–55). Keresztes mentions \**kürsü* ‘bread’ > Mo E, M *kši* as a parallel. Although this explanation is probably correct, the proto-form \**kürsü* already has a sibilant, unlike \**kūrtnV*; however, \**s* does not regularly develop into *š* in Mordvin, so also here the disappearance or \**r* has clearly played a role in the development of \**š*. It seems that \**ü* developed into Proto-Mordvin reduced vowel and then became lost. Although the exact environments are unknown, there is no obstacle to derive the Mordvin word from \**kūrtnV*.

It is possible to derive both the Mari and Mordvin words from a similar pre-form, and this is further supported by the notion that a similar derivational suffix appears in both words. This means that either the word has been borrowed into a common proto-language of Mari and Mordvin or that the word has been borrowed into pre-Mari and pre-Mordvin at roughly the same time when the two languages still were quite similar. As Mari and Mordvin are not closely related, contrary to some earlier views (Bereczki 1988: 314–315), the borrowing into the common proto-language of the two branches would have to be very early. As the taxonomy of Mari within Uralic is largely unclear (Metsäranta forthcoming), it is very difficult to solve the problem here in more detail. However, a word for ‘iron’ cannot have existed at Proto-Uralic times, so this puts some limits for the dating of the borrowing.

The Permic words *kert*, *kort* ‘iron’ (< PP \**kórt*) that are connected with the Mari and Mordvin words in the UEW have to be separate loans from Iranian. They can reflect earlier

\**ä*, which is a known substitution in Iranian loanwords. However, it is also possible that the word was borrowed into Permic at a time when Uralic \**ä* was already developed into a Proto-Permic reduced vowel that later produced the correspondence set Komi *e, o* - Udmurt *o*, and that the schwa-like Iranian vowel \**a* was substituted by this Permic vowel. However, Zhivlov (2010: 175) still reconstructs this vowel as \**ä*. Csúcs (2005: 347) reconstructs Proto-Permic \**kört*.

Regarding the Iranian original, it is tempting to assume that the original was similar to OI *krtí*: the syllabic \**r*<sub>1</sub> could have been substituted by \**ür* in the Uralic side, but it is much more difficult to assume similar substitution for Iranian full-grade \**ar* that is reflected by the attested Avestan forms. Also the vocalism of the Middle Iranian languages fits poorly with the \**ür* that has to be reconstructed for the pre-form of the Mari and Mordvin forms.

The semantics of the word are problematic, as the Indo-Iranian words denote ‘sword’ or ‘knife’, whereas in the Uralic side ‘iron’ is the primary meaning. In some later Iranian languages, at least in Shughni (*kârč*) the word means ‘steel’; perhaps a similar semantic from ‘sword’ to a ‘metal’ development has occurred in the languages that were spoken in the Eurasian steppe, but this meaning should be quite early as the borrowings in Mari and Mordvin have gone through similar vowel changes as the inherited Uralic vocabulary.

Blažek (2017: 291, footnote 20), who supports the Iranian etymology of the Uralic words, mentions a semantic parallel for the development ‘tool’ > ‘metal’: Middle Persian *pwl’wd* ‘steel’ (from which Armenian *połopat, połovat* ‘steel’), cf. Vedic *pāvīravat-*, *pavīrávat-* ‘armed with a lance or a goad’, derived from *pávi-* ‘metallic point of a spear or an arrow; tire of a wheel’. The development from ‘metal’ to ‘sword’ is also attested among the meanings of German *Eisen* according to *DatSemShift* (s.v. *metal*).

Joki also mentions Mari *kerde*, and Mordvin *keřet* ‘Pflugmesser’ that might be later, clearly independent borrowings from the same Iranian word. This idea is convincing. These etymologies do not involve any semantic problems.

The Iranian word was borrowed into Tocharian as well (Tocharian B *kert(t)e* ‘zwaard’, ‘straal’), but there the more original meaning is retained (Isebaert 1980: 89; Joki 1973: 273).

An Indo-Iranian etymology

**Ko *kert*, Ud *kort* ‘iron’ (< PP \**kört*)**

See \**kürtńV*

An Indo-Iranian etymology

**Mansi North *kēr* ‘iron (Eisen; eiserne Spitze; Schneideeisen; Glocke, Kirchenglocke, Schelle (rauta, rautainen kärki; leikkuurauta ?)’, East *kēr*, West *-kēr* (in compounds), South *k<sup>o</sup>ēr* < PMs *\*kīrV* (Kannisto et al. 2013: 345 s.v. *küer*)**

← ? Iranian, cf. PI *kartá-* > Avestan *karəta-* M ‘knife, dagger’ < PI *\*kartá-*; cf. also Av *kaša-* < *\*kárt-a-*; < PIE *\*kworto-*, root *\*kwer-* ‘to cut’ (EWAia I: 390 s.v. *kr̥tí-*; Isebaert 1980: 89)

(Munkácsi 1901: 390–; Korenchy 1972: 58–59 No. 10; Joki 1973: 273, No. 62; Katz 1985: 325–327; 2003: 260–261; Helimski 1997b: 121)

See also *\*kürtńV*, *\*kārtV*

The Mansi word is obviously not a cognate to Khanty *\*kārtV* but a parallel borrowing, as stated already by Korenchy. Korenchy reconstructs the Proto-Mansi vowel as *\*ī*, which does not fit the vowel of the attested Iranian words very well. Also the absence of *-t-* in the Mansi forms points to the conclusion that, unlike Khanty *\*kārtV* and the Volgaic and Permic words, the Mansi word is not borrowed from a form that is closely related to Avestan *karəta-* ‘Messer’. Korenchy notes that the loss of *-rt-* cannot have happened in Mansi, but the donor form has simply *-r*. Korenchy also mentions that the “long, closed palatal vowel” was already present in the donor language, but she does not attempt to determine the donor form in more detail. Joki mentions that the Uralic words are borrowed separately from various Iranian languages but he does not comment on the phonological problems with the Mansi word in any detail.

Regarding the consonant cluster *\*-rt-*, it is interesting that in the Middle Konda dialect of Mansi, the cluster *\*-rt-* of Komi loanwords was simplified to a voiceless *-R* in the word-final position (Liimola 1971: 89). It would be anachronistic to project a development of one Mansi variety back to a Proto-Mansi level (and the resulting *r* is also different), but perhaps the loss of *t* happened in Mansi because of phonotactic reasons.

Regarding the Proto-Mansi vowel, Korenchy’s reconstruction of *\*kīrV* is probably correct, as the Mansi cognate set has similar vocalism as the word *\*mōŋki* ‘forest spirit’ which is reconstructed with Proto-Mansi *\*ī* by Zhivlov. However, this does not automatically mean that the Iranian vowel was substituted by *\*ī*, as the Mansi vowel can also reflect earlier (Proto-Ob-Ugric?) *\*ē*.

Some Iranian languages have forms that have simplified the *\*rt* cluster. These include Kurmanci Kurdish *kēr* ‘knife’ and Yidgha *keṛo*, Munji *kēṛa* (Cheung 2002: 196.) It is tempting to derive the Mansi word from a form akin to these. Obviously, the Mansi speakers cannot have borrowed the word from Kurdish, and direct contacts to Pamir languages are also an implausible idea. Perhaps the word has spread as a Wanderwort from the south. Another option is that we have to assume that the Iranian languages of the steppe have participated in similar sound changes as the Pamir languages or Kurdish, but it is impossible to prove this. In Chwarezmian a form *krc* is attested that shows affricatization of earlier *\*kartiya-*. Assuming that the Mansi word could be borrowed from such a form and this would explain the simple *\*r*, would not however solve this problem, as clusters of *\*r* and affricate, at least *\*-rč-*, were allowed in Proto-Mansi, cf. *\*wārčək* ‘ein Vogel’ (Zhivlov 2006: 116).

It is also problematic that the Mansi word denotes simply ‘iron’, not ‘knife’, whereas ‘knife’ is the sole meaning of the Iranian words. The same problem is connected to the semantics of the other Uralic words for ‘iron’ (\**kürtńV*, \**kārtV*) as well. Hungarian *kard* ‘sword’ is a transparent late Alanic borrowing that better reflects the semantics of the Iranian word (EWUng s.v. *kard*).

According to Liimola (1957: 191–193), the meanings of *kēr* that refer to ‘edge’ or ‘point’ or the like reflect not the original meaning but are a later development which is parallel to the semantic development of Khanty *wax* ‘iron’ in various compounds. Liimola is aware of the Iranian origin of the Mansi word (he refers to the etymology of *Munkácsi*) and its semantic development. Liimola’s conclusion can be reconsidered; it is not unthinkable that the Mansi word originally referred to various sharp objects that were made of iron, and later ‘iron’ became the primary meaning. This polysemy could then have influenced the semantic developments of the Khanty word for ‘iron’. This could explain the semantics of the Mansi word, although it has to be admitted that the explanation remains speculative.

Katz agrees with Korenchy in that the Mansi word has to be separated from the other Uralic words, but he does not consider it a loan from Iranian \**karta-* but from a hypothetical formative of the same Indo-Iranian root; this form would be \*\**karam* ‘das Ergebnis des Schmiedens, das Geschmiedete’ in standard reconstruction, “\*\**k(u)ärɔm*” in the system of Katz. The explanation of Katz is problematic because of its hypothetical nature, and it is uncertain how well this could explain the Mansi vocalism. The etymology of the Mansi word remains unsolved.

? Indo-Iranian etymology

**Khanty North *karti*, *karte*, (Obdorsk dialect) *kordi* ‘Benennung des Kessels; Eisen’ (< PKh \**kārtV*)**

← Iranian \**karta-* (Korenchy: “MI”). cf. Avestan *karəta-* M ‘knife, dagger’ < PI \**kartá-*; also Av *kaša-* < \**kārt-a-*; < PIE \**kʷorto-*, root \**kʷer-* ‘to cut’ (EWAia I: 392 s.v. *kr̥tí-*; Isebaert 1980: 89)

(Korenchy 1972: 58–59, No. 11; Joki 1973: 273, No. 62; FUV: 131; UEW s.v. *kärtɜ*)

See also \**kürtńV*, \**kīrV*

The Iranian origin of the Khanty word is convincing, and it is obviously a separate loan from both Mansi \**kīrV* as stated already by Korenchy. More difficult, however, is the question of whether the Khanty word is directly borrowed from Iranian, or whether the Khanty words are borrowed from Komi as the UEW assumes.

Korenchy argues that the Khanty word and the other assumed cognates cannot represent an early borrowing at the time of Proto-Uralic because the Khanty word reflects \**ä* and there are no known cases of PU \**ä* reflecting Iranian \**a*. Korenchy even states that the word cannot be old because there was no \**ä* from earlier \**a* in Proto-Iranian. This is circular reasoning, however, because it is phonetically completely possible that PI \**a* would have been substituted by Uralic \**ä*. Korenchy is right that there are no examples of this substitution among the earliest loans with a wide Uralic distribution, but some of the

earliest loans into the Ugric languages actually point to the rather early occurrence of such a substitution (see \**päčäy*, \**säptä*).

The Permic words (*kort*, *kert*) could actually go back to an old \**ä*-word (compare the reflexes of Uralic \**tälwä*, \**jäsi*, Sammallahti [1988: 549, 552]). FUV (131) still reconstructed the word back to Proto-Finno-Ugric, but the UEW considers the Khanty words to be Permic borrowings and the Mansi word possible separate borrowings from Iranian. See \**kürtńV* for the etymologies of the Mari and Mordvin words that have been connected here. The word does not to appear in Sammallahti (1988).

If the word is a direct borrowing from Iranian, it is difficult to determine the age of the borrowing. The Proto-Khanty vowel \**ā* is found in a number of Iranian loans, some of which are quite old (and this vowel can reflect PU \**ä*), so the word can be an early borrowing but it can also be the case that the word was borrowed later into Proto-Khanty. There are no specific Alanisms that would force us to consider the word borrowed from an Ossetic-type language; even Proto-Iranian \**kartá-* would be a possible source, as \**ā* appears as the substitution of Iranian \**a* quite early in the loans that were borrowed into the Ob-Ugric languages. But the word naturally cannot be older than the beginning of the Iron Age.

Toivonen considers only the forms in the Obdorsk dialect with a labial vowel to be borrowed from Komi; the other Khanty forms cannot be loans from Komi because of the vocalism. The other Khanty words, as well Mansi \**kĩrV* are cognates of the Permic words in Toivonen's view. The Mansi word obviously cannot reflect the same proto-form as the Khanty word, but as the Khanty word could reflect Uralic \**ä*, so the words in Permic and Khanty can either be parallel borrowings or borrowed separately from Iranian. Korenchy accepts Toivonen's argumentation, noting that the Obdorsk word is either borrowed from or influenced by Komi.

As is the case with the other 'iron' words that are borrowed from Iranian, the semantic development from 'sword, knife' to 'iron' is troubling. It is possible that the word was first borrowed in the meaning of 'sword' and then developed later into 'iron', but Korenchy actually assumes that the meaning 'iron' is primary. Unfortunately, there are no traces of the meaning 'knife, sword' in Khanty that would prove this claim correct. The similar meanings in Khanty and Permic makes it likely that the Khanty words are borrowed from Komi, even if the vowel substitution seems atypical.

An Indo-Iranian etymology

### **SaN *gohpi* 'pit' (found in all Saami languages < PSa \**kopē*)**

< PreSa \**kupa*

← PII (or PI) ? \**kūpa-* > Old Indic *kūpa-* 'pit; well' (EWAia I: 325, s.v. *kūpa*)

(Koivulehto 2003; Sammallahti 1998: 7; K. Häkkinen 2004; Holopainen 2018b: 161–162; Rauhala 2019: 163)

Koivulehto's etymology was never published by him, but it can be found in the 2003 handout and the unpublished manuscript. The Indo-Iranian etymology is also mentioned

by Sammallahti (1998), so it is possible that the etymology was invented separately by both Koivulehto and Sammallahti.

According to Seldeslachts (2006: 131), who refers to an old idea by Bernhard Kölver, the Vedic word *kūpa* is actually a Middle-Indo-Aryanism, and the projected Sanskrit form would be *\*ku-āpa-* ‘holding water’. This casts some doubts on the etymology, as we do not know how Indo-Iranian *\*ku-āpa-* would be substituted on the Uralic side.

Koivulehto (2003) mentions the possible cognates Germanic *\*hūfa-* (> Old Icelandic *húfr* ‘bow [of a ship]’, Old English *hyf* ‘beehive’) and Latin *cūpa* ‘barrel’, but EWAia considers the Indo-Iranian background of the word uncertain. If the explanation of Seldeslachts is correct, these cognates have to be rejected. The Indo-European affinity of Latin *cūpa* is in any case doubted by de Vaan (2008: 155, s.v. *cūpa*).

Contrary to Holopainen (2018b: 162), the Finnic word *kuoppa* is probably not connected etymologically to either the Saami or the Indo-European words (see Ylikoski 2019: 195–196). The Komi word *gep* can be a parallel loan from Iranian, although the voiced *g-* in this case makes this assumption uncertain, and also the vocalism does not fit the Indo-Iranian form very well (the Komi word cannot reflect a Pre-Permic form with *\*u-*).

? Indo-Iranian etymology

**Fi *kota* ‘chum’, *koti* (der.) ‘home’, Ka *koti*, *kodi* ‘house’; SaN *goahti* ‘chum’; Mo E *kud*, M *kudo*; Ma *kudo*, *kudâ* ‘summer-hut’; Ko *kola* ‘hunter’s cabin’, Ud *kwala* ‘sanctuary, summer-kitchen’; Hu *ház* ‘house’, *haza* (der.) ‘home, fatherland’; Kh North *χat*, East *kat*, South *χot* (< PKh *\*kāt*)**

< PU *\*kota* (UEW s.v. *\*kota*; Sammallahti 1988: 543; Zhivlov 2014: 137; Aikio 2015b: 61)

← PII/PI *\*kata-* > YAv *kata-* ‘storage room, cellar’ (AiWb 431, s.v. *kata-*)

(Jacobsohn 1922: 225; FUV: 130–131; KESK: 114; Joki 1973: 272–273; Katz 1985: 296–297; 2003: 237–238; Rédei 1986c: 44–45; UEW s.v. *\*kota*; Gulya 1992; 2003: 81–82; SSA I s.v. *kota*; Koivulehto 1999a: 218 = 2016: 220)

It is an old idea that the Uralic word for ‘chum’ (from which meaning ‘house, home’ in several Uralic languages) is a borrowing from Indo-Iranian, and this etymology is accepted by Joki (1973) and Rédei (1986c), as well as Koivulehto (1999a), who briefly mentions the etymology but does not comment on it in more detail. However, Kümmel (2017; 2018a) has argued that the borrowing could have happened to the other direction, as there are no certain cognates of the Indo-Iranian word elsewhere in the family. The old explanation (IEW: 586) that the Iranian word is derived from a PIE root *\*ket ~ kēt-* with such cognates as Slavic *\*kotъ* ‘booth’ and Greek κοτύλη ‘bowl, dish’ has to be considered uncertain: Derksen mentions the possibility that the Slavic and Iranian words are connected, so the Indo-European origin of the Iranian word cannot be ruled out for certain.

On phonological grounds alone it is impossible to solve the direction of borrowing. Most other tentative examples of reverse borrowings are not very good, leaving *\*kota* quite alone in this respect (see Appendix II for the other possible examples). However, it is interesting that as the *o*-vowel usually appears either in labial environments, before *\*r* or

in some cases as the reflex of PIE \*o, this word fits poorly with the other o-loans, as it fills none of the environments here. Contrary to Joki's ideas, this cannot be a Pre-II loan, as the Avestan word cannot reflect PIE \*o- (due to Brugmann's law it would have become \*ā here). Also Lushnikova (1990: 31, 220) considers this as a very early loan that reflects the retained PIE vocalism.

It is also possible that the Uralic and Indo-Iranian words are borrowed from some parallel source. There do not seem to be many words borrowed from a third source into Uralic and Indo-Iranian, but a cultural term like this can of course be an early Wanderwort. This possibility was mentioned already by Joki (1973: 272–273). The Iranian word was also possibly borrowed into East Slavic: Ukrainian *kháta* 'hut' (Kim 2018: 13).

Katz (2003: 237–238) argues that the various Uralic words represent parallel borrowings from Indo-Iranian, but there is actually no reason to support this, as all the words can go back to Proto-Uralic \*kota. Also Gulya (2003: 81–82) has argued in favour of parallel borrowings, although he does not present any detailed argument for his assumption that the Ugric words reflect a separate loan (Gulya argues that the Ugric languages reflect a Proto-Ugric word with \*a, but this is erroneous, and also the Ugric, in this case Khanty and Hungarian, words can be regularly derived from PU \*kota).

? Indo-Iranian etymology

**Est *kärg* 'honeycomb'; ? Md *keras*; ? Mari *karaš*, *käräs* 'Honigscheibe (allg.) Bienenzelle; honeycomb', *karaš müj* (TschWB 224) < PMa \**kāras*; ? Ud *karas* < ? PU<sup>23</sup>**

← ? Indo-Iranian \**kāras* (unattested)

(Räsänen 1920: 245; 1956: 275; Joki 1973: 226–7, 268; Carpelan & Parpola 2001: 123)

The word is discussed here because Joki has considered it a possible Indo-Iranian borrowing, although he also mentions the possibility that the Uralic words are borrowed from Baltic. However, in reality there is no corresponding form attested in Indo-Iranian, and in earlier literature it has simply been assumed that an Indo-Iranian loan from a non-attested cognate of Baltic *korys* 'honey-comb', Latv *kāre(s)* (from Baltic \**kār-*), would be more logical than a true borrowing from Baltic. This is illogical, and the existence of the word in Indo-Iranian cannot be proven. The word is scarcely found in Indo-European, and it is not at all clear that the forms in Greek κήρος 'wax, honeycomb' and Baltic are truly inherited from PIE. Beekes (2010: 689–690) considers them as possible parallel borrowings from a substrate language (for an attempt to explain the relationship through secondary ablaut, see Derksen 2015: 254). Carpelan & Parpola have rightly argued that the Uralic words are, in reality, Baltic loans. Derksen also notes that the Baltic words might be related to the verb *kārti* 'hang' which is also of uncertain origin (possibly from from PIE \**kerH-*).

Carpelan & Parpola and Joki (referring to Räsänen 1969: 256) mention that the word has been borrowed from Uralic, probably from Mari, into the Turkic languages of the Volga

<sup>23</sup> I am grateful to Niklas Metsäranta and Riho Grünthal for useful discussions on the etymology of this word.

region, where it appears as Kazan Tatar *kärüz*, *käräs*, Bashkir *käräd*, Chuvash *karas* ‘honey-comb’. Katz (1985: 283) argues that the Uralic words are not Indo-Iranian but Turkic loanwords, but as the Turkic words have an uncertain etymology and the word is not found in other Turkic languages outside of Central Russia (Joki 1973), it is more probable that the Turkic words are borrowed from Mari, like Räsänen suggested.

Junttila (2012: 271) has considered the Baltic etymology to be “dubious” but without giving any reasons for this. He assumes that the Udmurt and Mari words are possibly borrowed from the Finno-Mordvin word, and mentions the borrowing of the word into Chuvash and Tatar as a parallel. However, one has to note that no unitary Finno-Mordvin/West Uralic proto-form can be reconstructed on the basis of the Estonian and Mordvin words, and it is more natural to assume that the word has been independently borrowed from Baltic or Balto-Slavic into several branches of Uralic. It is also obvious that the Udmurt word *karas* is no regular cognate of the Mari or Mordvin words, as Udmurt *a* cannot correspond to Mordvin *e*, Mari *a/ä* in inherited vocabulary. Because this goes beyond the scope of this dissertation the etymology will not be dealt in more detail here, but one can simply state that there is no reason to consider this an Indo-Iranian loan and the possibility that these words are borrowed from Baltic should be examined in more detail.

Not an Indo-Iranian etymology

**Fi *kärsä* (cognates in all Finnic languages; Vo and Izh possibly borrowed from Finnish; SSA s.v. *kärsä*; EES s.v. *kärss*)**

< ? Pre-Fi *\*kärsä*

← PII/PI *\*karšū-* ‘Furche, Einschnitt, Graben’, OI *karšū-* ‘Furche, Einschnitt, Graben’, root *karš-* > OI *karš-* ‘ziehen, schleppen, pflügen’, Av *karš-*; < PIE *\*k<sup>w</sup>el-s-* (EWAia I: 319–320, 397, s.v. *karš-*, *karšū-*; Werba 171, No. 31 s.v. *krš*; Cheung 2007 s.v. *\*karš/xrah*)

(Parpola 2005: 42–45)

Parpola’s recent etymology involves several problematic details which have to be taken into account when reviewing it. Parpola’s argumentation of the boar’s importance in ancient Iranian culture (Mithra and the boar) makes it plausible that a word for a relatively basic concept would be borrowed. It is also logical to think that the word for a pig’s snout is also borrowed in connection with other hyonymic terms, such as *\*woraći* (see this entry), which is also discussed by Parpola. The importance of the boar is mainly attested in Iranian and not on the Indo-Aryan side, but it is difficult to determine whether the Iranian practice denotes a Proto-Indo-Iranian cultural practice or not.

The Sanskrit noun *karšū-* ‘Furche, Einschnitt, Graben’ is not semantically very close to the Finnic word. It is mainly the verbal root *karš-* ‘ziehen, schleppen, pflügen’ that could be the source of *kärsä*, with the idea that the primary meaning of the Finnic word would have denoted the boar’s use of its snout to root around in the ground.

Two phonological problems make the etymology uncertain. Both are commented on by Parpola, but his arguments are not very convincing. First of all, the *\*-rs-* cluster is strange if the word is borrowed from PII/PI *\*karša-*; here one would expect *\*-rš-* also in the Uralic side. Parpola mentions several possible solutions for this problem: one is that the RUKI rule was not yet operative at the time of borrowing (the Uralic word would reflect Pre-Indo-Iranian/early Proto-Indo-Iranian *\*karsa-* and not later *\*karša-*). Parpola is cautious of this idea, noting that since the RUKI rule is common to Indo-Iranian and Balto-Slavic, it has to be a very old development. While it can be noted here that the irregularities of RUKI rule in Baltic have led some researchers to assume that it operated only partially/incompletely in Balto-Slavic (Hock 2016: 11), a more logical view is to assume that its operation originally was complete in both Indo-Iranian and Balto-Slavic, and that the Baltic irregularities can be explained as secondary, especially as the early Baltic loans into Finnic show RUKI reflexes sometimes when the attested Baltic forms lack the sibilant *š* (Kallio 2008: 267).

Another solution offered by Parpola is that by the time of borrowing, *\*š* was not yet a phoneme in Proto-Uralic, at least not in the cluster *\*-rš-*. Parpola bases his idea on Sammallahti (1988), where *\*š* is not reconstructed for Proto-Uralic but only for the later Finno-Ugric proto-language. However, this kind of assumption is now outdated. We now know that there is no reason to suppose that *\*š* was not present in the earliest phases of the common proto-language of the Uralic languages, as many Proto-Uralic words with *\*š* can be reconstructed, such as *\*šira-* ‘to dry’ (Zhivlov 2014: 133) and *\*kajšV* ‘sickness’ (Aikio 2015b: 57). Moreover, other early borrowings, notably *\*mekši* ‘bee’ (from Pre(!)-Indo-Iranian *\*mekši-*) show that *\*š* existed as a substitute of Indo-Iranian *\*š* already at a very early period of contacts. The word *\*kärsä* could not be borrowed from any other branch, as *\*r* here is from earlier *\*l* (Werba s.v. *krš*), so the absence of RUKI cannot be explained with this either. Thus Parpola’s explanation is problematic. It is, however, true that no Proto-Uralic words with the cluster *\*-rš-* can be reconstructed, so it would not be impossible to assume a substitution of *\*-rš-* by *\*-rs-*.

Another problem, also noted by Parpola, is Finnic *\*ä* instead of *\*a* (or *\*o*). The vowel *\*ä* can be found in place of Iranian *\*a* in many loanwords in Permic and Ob-Ugric, but in the early borrowings that have reflexes in Finnic this substitution is not found. Parpola suggests that this might be an instance of a secondary front-vowel in Finnic – the word would have been borrowed as *\*karsa*, but secondarily developed to *\*kärsä*; parallels for this development are found (see also *\*mükkä*). The Finnic word could not, in fact, reflect a Proto-Uralic *\*ä-ä* stem, as this would have developed into an *a-e* stem in Finnic (Aikio 2015b: 39–47).

These things taken into account, the etymology cannot be considered certain. Interestingly both *\*ä* and the absence of the RUKI rule could point to a later Iranian language, as is the case with West Uralic *\*warsa*. Ossetic has a cognate of the Sanskrit verb: I *xæssun/xast*, D *xæssun/xast*, *xæssun/xarst* ‘to carry, to breed...’, and phonologically it would be possible to derive the Finnic word from Pre-Ossetic/Alanic, although there is no evidence of a noun that could be a precise donor form.

Parpola also assumes that Fi *karhi* (< ? *\*karši*) ‘harrow made of logs by breaking off branches’, could be borrowed from the same Indo-Iranian word, more specifically from *\*krší-* (> OI *krší-*) ‘ploughing, field’ but this time showing *\*-rš-* as the substitution for Indo-Iranian *\*-rš-*. If both *karhi* and *kärsä* were correct etymologies, *karhi* should be an

earlier loan. Semantically the etymology for *karhi* is better than that of *\*kärsä*. However, because the word is found only in some dialects of Finnish and not in other Finnic languages, the Indo-Iranian origin is far from certain.

Possible cognates for *karhi* have been suggested in Mordvin M *kartša* ‘stick, dry twig (= Fi *risu*)’, ? E *kurtšt*, *kartšt* ‘dry twigs’ and Hill Mari *karša* ‘twigs rotten in the water; thicket’ (SSA I s.v. *karhi*). These words are not regular cognates even with each other (the Mordvin *a ~ u* vocalism is totally irregular, and the stem-vowel *-a* of the Moksha word does not regularly derive from an Uralic *\*i*-stem). Also semantically these words are rather far from the Finnic word. It is probable that the Mari and Mordvin words are of completely different origin, and the Indo-Iranian origin of Finnic *karhi* remains a possibility.

? Indo-Iranian etymology

**Fi *mato* ‘worm’ (cognates in all Finnic languages except Li); SaN *muohcu* ‘moth’ < PSa *\*muoçe* (distribution: SaU, SaL, SaI, SaN, SaKo, SaKld; Álgu s.v. *muohcu*)**

< West Uralic *\*mača*

← PI *\*matsa-* < PII *\*mača-*; Pashto *m’āšay* ‘mosquito’ < *\*mātsyaka-*, OI *maśáka-* < *\*mačaka-* (EWAia II 334–335, s.v. *maśáka-*; Morgenstierne 2003: 53, s.v. *m’āšay*)

(Koivulehto 1999a: 11; 2001b: 69)

The etymology is convincing both semantically and phonologically. The Finnic meaning ‘worm’ looks far-fetched, but Saami has retained the more original meaning of a ‘flying insect’, and parallels to such semantic development exist: for example, Old English *maða* denotes both ‘worm’ and ‘maggot’.

This is clearly a Proto-Iranian loan and one of the best examples of Koivulehto’s substitution rule *\*-č-* ← PI *\*-ts-* word-medially; an earlier Proto-Indo-Iranian source would have resulted in a Uralic form *\*\*mača*, and a later Iranian (“Alanic”) original would have given plain *-s-* in both Finnic and Saami.

The labial vowel of the non-initial syllable seen in the modern Finnish and North Saami forms is clearly secondary, as the illabial vowel is still retained in the forms of certain Saami languages, such as SaU *müehtsee*, SaL *muohtsē* (< *\*mača*), meaning that the Finno-Saamic word is an original *a*-stem like most of the Iranian loanwords.

As the Saami and Finnic words are regular cognates, this word is clearly an inherited word in both of them and not a borrowing into either direction. The word probably belongs to the same layer of loans as other Proto-Iranian loans showing the same affricate substitution, some having a “West Uralic” or Finno-Permic distribution.

Kallio (2000) has noted that Proto-Germanic *\*maþan-* ‘moth’ is borrowed from this Uralic word, which seems likely, as there is no convincing Indo-European etymology for the Germanic word; the Germanic word and Iranian *\*matsa-* cannot be regular cognates. Earlier the opposite direction of borrowing from Germanic to Finnic has also been suggested (T. Sköld 1984), but due to the uncertain Indo-European background of the

Germanic word, the borrowing from Finnic is more convincing. LÄGLOS (II s.v. *mato*) considers the Germanic etymology of Finnic *\*mato* uncertain (the Iranian etymology was not yet known to the editors of LÄGLOS).

Mallory & Adams (2006: 150) derive the Germanic words *\*muþþa-* ‘moth’ and *\*maþan-* ‘maggot’ from a PIE root *\*mat-*, which according to them appears also in Armenian *mat’il* ‘louse’ and Av *maðaxa-* ‘grasshopper’. This scant distribution raises questions about the status of such an “IE root”, and nothing of the sort can be found in LIV<sup>2</sup> or NIL. The Armenian word cannot be found in Martirosyan’s etymological dictionary. Kroonen (2013: 358) does not consider the Germanic words as reflexes of a PIE root *\*mat-*, but he considers the Germanic words for ‘maggot’ and ‘moth’ as “North-European” (Balto-Slavo-Germanic) words, and he mentions Russian *мотыл* ‘maggot’ as a cognate of these words. Kroonen assumes that both PG *\*maþan-* ‘maggot’ and *\*muþþa-* ‘moth’ split off from “an ablauting paradigm *\*maþō*, gen. *\*muttaz* continuing *\*mót-on*, *\*mt-n-ós*. The etymology of Russian *мотыл* is not mentioned in either of the dictionaries of Derksen (2008; 2015).

An Indo-Iranian etymology

**Fi *maksa-* ‘to pay’ (cognates in all Finnic languages); Md E, M *makso-* id.**

< West Uralic *\*maks-*

← PI *\*ma(n)dza-*, cf. Avestan *mąza.raii* ‘giving wealth’, OI *mamh*, Av *mąz* < PII *\*manʒh-* (EWAia II: 289 s.v. *mamh*; Werba 1997: 214, No. 136 s.v. *mamh*)

(Koivulehto 1999a: 221; 2001b: 255 = 2016: 223, 297)

The etymology is suspicious, as there are only two credible examples of the substitution *\*-dz-* > *-ks-* (see the entry on *\*paksu*). Semantically the comparison makes sense, and there is no other compelling etymology for the West Uralic word. Koivulehto has also rightly argued that as the word is homophonous with the word *maksa* ‘liver’, it would be strange for a similar word to arise as an *Urschöpfung* (which is a suspicious way to create words anyway). One could note, however, that the original vowel in the word for ‘liver’ was *\*i̯*, and this would have probably still been retained at West Uralic times, so the words *\*maks-* and *\*miksa* would not have been homonymous at the time of borrowing, before the Proto-Finnic stage.

Saami *\*mäksē-* (> SaN *máksit*) ‘to pay’ is mentioned by SSA as a possible cognate or alternatively a Finnic borrowing. Lehtiranta (2001: 74–75) and Sammallahti (1988: 253) consider the Finnic and Saami words as cognates. It is well-known that the Latvian word *maksāt* ‘pay’ is a late borrowing from Finnic. The regular reflex of Uralic *\*a* in *a*-stems is Saami *\*ō*, which points to the conclusion that *\*mäksē-* is a loan from Finnic.

The Iranian etymology cannot be ruled out, but due to lack of parallels the etymology cannot be considered certain.

? Indo-Iranian etymology

**Fi *mana* ‘death’, der. *Manala* ‘land of the dead’; *mana-* ‘schimpfen, (ver)fluchen, beschwören, tadeln; verklage’, also in Karelian, Est *mana-*, ? *mõna-* < ? PFi *\*mana-* (SSA s.v. *manata*)**

← PI/PII *\*manā-* ‘Andacht, Andachtslied, Frommer Eifer, Verdacht’, or *\*mánas-* ‘Sinn, Geist, Verstand, Gedanke’, OAv, YAv *manah-* ‘Gedanke, Denken, Denkkraft’, root *\*man-* ‘denken, meinen’ < PIE *\*men-* (EWAia II: 305, s.v. *MAN*, 307–308 s.v. *mánas-*; Werba 1997: 215–216 No. 137)

(Koivulehto 1996: 324, footnote 6)

Koivulehto mentions this as one of the Indo-Iranian loans confined to Finnic, but he did not discuss the etymology in detail in any publication. In previous research, cognates from other Uralic languages have been suggested: SaS *moene-* ‘say, mention’, Md E *muña-* ‘to conjure’, PMa *\*mana-* ‘to say, utter’, Hu *mond* ‘to say, utter’ and PS *\*mā(n)-* (> TN *mān-*) ‘to say’ (UEW, s.v. *m8n3*) but these are rejected as irregular by Aikio (2014d: 62, 71). The word could belong to the category of mythological influence from Indo-Iranian (see *\*juma*).

The etymology is phonologically and semantically possible, but also a Germanic etymology (PG *\*manōjan-* > OE *manian* ‘mahnen’) has been suggested, and this is considered as possible by LÄGLOS (II s.v. *manata*), although it is also noted that the Finnic words can be later loans from Old Swedish *mana* ‘mahnen’. Aikio (2014d: 62) supports the Old Swedish etymology. LÄGLOS discusses only the etymology of the verb *manata*, not the noun *mana*. It remains uncertain whether the noun and the verb are etymologically connected. The meanings of the Indo-Iranian word fit better the verb *manata* than the noun which denotes ‘death, land of the dead’.

Even if the Finnic word is from Germanic, the assumed cognates in Mordvin, Mari, Hungarian and Samoyedic could be borrowed from Indo-Iranian. As these words also do not stand in a regular relationship with each other, it is possible that these words are borrowed separately from Indo-Iranian. Mari shows *a* in both main varieties, which means that the word cannot be an old, inherited word, and it is unlikely that it is borrowed from Indo-Iranian. SaS *moene-* has been explained as a Germanic loan from PG *\*man-* ‘to think, remember’, which is cognate to the Indo-Iranian word discussed here (Aikio 2006b: 32–33; 2014c: 62). The Hungarian word can hardly be derived from a form with earlier *\*a*, as *o* usually reflects *\*u*, and it is very unlikely that the Hungarian word could have been borrowed from Indo-Iranian. Samoyed *\*ā* could reflect earlier *\*a*, so deriving this word from Indo-Iranian *\*manā-* or *\*mánas-* could be possible. The simple meaning ‘to speak’ could have developed from the meaning ‘to think, to mean’, although this is not a very obvious semantic development. Mordvin *muña-* ‘to conjure’ fits better the mythology-related meanings of the Indo-Iranian words. The word could reflect an old *\*o-a* or *\*a-i-* stem, but the palatal *ń* here can hardly be secondary, so the pre-Mordvin form should be *\*mońa*. Formally the word could also reflect an earlier *\*a-i-* stem, but the stem vowel *-a* is unexpected in this case.

If the Mordvin and Samoyed words are Indo-Iranian loans, it is very difficult to determine the exact form from which they have been borrowed. As most Indo-Iranian *s-* stems were borrowed as *a-* stems in Uralic, the word could have simply been borrowed from *\*mánas-*. Also the collective *\*mánās* would have led to the same kind of result. If the

word were really a late Iranian borrowing, where \*-s would have already become \*-h, this would also result in Uralic/Pre-Finnic *mana-*, but as noted, even this form could have been borrowed as a Uralic *a*-stem. *\*manā-* would be semantically more convincing, however. The Indo-Iranian word *\*manas-* continues the PIE proterokinetic *s*-stem *\*mén-(o)s-* > PII *\*mánas-* > Av *manah-*, OI *mánas-* ‘Sinn’, cf. Gr *μένοϛ*; > amphikinetic collective *\*mén-ōs* > PII *\*mánās* Av NA.PL:N *manā*° (Steer 2015: 256–258; EWAia II: 307–308)

To sum up, the Indo-Iranian etymology for Finnic *mana* or *manata* is not very convincing, and there remains a possibility to derive the Samoyed and Mordvin words from Indo-Iranian, although this entails various problems.

? Indo-Iranian etymology.

**Hu magyar ‘Hungarian’; Ma North *mańsí*, South *mǎńcǐ*, East *möǎńś*, West *mǎńś* ‘Mansi; child that has not been baptised’ (< PMs ? *\*mǎńć-*); Kh North *mǎś*, East *mańt* ‘name of a phratry’ (< PKh ? *\*mǎńć*)**

< ? PUG *\*mańćV*

← ? PI/PII *\*manu-š* > Av *\*manu-š-*, *manuš.čǐθra-* personal name, ‘descendant of *Manuš*’, OI (Vedic) *mánu-* ‘Mensch, Mann, Menschheit’ < ? PIE *\*man-* or *\*mon-* (EWAia II: 309)

(Korenchy 1972: 60, No. 13; UEW s.v. *mańćś*; Katz 1985: 148; 2003: 124, 168; Ligeti 1986: 136; Lushnikova 1990: 225–226; Helimski 1991: 221; Abondolo 1996: 84; Zimmer 1990: 15, footnote 43)

This etymology is accepted in most sources despite its obvious phonological problems. Interestingly, Joki (1973) does not even mention the etymology. Semantically the etymology is convincing, and there are other borrowings from Indo-Iranian to Uralic denoting ‘man, people’, such as *\*mertä* and possibly *\*mari*. While it has been seen as problematic by some that ethnonyms and words for people are borrowed from neighbouring languages, this seems to be surprisingly frequent, and can probably be explained through the prestige of the contact languages.

However, we do not know what this ethnonym originally referred to: it is uncertain in what kind of situation this was borrowed, and whether the word was used as an ethnonym in the Indo-Iranian side, or whether the word was borrowed in the meaning of ‘man, human’.

Zimmer (1990: 15, footnote 43) briefly criticizes the etymology, arguing that the borrowing of the ethnonym in the meaning of ‘human’ from another ethnic group looks unlikely, even if these group was a *Herrenvolk*. Ethnonyms in general do indeed get borrowed, so Zimmer’s argument here is not very strong: for example, the ethnonym of the Russians (*rus*, *russkiy*) is probably borrowed from Finnic *\*rooči*, which originally referred to Swedes (see the recent discussion about this etymology by Schalin 2014: 428).

It is uncertain whether PII *\*manu-* reflects a Proto-Indo-European word or not. EWAia mentions that Germanic *\*man-n-* ‘man’ and OCS *mǫžǐ* ‘man, husband’ (< ? *\*mon-g<sup>(w)</sup>yo-*) are related to the Indo-Iranian word. However, Kroonen (2013 s.v. *\*mannan-*) notes that

the phonological connection between *\*manuš-* and Germanic *\*mannan-* ‘man’ is phonologically impossible (the assumed IE proto-form *\*menu-* would give *\*minu-* in Germanic). Derksen (2008 s.v. *mōžī*) mentions the old explanation of these words as cognates. Also Matasović et al. (2016 s.v. *muž*) consider the traditional explanation as possible, but not certain.

In Hungarian the word is attested only in the obscure compound *magyar*, which denotes only ‘Hungarian’, and has been explained from *\*mańci-erkV*, where *\*erkV* is an old Uralic word for ‘man’<sup>24</sup>. This compound etymology is usually accepted, although it has been criticized recently by Németh (2017) in an unpublished presentation (Németh is also critical of the Iranian etymology of the Ugric word). In Mansi the word is an ethnonym. In Khanty the word denotes one of the traditional phratries of the Khanty.

The word is not reconstructed for Proto-Ob-Ugric by Zhivlov (2006), so even though there exists no commentary on the Ob-Ugric forms by Zhivlov, the absence of forms in his dissertation points to the conclusion that the word cannot be reconstructed for Proto-Ob-Ugric or preceding proto-languages. Honti (1982: 164–165) reconstructs the word for Proto-Ob-Ugric: the word appears in Honti’s list of reconstructed Proto-Ob-Ugric words, but he reconstructs the alternation of a palatal and a velar vowel already for Proto-Ob-Ugric, giving the forms *\*māńć*, *\*māńć*. The impossibility of reconstructing a unitary proto-form points to the conclusion that the words are parallel borrowings.

Although semantically the etymology is more or less problematic, phonological problems appear in both the substitution for the tentative Indo-Iranian original and the irregular relationship of the Ugric words. Starting from the latter, it can be easily observed that the Hungarian word has to reflect earlier *\*mańci*, whereas Mansi points to a front-vowel form (*\*māńci* or the like). Mansi *\*ä* is caused by the palatal consonant according to Korenchy, but this seems an ad hoc explanation.

The irregular relationship of the “cognates” in the three Ugric languages could point to a parallel borrowing, as is the case with several other Iranian loanwords. In Hungarian also front-vocalic forms are observed, such as *megyer*, which is attested already in 9th-century Byzantine sources. Nevertheless, most early attestations clearly reflect the velar form *magyar* (TESz s.v.). As the Hungarian word is a compound, the varying forms are usually explained as generalizations of the vowel harmony of either the *mańci-* or the *-erke* part of the compound.

These problems taken into account, one cannot be certain about the Indo-Iranian etymology of this Ugric word. As a large part of the common vocabulary in the Ugric languages derives from an unknown source, this important term could equally well belong to this group of (substrate, superstrate?) words. Semantically the old Indo-Iranian explanation is tempting, but at the moment there is no possibility to explain the phonological irregularities.

---

<sup>24</sup> The reconstruction of a PU word *\*irkä* or *\*ürkä* is problematic, as similar words are found in Turkic languages also, and the Hungarian and Mari “reflexes” of this Uralic word can be also explained as Turkic borrowings. As the compound *magy-ar* does not have to be older than Old Hungarian/Proto-Hungarian, the element *-er* could be an early Turkic loan acquired during the early first millennium AD. The word has irregular reflexes in Finnic *yrkä*, where the situation is complicated by the parallel-form *ylkä* and its Saami cognates; their relationship has to be investigated in more detail elsewhere, but it looks unlikely that both *ylkä* and *yrkä* could reflect the same Uralic form *\*irkä* or *\*ürkä*, especially as the Finnic form with *-l-* has an exact cognate in Saami. Also Mari *erye*, *eryə* ‘boy’ cannot be regularly derived from either *\*irkä* or *\*ürkä*.

Another problem concerns the substitution of *\*manuš-* by *\*mańí-*. It is unclear why this donor form results in a *\*-ńć-* cluster on the Uralic side – one would not expect a palatal affricate in this position based on Indo-Iranian *\*-š-*. Korenchy explains this by assuming that because a cluster *\*-nš-* was unknown in Uralic, *\*-ńć-* was the most fitting substitution. This explanation is also supported by Abondolo (1996: 84). While there is logic in this explanation, it is not faultless. Here we have to assume that the *u* was syncopated at the time of borrowing, but why could the Uralic speakers not borrow the word simply as *\*\*manV* or as a three-syllable word? Another option would be to suppose that *\*ć* is a derivational element on either the Indo-Iranian or Uralic side, but no such element can be found.

The connection of the Hungarian ethnonym *magyar* and the word *mese* ‘tale’ and its Khanty cognates (Khanty *V mańt’*, DN *mońt’*, O *maś* ‘Sage’) is an old problem of Ugric etymology, even though in the Neogrammarian framework of regularity it is impossible to derive these two words from a common source, it has been attempted (see the UEW with references). Already Abondolo (1996) criticized these attempts. Recently Honti (2017: 129–147) and Róna-Tas (2017: 66–68) have discussed the etymology and concluded that the two words are originally distinct but have become etymologically connected. However, this does not help much in determining the possible Indo-Iranian origin of the ethnonym.

? Indo-Iranian etymology

**Ma marij, marâ ‘man, husband; Mari’** (TschWB: 372)

← PI/PII *\*mar(H)ya-* > YAv *mairiia* ‘Schurke, Bube’, OI *márya-* ‘Jungmann, Jüngling (auch junger Krieger, junger Mann einer Frau, Liebhaber’; < ? PIE *\*mer(H)yo-* ‘junger Mensch’ (EWAia II: 329–330, s.v. *márya-*; NIL s.v. *mer*)

(Munkácsi 1901: 456; Jacobsohn 1922: 245–246; Joki 1973: 280; E. Itkonen 1956: 78; Bereczki 1994: 86; Parpola 2015: 85–86; Milanova, Holopainen & Bradley [forthcoming])

Munkácsi’s etymology includes problems that have been pointed out by the subsequent sources. Semantically the etymology is completely plausible: the Indo-Iranian meaning ‘(young) man, mortal’ (attested in both Old Iranian and Vedic) and the Mari word denoting ‘man’ as well as ‘Mari’ fits this semantically well (it is not uncommon for ethnonyms to develop out of words meaning ‘human’).

The more distant etymology of the Indo-Iranian word is open to different interpretations: LIV<sup>2</sup> (s.v. *\*mer-*, 490, footnote 2) mentions that the IE word the word has been derived from the root *\*mar-* (< PIE *\*mer-*) ‘to die’ by McCone (1987); McCone also connects Old Irish *muire* ‘Anführer’ here. This is not universally accepted, however. It is usually assumed (EIEC: 655–; Kölligan 2018: 2273) that the PII word comes from PIE *\*meryo-*, with a cognate in Greek *μεῖραξ* and possibly in Celtic (cf. Welsh *morwyn* ‘girl’, *merch* ‘daughter’) and Baltic (Lithuanian *martì, mergà*).

Peters (1980: 222) refers to Thieme (1963: 238–243), who reconstructs Latin *\*marī* as the basis for *maritus* (Watkins [1957: 280] reconstructs an *i*-stem *\*maris* ‘young person’ as the basis of the Latin word). AIEW (618–619: s.v. *martì*) is of the opinion that Baltic *\*marti-* could formally continue a PIE formation *\*mortih<sub>2</sub>-* from the root *\*mer-*, from

which the aforementioned Indo-Iranian and Greek words can also be derived, as well as Latin *marītus*, although there it is considered problematic that the other IE words feature no *-t*. ALEW also mentions the possibility to derive these Indo-European words ultimately from *\*mor-to-* ‘Mensch’.

Beekes (2010: 921–922) derives the Greek and Indo-Iranian words from a PIE noun *\*mer-io-* and states that further connections are uncertain. Derksen (2015: 306, 342) is supportive of the connection of the Baltic words *marti* and *merga* to the Indo-Iranian and Greek words and Welsh *merch*.

However, although the resemblance of the modern Mari and the Old Iranian and Indo-Aryan words is also transparent, the initial syllable vocalism is difficult to explain from a diachronic point of view. PU *\*a* develops into Mari *o* or *u*, and as a rule it can be stated that Mari *a* (= in both Meadow and Hill Mari) never appears in words that are inherited from times earlier than Proto-Mari (*a* can thus regularly reflect no other Pre-Mari vowel either, meaning that Pre-Mari *\*o* and *\*i*, which could substitute for Indo-Iranian *\*a*, are also out of the question).

The problem of Mari *a* has been long noted by loanword research too, but Joki (1973: 280) considers the Mari word an Iranian borrowing despite this problem. Joki’s solution is to consider the Mari *a* a new phoneme that would have risen in the Proto-Mari vowel system, but the problem here is that Joki considers Proto-Mari significantly too early (1000–500 BC) than is plausible in the light of modern research. The problem with *\*a* is also remarked already by E. Itkonen (1956: 78).

It remains unclear how to solve this situation. One could argue that there is a way for Pre-Mari *\*a* to be reflected as Mari *a* in some environments that have simply gone unnoticed until now, but one would need much more examples of “early” *a* to substantiate this kind of claim. Another chance would be to suppose that here *\*a* is retained irregularly – the use of the word as an ethnonym (proper noun) could mean that it could escape some sound laws (this is assumed by Bereczki 1994), but the fact that the word is primarily a word for ‘man’ speaks against this.

One possibility would be to consider the word a late borrowing from some neighbouring language which would have borrowed it from Indo-Iranian earlier. The problem is that there is no such word in the neighbouring contact languages, such as Chuvash or Tatar.

It is, of course, possible to assume that the Iranian presence near the areas where Mari is spoken was so late that Mari could have borrowed the word after the sound changes which made Pre-Mari *\*a* disappear had already taken place. From the point of view of Iranian ethnohistory this is an important question. It is, of course, known that before the spread of the Bolgars in early Middle Ages the Iranian languages were widespread in central parts of European Russia, but a borrowing before the early Middle Ages would probably be too early, as based on later Turkic borrowings the divergence of Proto-Mari is dated to the later Middle Ages (Bereczki 1992: 12–15). However, as the Alans played an active role in the Mongol army in the high Middle Ages (Alemany 2000: 372), it is not impossible to assume that some forms of Iranian could have survived around the Volga until such late times too (the mere presence of the Alans in more southerly areas is of course not enough to prove this). Joki specifically notes that this word is probably borrowed from Alanic, but here the problem is that a suitable donor form is not found in Alanic (or Ossetic). For late Mari *a* reflecting Iranian *\*a* one would need more parallel examples, but there are almost no other potential Iranian loans with Mari *a*.

? Indo-Iranian etymology

**Fi *marras* (: *marta-*), also in Karelian (< PFi *\*martas*)**

← PII/PI *\*marta-* ‘dead’ > OI *márta-* ‘mortal, human’ (< ? PIE *\*morto-*)

**or**

← PII/PI *\*mr̥ta-* > OI *mr̥tá-* ‘dead’, verbal adjective from the root *mar-* ‘to die’, PIE *\*mer-* ‘to die, to disappear’ (EWAia II: 318–319, s.v. *MAR*, 327, s.v. *márta-*; LIV<sup>2</sup>: 488–491, s.v. *\*mer*; Werba 1997: 218–219, s. 143 PIE *\*mer-* > PII *\*mar-*, > OI *mr̥* ‘sterben’, OP/Av *mar*)

(Mikkola 1902: 72; FUV: 155; Joki 1973: 280–281; Katz 1983: 174–177; 2003: 123; Gamkrelidze & Ivanov 1984: 925; K. Häkkinen 1987: 254; Dolgopolsky 1989: 20; Lushnikova 1990: 227–228; Helimski 2000: 192; Koivulehto 1999a: 228–229; SSA s.v. *marras*)

This is a well-known old etymology. The Finnic word is a parallel loan from the same source as *\*mertä* (see below) – it is, however, obvious that the two words are parallel loans (see below). It is interesting that there are several words meaning ‘man’ or ‘mortal’ borrowed from Indo-Iranian. This word shows the substitution *\*-ar-* or *\*-i̯r-* for PII or PI *\*r̥*, which is found in a number of loans, alongside *\*er* and *\*or*. The forms with *\*-ar-* have a limited and westernly distribution, which probably points to a (relatively) late borrowing, perhaps from Proto-Iranian, but theoretically (= from the point of view of Iranian historical phonology) also a Proto-Indo-Iranian source would be possible. It is impossible to say whether Uralic *\*-ar-* here reflects truly the retained syllabic *\*r̥* or a later development into a sequence *-ar-* as seen in later Iranian languages. It is also impossible to determine whether Finnic *-ar-* reflects earlier *\*-ar-* or *\*-i̯r-*, as the two vowels merged in most environments in Finnic, and in *a*-stems their difference cannot be seen from modern Finnic alone.

There are various views about the exact donor form, as phonetically both *\*mr̥ta-* and *\*marta-* would be suitable sources. Lushnikova (1990: 227–228) and Helimski (2000: 192) have assumed that the Finnic word is borrowed from the full-grade form *\*marta-*. Gamkrelidze & Ivanov (1984: 925), on the other hand, assume that the Finnic word has to be from zero-grade *\*mr̥ta-*, as a substitution *\*a* by *\*a* here would contradict their rule of *\*o* in loans in front of *\*r̥*. Dolgopolsky (1989) mentions both possible donor forms.

In this case the semantics do not help much, as the Finnic meaning can probably be derived from either ‘mortal’ (attested for OI *\*márta-*) or ‘dead’ (attested for OI *mr̥ta-*).

If the word is borrowed from the zero-grade form, it would be tempting to argue that the different substitutions of *\*r̥* point to chronological differences. There are a few etymologies where *\*er* appears in Mordvin and Permic, and since these branches are not particularly closely related to each other, one could suppose that such words are early borrowings retained in these branches. The more westernly loans, such as *\*martas*, could then be later borrowings.

Also the ending *\*-as* appears in later borrowings, as Helimski (2000: 194) has noted. So although there are no phonetic means to show that the word is an Iranian borrowing, the limited distribution and the consonantal stem would point to that. Theoretically the Finnic word could be borrowed even from Proto-Indo-European: the syllabic resonant would have been substituted similarly, and the ending *-as* (if it were possible on the Uralic side already) could easily substitute for PIE *\*-os* too, due to the absence of labial vowels in the earliest stages of Uralic.

K. Häkkinen (1987: 254) criticizes Rédei's idea of dating this word to Proto-Finno-Permic. Rédei assumes (correctly) that the Finnic word has to be a relatively early borrowing, but Häkkinen considers it methodologically weak to assume that the word was borrowed into a Finno-Permic stage, as there is nothing in the distribution of the word to suggest this, and as the Finno-Permic and Proto-Finno-Ugric stages are phonologically identical, the borrowing could be considered just as easily Proto-Finno-Ugric (= Proto-Uralic) as Proto-Finno-Permic.

In very early sources such as FUV, this Finnic word has been etymologically connected to Permic *murt*, *mort* and Mordvin *miřd'e*, although it is clear that these words cannot be derived from the same proto-form but reflect parallel borrowings (see the entry on *\*mertä* for more detailed discussion of the Mordvin and Permic forms).

UEW considers the Permic words as cognates of the Finnic word, although it notes that the consonant correspondence Fi *-rt-* ~ P *-r-* is irregular. This would point to the conclusion that the Permic and Finnic words are not cognates but parallel borrowings at best. The meaning 'unfruchtbar infolge Alter' in Finnish is surely reminiscent of the meanings that are found in Permic, but these can be considered parallel developments. The Permic meanings 'barren, sterile' can be derived from II *\*mar-*, but it remains unclear from which form the Permic word is borrowed. NIL lists various meanings in other Indo-European languages, such as Lith *māras*, OCS *morĭ* 'disease, pestilence'. These are no nearer to the meanings of the Permic words than the Indo-Iranian ones.

An Indo-Iranian etymology

**Fi** *maa* 'earth, country'; ? **Md E** *mastor*, **M** *mastər*; **Ma (compound)** *mlande*, *müländə* (< PMari *\*mü-*); **Ko** *mu*, **Ud** *mu* (< (PP *\*mo*)); **Mns West** *mē*, **North** *mā* (< PMs *\*mī*); ? **Kh East** *məy*, **Kh North** *mÿ* (< PKh *\*miy*)

< **PU** *\*miya* or *\*miyi* 'earth' (UEW: 263–264; Zhivlov 2014: 133; Aikio 2015b: 58)

← ? **PIA** *\*mahī-* 'big' < **PII** *\*majhī-*, **OI** *mahī-* 'big; earth'; < **PIE** *\*meg-h<sub>2</sub>-ih<sub>2</sub>-* (EWAia II: 337–339, s.v. *mahānt-*)

(Koivulehto 2009: 88, footnote 10 = 2016: 418, footnote 10)

Koivulehto (2009) has proposed an Indo-Aryan etymology for this Uralic word with a wide distribution. He does not argue decisively for an Indo-Aryan origin and it is the readers' task to work out Koivulehto's reasons for an Indo-Aryan etymology. The main reason seems to be the substitution of the PIA *\*-h-* by the Uralic spirant *\*ɣ*. However, this would mean that the Indo-Aryan deocclusion of PII *\*ǰh* had happened very early, which is unlikely

due to Mitanni evidence which clearly shows that a sibilant was retained in Proto-Indo-Aryan and later (see also the entry \**sayi-* ‘to get’).

As the Khanty vowel is irregular (this is noted by Zhivlov 2014), it could be tentatively be suggested that this is a result of an inter-Ob-Ugric borrowing, or that the word is a parallel loan in Khanty and in different Uralic languages. Also the other Uralic forms include various irregularities. It is uncertain whether Mordvin *mastor* belongs here as some kind of derivative. The Uralic stem-vowel is reconstructed as *-a* by Zhivlov, but *-i* by Aikio, which points to the conclusion that the Uralic correspondences in this word have not been yet worked out properly.

Katz (1985: 103–104) has suggested that the Khanty word could be derived from Pre-Indo-Iranian FUA \**méžh<sub>2</sub>ih<sub>2</sub>* (> OI *mahí-*), but this idea involves impossible sound laws, and the etymology has been criticized by Rédei 2004b. Katz derives Fi *maa* from “frühurarisches” \**gžmā*, but this is likewise not very convincing. It is better to assume that despite the unexpected Khanty vowel, the word can be reconstructed for Proto-Uralic.

Not an Indo-Iranian etymology

**Fi *mehiläinen* ‘bee’ (derivative; cognates in all Finnic languages); Mo E *mekš*, *mākš*, M *meš* id.; Komi *moš*, Udmurt *muš*; Hu *méh*, *měj* (dial.) id.**

< PU \**mekši* ‘bee’ (UEW s.v. \**mekše*; Sammallahti 1988: 545)

← Pre-II \**mékš-* > PII \**mákš-* > OI *mákš-* ‘bee, fly’, (der.) *mákšikā*, also *mākšā-*; PII \**makši-* > YAv *maxši-* ‘fly’ (EWAia II: 287; KEWA II: 540–542, s.v. *mákšā*)

(Munkácsi 1901: 458; FUV: 132; Schindler 1972: 36; Joki 1973; Katz 1985: 282; 2003: 225–226; Gamkrelidze & Ivanov 1984: 923; Lushnikova 1990: 231; Parpola 1999: 199–201; 2005: 40, footnote 247; Koivulehto 1999a: 216; 2001b: 247, 248)

This is a widely-accepted etymology that was discussed already by Munkácsi, and one of the examples of a very early Indo-Iranian loan with retained PIE \**e*. The direction of borrowing has been at times suspected also to be from Uralic to Indo-Iranian (Jacobsohn 1922: 161; Moór 1956; KEWA; Schindler 1972: 36), but this idea is based on outdated views of the Proto-Uralic lexicon. Uralic \**mekši* is an opaque noun, but Schindler mentions the possibility that the word is derived from \**meki* ‘honey’. A word \**meki* has been reconstructed for Proto-Uralic on the basis of Mari *mü* ‘honey’ and the Ob-Ugric words for ‘honey’ (Mansi *may*, Khanty *mäχ*); in the UEW the reconstruction is given as \**mäkä*. However, Aikio (2014) has shown that the Mari word is borrowed from Udmurt *mu* (which is a regular reflex of Uralic \**meti* ‘honey’), and the word \**mäkä* is thus confined to Ob-Ugric only. There is no Proto-Uralic word from which \**mekši* could be derived from, and even if the Ob-Ugric word had a Uralic predecessor, the reconstruction would be \**mäkä*, and from such a form \**mekši* with \**e* could not be derived from. Katz (2003: 226) has argued that the Ob-Ugric word for ‘honey’ could be a loan from PII \**makš-* ‘bee’ but this is unlikely for both semantic and phonological reasons.

Gamkrelidze & Ivanov (1984: 923) argue that the Uralic word could be a later loan from Iranian, and the vocalism can have been influenced by *\*meti* (which is a Tocharian loan according to them; see the entry *\*meti*), but this is an *ad hoc* explanation.

The Uralic words are regular cognates to each other, so the word was clearly borrowed into Proto-Uralic already and not separately into the already distinguished branches. It would be unlikely that a loanword this early could have been borrowed parallelly into the Uralic daughter languages.

UEW mentions, rather confusingly, that the word is an “Indo-European” borrowing (“< ieur. *\*mekš-*”), without describing the situation in more detail. The word is obviously not borrowed from Proto-Indo-European, as the Indo-Iranian word has no cognates elsewhere in the family, and the word cannot be reconstructed for Proto-Indo-European. Moreover, Uralic *\*š* clearly reflects the Indo-Iranian *\*š* (resulting from the RUKI development), which was not yet a phoneme in Proto-Indo-European. The word belongs obviously to the same layer of borrowings as several other words that display retained PIE vocalism but show Indo-Iranian developments in the consonants, such as SaN *geavri* < *\*kekrä* ← Pre-II *\*kekro-* < PIE *\*kwekwlo-* and Fi *kehrä* ‘spindle’ < *\*kečrä* ← Pre-II *\*ketstro-* (cf. Koivulehto 1999, 2001).

There are various views about the background of the Indo-Iranian noun *\*makš-*. Lubotsky (2001b: 314) lists the word *\*makš-* as one of the inherited words which are confined to Indo-Iranian but he does not take a more precise stance on the origins of the PII word. Lubotsky seems to assume that the word is not a substrate word even though it has no certain cognates elsewhere in Indo-European.

Mallory & Adams (2006: 149) state that Proto-Indo-European had two variants of the same insect-word (denoting ‘fly’ or ‘gnat?’), *\*moko-* and *\*moko-*, and according to them, the latter depalatalized variant gave Indo-Iranian *\*makš-*, as well as Lithuanian *makatas* ‘gnat’. Also EWAia (II: 335 s.v. *másaka-*) mentions this possibility but does not discuss the matter further (“... daher ist auch der Anschluss von *mákš-* (iran. *maxš-*) versucht worden”). Since the exact path of word formation is totally obscure here (it receives no comment from Mallory & Adams), one cannot support this kind of explanation. Derksen (2015: 306, s.v. *mašalas*) notes that the relationship of Lith *mašalas*, Latv *masals* and OI *másaka-* is not entirely clear. He supports the reconstruction of a PIE root *\*mek-* and assumes that Lith *makatas* ‘midge’ can be explained through depalatalization and that Pali *makasa* ‘horsefly, mosquito’ and MP *magas* ‘fly’ can be explained through metathesis (here he refers to KEWA II: 603). It is interesting that the Lithuanian word *makatas* is not mentioned by Matasović (2005) in his account of centum elements in Balto-Slavic. In any case the Baltic form with *k* has to go back to a relatively early stage, as it is reflected by the Finnish loanword *mäkärä* ‘blackfly, buffalo gnat’ (Junttila 2015: 122, 126).

In any case it remains unclear how Pre-PII *\*mekš-* would have been derived from a root *\*mek-* (one is forced to assume some irregular depalatalization in this case). Derksen does not discuss the *\*mákš-* word group here at all. The existence of a root *\*mek-* is on rather shaky grounds, and one cannot find such an entry in LIV<sup>2</sup>.

Schindler (1972: 36) notes that the word *\*mákš-* could theoretically contain the same *s-* suffix as some other Indo-European animal-words, such as *\*ghans-* ‘goose’.

The old explanation by Thieme (1965: 224, footnote 1) for the etymology of the Indo-Iranian word is mentioned by Schindler (1972) as a possibility. Thieme notes that there is no reason to assume a meaning ‘bee’ for the Vedic *mákš-*, *mákšikā* words, but that ‘fly’ is

the only meaning found. The same meaning prevails in Avestan. Thieme also states that this ‘fly’ word is derived from an Indo-Iranian root *\*makš-* ‘beschmutzen (liata, sotkea)’ which is not found in Old Indo-Aryan (nor Old Iranian), but which according to Thieme is reflected in Pali *makkh-* id.

Parpola (1999, 2005) has attempted to explain the Indo-Iranian word as an old compound: he considers the word a compound of the words *\*med<sup>h</sup>u-* ‘honey’ and *\*k<sup>w</sup>ey-* (> PII *\*č(/k)ay-* > OI, Av *ci-* ‘sammeln, schichten’, Werba 1997: 181–182, No. 55) ‘to gather’, so *\*mekši-* ‘bee’ would be a ‘gatherer of honey’. Semantically the explanation is plausible. However, it is dubious that in the attested early Indo-Iranian languages, the word means usually ‘fly’, rather than ‘bee’; both meanings are attested in both Avestan and Old Indic, but the meaning ‘fly’ prevails. It is, of course, possible that the initial meaning was ‘bee’ and it has been secondarily ousted by ‘fly’ in both Iranian and Indo-Aryan, but it is not obvious that ‘bee’ was the initial meaning of the Indo-Iranian word, and this makes Parpola’s explanation dubious. The word could have acquired a more general meaning of ‘flying insect’ in Proto-Indo-Iranian. Divergent meanings in words denoting insects can be found also in the Finno-Saamic *\*mača*, a borrowing from an Iranian source meaning ‘mosquito’, but denoting ‘worm’ in Finnic and ‘moth’ in Saami (Kallio 2000; Koivulehto 1999b). Interestingly, also Pashto has borrowed the Indo-Aryan word for ‘fly’ in the meaning of both ‘bee’ and ‘fly’: Pashto *məč* M ‘fly’, *məčəy* ‘bee’ ← IA *makšā-* (Morgenstierne 2003: 48, s.v. *məč*).

The development according to Parpola is through the following phases: *\*med<sup>h</sup>u-* + *\*kw(e)i-* > *\*met-kwi-* > *\*mekši-*. Here the *\*š* would be the result of the so-called thorn cluster (compare *\*d<sup>h</sup>g<sup>h</sup>m-* > OI *kšám-* ‘earth’). The explanation is thus plausible, but it is difficult to prove that the word really derives from these elements. It is also uncertain whether the PII verb really reflects a PIE root *\*k<sup>w</sup>ey-* (Werba mentions the PIE etymology with a question mark). LIV<sup>2</sup> (378–379) does present such a root with reflexes in Greek (ποιέω ‘mache, schäffe, bringe hervor; to do, make, produce (also of poetry), act’, in mediopassive also ‘to choose, deem, appraise’) and Slavic (Ukr. *kojú, kojíty* ‘anzetteln, (etwas böses) bereiten’) besides Indo-Iranian. Beekes (2010: 1216) accepts the PIE etymology of the Greek and Indo-Iranian verbs.

If Parpola is right and the word really can be explained like this, the word cannot serve as evidence for the early occurrence of the RUKI rule, as has been argued. However, phonetically Parpola’s idea is problematic as he does not explain at all why *-u* would have been syncopated in the middle of the word. In the light of the recent argument by Lipp (2009) that the clusters of dental and velar stops did not develop into *\*kš* but *\*tš, \*d<sup>h</sup>ž* in Proto-Indo-Iranian, it becomes obvious that there is little reason to support Parpola’s etymology for the Indo-Iranian word.

Even if the etymological background of the Indo-Iranian word *\*mekši-* remains uncertain, the direction of borrowing is still rather from Indo-Iranian to Uralic than the other way around. As the central apiculture word *\*meti* ‘honey’ is certainly an Indo-European (? Pre-Indo-Iranian) borrowing (from *\*med<sup>h</sup>u-*), it is logical to consider the word for ‘bee’ also an Indo-Iranian loanword. Another Indo-Iranian loan-word in the field of apiculture is Uralic *\*síšta* ‘beeswax’ ← PII *\*síšta*, cf. OI *madhu-sišta-* (Carpelan & Parpola 2010), although this can also be a slightly later borrowing. The direction of borrowing is further supported by the fact that there are no words related to apiculture that can be reconstructed for Proto-Uralic, and the word for ‘bee’ thus looks like a cultural innovation

in Uralic – but not in Indo-Iranian, where other words related to honey and its production were inherited from Indo-European.

It is interesting that the word was borrowed into Uralic as an *\*i*-stem, as these are much rarer than *\*a*-stems among the early loans. This can be perhaps explained by viewing the Uralic word as borrowed from an Indo-Iranian form leading to Avestan *maxšī-* (so Pre-II *\*mékšī-* or the like), and in this case *\*-i* would be phonetically an obvious sound substitution. It is also possible that the root noun *\*mékš-* was simply acquired as an *\*i*-stem. The prevalence of Uralic *\*a*-stems among the loanwords can result from the fact that many of the loans are *a*-stems in Indo-Iranian.

An Indo-Iranian etymology

**Ma *mež*, *miž* ‘wool’; Ko *mež* ‘ram’ (< PP *\*mež*)**

< ? PU *\*mešV*

← Iranian *\*mēš(a)-* < PII *\*maiša-* M, *\*maišīH-* F > Av *maēša-* ‘ram’, *maēšī-* ‘sheep’, MP *mēš-* ‘sheep’; OI *mēša-* ‘ram’, *mēšī-* ‘sheep’ (EWAia II s.v. *mēša-*)

(KESK: 171; Joki 1973: 285; UEW s.v. *\*meš3*; Rédei 1986c: 72; Lushnikova 1990: 232; Bereczki 1992: 92–93; 2013 s.v. *mež*)

The Indo-Iranian origin of the Permic and Mari words is convincing, but there are various views about the exact donor form and the relationship of the words in these two branches of Uralic.

The Permic vocalism probably reflects Middle Iranian *ē* that has developed from the earlier diphthong *\*ay*. Lushnikova considers the borrowing to be Scytho-Sarmatian. However, in the Alanic name Ματοης, which presumably is derived from this Indo-Iranian word, the diphthong is still retained (Zgusta 1955: 113), so it seems uncertain whether the Permic word can be borrowed from a form of Alanic.

Mari *mež*, *miž* ‘wool’ is borrowed from Permic according to Rédei (1986c), but Joki (1973), UEW (s.v. *\*mišV*) and Bereczki (1992: 92–93; 2013 s.v. *mež*) argue that the word can have been separately borrowed into Mari and Permic. The Mari word is probably also a late loan in any case, as it does not show any trace of the diphthong *\*ai*. Metsäranta (personal communication) notes that the Permic word should be borrowed from a disyllabic form where *\*s* is between vowels, as the voiced *ž* can only be explained through this. Based on vowel correspondences, Aikio (2014c: 139) argues that the Mari word has to be borrowed from Permic.

An Indo-Iranian etymology (Mari a loan from Permic)

**Md E *mirǵ'e*, M *mirǵ'ε* 'man, human'; Ko *mort*, Ud *murt id.* < PP *\*mǝrt***

**< PU *\*mertä***

**← PII *\*mr̥tá-* > OI *mr̥tá-* 'dead', YAv *mər̥ta-* *id.*, OP <mrt-> *id.* (< PIE *\*mr̥tó-*)**

**or**

**← Pre-II/PIE *\*mérto-* ? > OI *márta-* 'mortal, human', OP *\*marta-* (attested only indirectly in Elamite <pír-šá-mar-da>) ? OAv *mar̥ta*<sup>25</sup> < PIE root *\*mer-* 'to die, to disappear', PII *\*mar-* > OI *mr̥-* 'to die', OP *mar-*, Av *mar-* *id.* (EWAia II s.v. *márta-*; LIV<sup>2</sup>: 488–491, s.v. *\*mer-*; Werba 1997: s.v. *mr̥-* 'sterben'; NIL s.v. *\*mer-* 489–491)**

(Setälä 1928: 17; Munkácsi 1901: 455; FUV: 130; KESK: 175; Joki 1973: 281; Katz 1983: 175; 1985: 147 2003: 123; Gamkrelidze & Ivanov 1984: 926, footnote 1; UEW s.v. *mertä*; Rédei 1986c; Lushnikova 1990: 39, 235–236; Koivulehto 1999a: 221; 2016: 223)

See also *\*martas*

This is an old and widely accepted etymology, and there is no reason to doubt the Indo-Iranian origin of this Uralic word. It is also well known that *\*martas* > Fi. *marras* (see the entry above) is a parallel loan from the same Indo-Iranian word group (SSA s.v. *marras*), despite confusing claims by earlier sources such as FUV, where these two words are lumped together. However, some important details of the etymology of *\*mertä* are obscure and require a rather detailed analysis. The central problem is whether the Uralic word has been borrowed from a predecessor of the OI noun *márta-* 'mortal, human' or that of the verbal adjective *mr̥tá-* 'dead'. Although this problem might seem trivial, it is extremely relevant for the dating of the borrowing and it can also offer indirect evidence of Indo-European and Indo-Iranian word formation.

Koivulehto (1999c) argues that the word is borrowed from the Indo-Iranian zero-grade verbal adjective *\*mr̥tá-* with Uralic *\*er* as the substitution for the Indo-Iranian syllabic resonant. A similar argument was already presented by Lushnikova (1990). This would be phonetically perfectly fitting as the original of the Uralic, but for this Indo-Iranian adjective no meaning of 'human' is attested (the word explicitly denotes 'dead' in the attested ancient Indo-Iranian languages), making this explanation slightly uncertain. Although it is not noted by Koivulehto, already the Indo-European pre-form *\*mr̥tó-* could also be an equally good original. This verbal adjective is reflected by the Armenian *o*-stem noun *mard*, which denotes precisely 'human', and this meaning has clearly developed from earlier meaning 'mortal' (NIL). Also in Greek the cognate of *mr̥tá-*, Gr *βροτός* means 'mortal'. If the PIE verbal adjective already had the meaning 'mortal', there would be no phonological or semantic problems in deriving the Uralic word from this Proto-Indo-

---

<sup>25</sup> As noted by both EWAia and NIL, the Avestan form might be a later, Avestan-internal formation because it points to a pre-form *\*martás* (with different accentuation than the Vedic word). EWAia notes that the Old Persian word can be derived from *\*mártas*, so the Avestan form is clearly isolated. This does not have effect on the evaluation of the Uralic etymology, and the analysis of the Avestan form has to be left to researchers of Old Iranian philology.

European (or Core Indo-European/Late Proto-Indo-European) form. Lushnikova glosses the PII form *\*mr̥tá-* as ‘mortal, human’ but does not explain this choice in more detail.

Proto-Indo-Iranian also had the adjective *\*a-mr̥ta-* ‘immortal’, reflected by Vedic *amr̥ta-* and Young Avestan *aməš̥’a-*. Young Avestan has also *məš̥a-* ‘dead’ which continues a similarly accented form *\*mr̥ta-* that is attested in these compounds. As the forms with the negative prefix clearly denote ‘immortal’, one could assume that the verbal-adjective could have denoted ‘mortal’ too, even if it is not attested in this meaning.

The earlier explanation by Katz (1983b) is that the word is borrowed from Pre-Indo-Iranian *\*mérto-*. This explanation supposes that the Indo-Iranian noun *\*márta-s* ‘human’ is a substantivized vrddhi-derivative of *\*mr̥ta*. This noun indeed has the meaning ‘mortal, human’. According to Katz, the Uralic word has been borrowed from the Pre-Indo-Iranian stage of this word, *\*mérto-* (~ *\*mérta-*). Tremblay (1998) and Steer (2015) likewise consider PII *márta-* a reflex of *\*mérto-*, which is a substantivizing vrddhi-derivative from *\*mr̥tá-* (in a similar way as PG *\*ken̥pa-* ‘child’ < PIE *\*géh<sub>1</sub>to-* is a derivative from verbal adjective *\*géh<sub>1</sub>to-* ‘born’). The explanation of Katz is also accepted by Rédei (1986c), but the UEW mentions both possibilities. The Pre-II form in the UEW has been reconstructed as *\*mérta-*, with the underlying (but not expressed) assumption that *\*o* has already changed to *\*a* but *\*e* has not. The UEW mentions that the Uralic (Finno-Permic in the UEW’s terminology) word is borrowed from a nominative form. This is, indeed, probable, but the “Pre-II” forms UEW cite are anachronistic: *\*mr̥taḥ* and *\*mertaḥ* (obviously the word-final *-s* had not yet become visarga at the point when PIE *\*e* was still retained). The UEW presents the data in a way that gives a reader who has no background in Indo-Iranian linguistics a misleading picture of the detailed problems that are involved in the precise etymology of the Uralic word and the Indo-Iranian word family.

*\*merto-* → *\*mertä* fits well into the system of front-vocalic *\*e* loans, such as *\*kekro-* → *\*kekrä*; it seems that Pre-II words with *\*e - o-* vocalism usually became *\*ä*-stems in Uralic.

Although the explanation by Katz is unproblematic *per se*, there are also reasons to consider *\*mártas* a reflex of earlier *\*mórtos*: this is the view of Koivulehto (1999a). Koivulehto mentions the old idea that Greek *μορτος* is an exact cognate<sup>26</sup> of PII *\*mártas*. Both possible explanations of Indo-Iranian *\*mártas* have received support in works that deal with Indo-European etymology. LIV<sup>2</sup> mentions both possibilities for the origin of PII *\*martas*. LIV refers to Schwyzler (1939) who has argued that *μορτος* continues the zero-grade form with different (Aeolic) vocalization of PIE *\*r̥*. Beekes (2010) considers the Greek and Indo-Iranian nouns to be reflexes of PIE *\*mortos* and has argued that the word cannot show Aeolic vocalization because the form is attested in Doric names. Also Mallory & Adams (2006: 199) support the view that both the Greek and Indo-Iranian forms continue PIE *\*mórtos*. Kulikov (2017) also reconstructs *\*mártas* as *\*mórtos*. NIL (489–491) considers PIE *\*mérτος* as the most probable pre-form of the Indo-Iranian word and specifically mentions Katz’s etymology as evidence, although the possible Greek evidence of earlier *\*mórtos* is also mentioned.

Schwyzler (1939) also mentions a different view by Schulze (1933) of the *o*-vocalism in this word: Schulze reconstructs the PIE word as *\*mrtvós* and argues that the Greek *o* in *βρότος* results from PIE *\*w*; in reality this *w* cannot be reconstructed for the PIE form and

<sup>26</sup> The Greek word is attested only in a gloss and its accentuation is unclear

only appears in the “Northwest-Indo-European” pre-form of the Italo-Celtic and Balto-Slavic words for the ‘dead’, cf. Latin *mortuus*, OCS *mr̥tvъ* where the *w* results from analogical influence from PIE *\*g<sup>w</sup>ih<sub>3</sub>-wo-* ‘living’.

In his discussion of Indo-Iranian loanwords in Uralic, Dolgopolsky (1989: 20) mentions that *\*mertä* is borrowed from the zero-grade form, but he offers no arguments for this and does not refer to Katz’s article about the topic, only to Joki (1973: 281) and FUV (1955: 130).

Recently, Höfler (2018) has argued that OI *márta-* is an example of “*o*-vrddhi”, a process of substantivization where the *\*o*-grade is used instead of *\*e*-grade. This noun fits the system of word formation that is attested in a number of *o*-grade forms in Greek and Indo-Iranian. Höfler mentions the Greek form *μῆρτος* as a cognate to the Indic word.

Phonologically, both *\*mértos* and *\*mrtás* (or *mrtós*) would be plausible donor forms for the Uralic words. *\*er* appears as a substitution for PII *\*r* in some loans. At least Permic *gort*, *gurt* is an apparent parallel: the Permic words could reflect Uralic/Pre-Permic *\*kertä*, and Koivulehto (1999a) assumes that this is borrowed from PII *\*gr̥dha-* with the substitution *\*er* ← *\*r̥*. The word-initial voiced *g* is slightly problematic as it usually does not appear in inherited Pre-Permic vocabulary, and already Koivulehto mentions that the word can also be a later borrowing, acquired at a time when the Permic vowel system had already changed considerably, and the Iranian syllabic resonant or its later vocalic reflex could have been substituted by a Proto-Permic vowel that later led to Komi *o* and Udmurt *u* (in Sammallahti’s 1988 system this vowel would be Proto-Permic *\*o*, which would be a phonetically fitting substitution of the vowel in a later Iranian form that could reflect PI *\*gr̥da-*). See the entry *\*kertä* for more detailed discussion. Gamkrelidze & Ivanov (1984: 926, footnote 1) have assumed that the Permic word is a later loan, although they do not go into detail when arguing this.

A similar explanation could apply for Permic *mort*, *murt* too, as the word does not have to be an extremely old borrowing but the Proto-Permic vowel could substitute for the vowel *a* that is found in later Iranian forms such as MP *mard* or Parthian *mrd* [*? mard*] which continue PI *\*marta-* (this is suggested tentatively by Metsäranta, p.c.). However, because the Permic and Mordvin words regularly correspond to each other and reflect a pre-form *\*mertä*, it is probable that the word indeed is an old borrowing. The apparent similarity of *mort*, *murt* and *gort*, *gurt* need not mean that the words belong to the same layer of loanwords.

A form *\*mórtos* that could be reconstructed based on OI *márta-* and Greek *μῆρτος* would be a more problematic donor form for Uralic *\*mertä*, as there are no known parallels to the substitution *\*o* → *\*e*, and this would be phonetically dubious.

To sum up, the exact donor form remains inconclusive, as there are arguments that support both of the suggested forms. Because of this uncertainty, this word cannot be used as major evidence for various substitutions of the PII syllabic resonant.

As an excursus it can be mentioned here that Rédei (1988b) has considered the ethnonym *Mordva* ‘Mordovian’ a separate borrowing from the same Indo-Iranian source. The Indo-Iranian etymology of this ethnonym is also mentioned by Carpelan & Parpola (2001: 111–112) and Grünthal (2012: 302). Rédei sees Russian *Mordva* as a loan from originally Mordvinic word *\*morda* (the Russian *-v-* can be clearly secondary, and Rédei provides parallels for this development), and Rédei derives *\*morda* from Pre-Mordvinic *\*murta*; he argues that this is a borrowing from the form *\*mrtā-*, and that here vocalic *\*r̥*

was substituted by \**ur*. This substitution is found in many other tentative early Indo-European loans, but not in many Indo-Iranian loans. However, because of the labial consonant environment, this would be a natural substitution of \**r*<sub>l</sub> in this environment. Further assessment of this etymology is very difficult, because the Mordvin word \**morda* is not directly attested.

Rédei critically comments on another earlier etymology for this ethnonym too, namely the possibility to derive the ethnonym *Mordva* from an Iranian compound \**mr̥ta-χwa-* ‘man-eater’, connecting this hypothetical word with Greek ἀνθροφάγοι ‘man-eaters’, an ethnonym mentioned in ancient Greek sources. Rédei dismisses the etymology because there is no proof of a previous existence of this Iranian compound (it is not attested in any Iranian language), and because the *-v-* of the ethnonym *Mordva* is a secondary development (in Rédei’s view).

The connection with this hypothetical Iranian term is interesting, as the ethnonym *Sarmatian* has also been derived from a compound ‘man-slayer = hero-slayer’ \**Saura-mar-* > \**Sūra-mar-* according to Schmeja (1974), and their ethnogenesis is connected with the Amazons, Οἰόρπαιτα ‘men-slayers’ in the classical Greek sources. However, it is doubtful whether the Mordvians can be connected here, and thus the relationship of their ethnonym to the Iranian ‘man-slayer’ and ‘man-eater’ words remains hypothetical though interesting.

There is no compelling reason to doubt Rédei’s etymology for \**morda*, but one has to admit that the existence of this Proto-Mordvin word is hypothetical, as the word has not been attested in either Mordvin language. In neither Erzya nor Moksha is there a word denoting ‘Mordvian’, although Rédei plausibly argues that the word could have become lost after the Mordvians were separated into two ‘tribes’, Erzyans and Mokshans. It is however not certain that the ethnonym derives from one that was used by the Mordvians themselves; cf. the ethnonyms *Finnish* and *Lappish* or *Hungarian*, which have never been used by the *suomalaiset*, *sápmelaččat* and *magyarok*, respectively. It is probably true that *Mordva* is in one way or another derived from \**mr̥ta-*, but it is difficult to prove that it happened the way Rédei describes.

An Indo-Iranian etymology

**Fi *mesi* ‘nectar’ (cognates in all Finnic languages); Md E *med’*, *mäd’*, M *med’*; Ko *ma*, Ud *mu* (< PP \**ma*); Hu *méz* ‘honey’**

< PU \**meti* ‘honey; nectar’ (UEW s.v. \**mete*; Sammallahti 1988: 545)

← Pre-II or PIE \**médʰu-* ‘mead, honeywine’ > PII \**mádʰu-* > OI *mádhu-* ‘sweet drink, honey, sweetness, soma, milk’, Av *maδu-* ‘Beerenwein’; other IE: Gr. μέθυ ‘wine’, English *mead*, OCS *medъ* ‘honey, mead’, Tocharian B *mit* ‘honey’ < PIE/core-IE \**médʰu-* (EWAia II s.v. *mádhu*; Dunkel 1995: 12; SIL: 306; NIL: 467–468; Kölligan 2018: 2275)

(FUV: 132; Joki 1973: 283–285; Katz 1985: 280; 2003: 224; UEW s.v. \**mete*; Rédei 1986c; Lushnikova 1990: 230; Abondolo 1996: 87; SSA II s.v. *mesi*; Koivulehto 2001b; Parpola 2012: 161; 2015: 66)

In current research, this word is usually considered to be an Indo-Iranian borrowing, but there are actually no phonological arguments to support this, as the original *\*med<sup>h</sup>u-* could equally well be Proto-Indo-European, and this actually speaks against the criteria outlined in the Introduction to include this word in our research material. One of the main arguments to consider this word Pre-Indo-Iranian is that other apiculture words were also borrowed from Indo-Iranian into Uralic, and the word *\*mekši*, which is clearly a Pre-Indo-Iranian borrowing because of retained *\*e* but RUKI *\*š*, could logically belong to the same layer of loanwords as *\*meti* (see the entry *\*mekši*).

Semantically an early form of Indo-Iranian would be more fitting than PIE donor. It is possible that the PIE word denoted ‘mead’, and the meaning ‘honey’ developed later in Indic, Slavic and Tocharian (Kölligan 2018: 2275). In Avestan, the meaning of ‘honey’ is not found, however.

On the other hand, according to Dunkel (1995: 12), ‘honey’ is the original meaning, and the meaning ‘wine’ in Greek and Luvian is secondary (developed through the meaning ‘mead’ when the use of wine in rituals ousted that of mead, which had ousted *soma*). The meaning of ‘honey’ can be reconstructed for Proto-Indo-Iranian in any case, but various views remain for the reconstructed Proto-Indo-European meaning.

PU *\*meti* is usually considered a loanword even by the supporters of the Nostratic and Indo-Uralic hypotheses, who consider many potential early IE loans to rather be cognates between Indo-European and Uralic, such as Dolgopolsky (1989: 20). In the light of this, it is interesting that Kümmel (2009: 9) lists this as one of the words which he considers ambiguously either loans from PIE to PU or cognates between the two language families.

It is difficult to present further criteria to argument for or against the Indo-Iranian origin of this word. It’s Indo-European origin is, in any case, clear and settled.

Gamkrelidze and Ivanov (1984) argue that the word is borrowed from Tocharian. This is not a very convincing idea for areal and chronological reasons, and also phonetically there is nothing specifically pointing to a Tocharian source.

The Uralic forms are regular cognates. Mari *müj*, *mü* has sometimes been connected tentatively to PU *\*meti*. Bereczki (2013. s.v. *mü*) notes that the UEW connects the Mari word also to *\*mäkV* ‘honey’, a hypothetical PU word stem reconstructed on the basis of Mari and Ob-Ugric forms Mansi *\*mäy*, Khanty *\*mäy*. This old explanation is found also in Abondolo (1996). Bereczki does not support this explanation but does not explicitly refute it. Aikio (2014: 133) has noted that the Mari word is borrowed from Permic. This is a convincing explanation for the Mari word, and the existence of Uralic *\*mäkV* can be rejected.

The Saami forms like SaN *miehta* are loans from Finnic (SSA s.v. *mesi*; Holopainen 2018b: 162). Aikio (p.c.) has noted that Lule Saami *mieda* is rather borrowed from Scandinavian *\*meðu-*.

Non-initial syllable *\*-i* reflects PIE/PII *\*-u* here, and same kind of development is found in other loans from Indo-European *\*u*-stems (Holopainen, Kuokkala & Junttila 2017: 131).

An Indo-Iranian etymology

**Mansi South *mūj*, West *muj*, North *mūj* (< PMs \**mūj* (-ə)) ‘Gast’; Khanty North *maj*, East *maj*, South *moj* (< PKh \**māj*) ‘Hochzeit’**

< **POUg** \**mēja* (Zhivlov 2006: 164)

← **PII or PI** \**máy(H)as-*, **OI** \**máyas-* ‘enjoyment, pleasure’, **Av** \**maiiah-* ‘mating’, **Alanic name** Μαυακος < **PIE** \**me(y)(H)es-* (EWAia II: 315–316, s.v. *máyas-*; Zgusta 1955: 113)

(Blažek 1990: 40; Helimski 1997b: 121)

There are no phonological problems in supposing an Indo-Iranian loan, but the semantic connection is not straightforward. In the Ob-Ugric languages there is no meaning of ‘enjoyment’ or ‘mating’, and the Ob-Ugric meaning ‘guest, visit’ is quite far from that of the Indo-Iranian words. One could suppose a semantic development ‘mating’ > ‘wedding’ > ‘guest’, but this remains speculative.

The Ob-Ugric vocalism could reflect an earlier (PU) \**o*-*a*-stem (the vowel combination is similar as in the reflexes of the well-known words \**woča* and \**koδwa*; Aikio 2015b: 31 notes that PKh *ā* often reflects PU \**o*-*a*-stems), meaning that the word is probably an “*o* loan”. Here the Indo-Iranian word reflects PIE \**e*, so the word cannot be borrowed from a word with PII \**o*. However, also an \**a*-*a*-stem would be possible. It seems uncertain whether the word has been borrowed into a Pre-Ob-Ugric stage with \**a* or \**o* or into Proto-Ob-Ugric with \**ǣ*.

If the etymology is correct, the word is probably an old loan (Pre-Ob-Ugric), even though its distribution is limited to Khanty and Mansi. The words for ‘wedding’ are usually loanwords in the Uralic languages (Hungarian *esküvő* is derived from *eskü* ‘vows’, a Turkic loan [EWUng s.v.], and Finnish *häät* is borrowed from Middle Low German [Koivulehto 2016: 306]).

? Indo-Iranian etymology

**Mansi East *mex*, North *mēŋk*, West *meŋk* ‘devil’ (< PMs \**mīŋ<sup>w</sup>k<sup>w</sup>* (~ -ə)); Khanty East *mōŋk*, South *meŋk*, North *meŋk* etc. ‘forest spirit’ (< PKh \**māŋk*)**

< **POug** \**mōŋki* (Zhivlov 2006: 174–175; Honti 1982: 165, No. 402)

← **Iranian, compare Middle Persian** *mēnōg* ‘spirit’ < **PII** \**manyu-ka-* (EWAia II s.v. *manyú-*)

(Munkácsi 1901: 457–458; Korenchy 1972: 60–61 No. 14; Lushnikova 1990: 231–232; Helimski 1997b: 121; Sipos 2002: 32–33; Blažek 2005: 166)

The etymology is convincing: the Khanty and Mansi word are transparent Indo-Iranian borrowings, and the phonological shape of the Ob-Ugric words points clearly to a relatively Middle Iranian origin (the vowels are clear indications of a relatively late Iranian form). Korenchy reconstructs the donor form as “Middle Iranian” \**mēnūk* ‘spirit’. According to Lushnikova (1990: 231–232), the Ob-Ugric word is a Scytho-Sarmatian loanword. Zhivlov

reconstructs \*ǰ already for Proto-Ob-Ugric. Korenchy also accepts Steinitz's conclusion that Ob-Ugric \*ö is retained in the archaic East Khanty forms, but still she reconstructs \**mēŋkV* for some reason. Honti (1982) reconstructed \**mīŋkV*. This reconstruction tells very little of the actual quality of the vowel.

While the origin and the relatively late date of borrowing can be considered certain, there are various issues with the relation of the Khanty and the Mansi words. Khanty and Mansi probably were independent languages by the beginning of the “Middle Iranian” period (after 500 BC), and thus the word is probably a separate borrowing from Iranian to Khanty and Mansi. If takes a closer look at the vowel correspondences between the Mansi and Khanty words, one can note that this correspondence is not found in many Ob-Ugric words, and in only a few words that have a reliable etymology going back to reconstruction stages beyond Ob-Ugric. However, at least the Ob-Ugric reflexes of Proto-Uralic \**kälä-* ‘to wade’ (Aikio 2015b: 63; UEW s.v. \**kälä-*) have a similar vowel combination, so it seems possible to assume that Khanty \**māŋk* and Mansi \**mīŋ<sup>wk</sup>* are inherited words from Proto-Ob-Ugric. But because of the clearly Middle Iranian appearance of the word it is best to consider these two words independent borrowings from Iranian at the time when Khanty and Mansi were separate languages, even if they were spoken in a close-knit Sprachbund.

An Indo-Iranian etymology (parallel loans)

**Mansi West *mūl-*, South *mōl-* ‘to win’ (< ? PMS \**mul*’V-)**

← ? Iranian \**mūl-*, compare Ossetic *æmbūlyn* ‘to play, to win’, *mohun: muld* ‘to win’ (Abaev 1958–1989 I: 140, II: 128–129)

(Bouda 1933: 162–163; Korenchy 1972: 59–60 No. 12; Helimski 1997b: 121)

The unknown background of the Ossetic word, as well as the scarce distribution of the Mansi word make this etymology unlikely. However, it has to be noted that the distribution is not as scarce as Korenchy implies (she only mentions the P forms). Kannisto et al. 2013 give a distribution TJ, TČ, P, VN, VS, Munkácsi gives N, P, T. Korenchy reconstructs a Proto-Mansi form \**mul*’V-. This is probably correct, as the attested Mansi forms could reflect earlier \**u*.

Regarding the Iranian reconstruction, Korenchy gives a Middle Iranian form \**mūl-*, but the Ossetic forms are not regular reflexes of earlier (Proto-Iranian) \**u* (see Cheung 2002: 124–127). Abaev mentions possible Slavic cognates of the Ossetic verb: Russian *мылить*, Czech *mýlit*, Polish *mylić* ‘obmanyvat’, but these words are not mentioned at all by Derksen (2008) and it is doubtful that they are really cognates of the Ossetic verb. The Ossetic word remains isolated, and it seems unlikely that the Mansi word is borrowed from it. Also the Mansi *l* would be an irregular reflection of Iranian \**l*. Semantically the etymology would be unproblematic. Helimski (1997b) considers the etymology as problematic.

Not an Indo-Iranian etymology

**Proto-Mordvin \*morda**

< Pre-Md \*murta

← \*mr̥ta-

(Rédei 1988b: 178–79)

See \*mertä

Fi *mykkä* ‘dumb’ (cognatesin Izh, Ka, Va), verb *mykistyä* ‘to be muted’, *mykistää* ‘to mute’; SaS *makkalit* ‘to stop in one’s speech’;? Ma W *māk* (*manaš ogəl*) ‘to stop speaking’; Ko *miktav-* ‘änkyttää, mumista’, Ud *mik-mak veraške* ‘to speak unclearly’ (< PP \*mik-) (Saarikivi 2007: 333–335)

< PU \*mükkä

← PI \*mūka- ‘dumb’ > OI *mūka-* ‘stumm, lautlos, schweigend’ < ? PIE \*mu-, cf. Gr *μυκός*, Lat *mūtus* ‘stumm’ (EWAia II: 365, s.v. *mū́ ka-*)

(Parpola 2010: 312)

Parpola’s etymology is semantically convincing. The Finno-Saamic and Indo-Aryan words resemble each other to such an extent that it would be peculiar to consider this purely coincidental. However, Parpola’s etymology is problematic in the sense that it supposes a variation of \*u- and \*ü forms at a very early stage. It is true that in Finnic many words have *u ~ ü* pairs, such as *tuhma ~ tyhmä*, a word which is a Baltic borrowing, and this variation appears in several other borrowings from Baltic. However, there are no reliable examples of such variation at such an early proto-language level that both Finnic and Saami cognates could be derived from an alternating front-vocalic form.

It is of course possible to suppose a substitution U \*ü ← PII \*ū, as there of course was no opposition \*u : ü in Indo-Iranian. This assumption could easily save the etymology. Another possibility would be assume that Finnic *mykkä* is indeed a later variant of early (unattested) \*mukka, and the Saami verb is borrowed from this word. However, the Saami borrowing would have to be very early, making this explanation complicated.

The Mari word has to reflect earlier \*u, and the Permic words can reflect either \*u or \*ü (Metsäranta, personal communication).

? Indo-Iranian etymology

**Fi *myrkky* ‘venom’, Est *mürk* (< PFi \* *mürkkü*)**

← ? PII \**marká-*, OI *marká-*; Av *mahrka-* ‘death’<sup>27</sup>, Ossetic *marg* ‘poison’, PII root \**marč-* ‘to injure, hurt’ (EWAia II s.v. *MARC*)

(FUV: 133; Abaev 1958–1989 II: 72–73 s.v. *marg*)

It is an old idea that this word is an Iranian loan, but the etymology does not appear in newer sources on the topic. It is not considered convincing by SSA and not discussed in the articles of Koivulehto or Parpola. Also EES does not mention the etymology. FUV still considered Finnish *myrkky* and Hungarian *méreg* ‘venom’ cognates, although the Hungarian word has to be a separate loan, as Hungarian \*-*rg-* cannot correspond to Finnic \*-*rkk-*. It is usually assumed that *méreg* is a later borrowing from Alanic \**marg*, whence Ossetic *marg* (Abaev; WOT; Joki 1973: 282–283). UEW (s.v. *mirkkV*) considers the Hungarian and Finnic words possible cognates, but does not support the Indo-Iranian etymology for this Uralic word (the possibility that Hungarian *méreg* is a later loan is mentioned by the UEW).

As the cluster \*-*rkk-* cannot be reconstructed for earlier proto-languages preceding Proto-Finnic, the early origin of this word does not look very plausible. Semantically the word could be derived especially from Alanic, as the meaning ‘venom’ is attested in Ossetic *marg*, but this cannot be reconstructed for Proto-Indo-Iranian or Proto-Iranian. The Avestan and Old Indo-Aryan words mean ‘death’. Also the Finnic *ü* is difficult to explain as reflecting Iranian \**a* (no parallels exist). If the word was borrowed from an Indo-Iranian form with \*-*r-*, then \*-*ür-* would be a phonetically more plausible substitution. From the verbal root \**marč-*, zero-grade formations are of course attested (such as OI *mṛktá-* ‘beschädigt’ and *mṛkṇá-* ‘verletzt’), but these forms would not be semantically very good originals. On the other hand, it has been noted by de Smit (ms.) that if the original vowel was \**e*, the \**ü* could be explained through the sound change \**e* > \**ü* that is attested in Finnic for some words. Although de Smit mentions this change as “not unknown but not exactly regular either”, Aikio (personal communication) has noted that this sound change is actually regular before \**ü* in Finnic (no examples of the vowel combination \**e-ü* can be found in the Proto-Finnic vocabulary; see also \**sentü-* for a similar case). If we assumed \**e* substitution then, the old idea of an Indo-Iranian etymology would not be completely implausible, but as we have no examples of Alanic \**a* being substituted by Pre-Finnic \**e*, this does not help much.

Not an Indo-Iranian etymology

---

<sup>27</sup> Lubotsky (1988: 78) notes that the OI and Av words point to different accentuation.

**i naaras, dial naara, nairas ‘female animal’** (SSA II s.v. *naaras*)

< ? Pre-Fi *\*naras* ~ *\*nayVras*

← PII (or PI) *\*HnāriH-* > OI *nārī-* ‘woman, wife’, Av *nāiri-* ‘woman’ (EWAia II: 19–20, s.v. *nār-*)

(T. Sköld 2002)

Finnish *naaras* has usually been considered a derivation of Uralic *\*ne-* ~ *\*näyi* ‘woman’, although the derivational process and even the whole reconstruction of the Uralic word ‘woman’ are unclear. According to SSA, the word *uros* ‘male’ contains the same suffix (the original form would have been *\*uras*), and the stem would be *\*u* (see the entry on *úr* for more details on the word *uros*). The word *naaras* might contain the same suffix, being analogically formed on the basis of *uros* (this has been suggested by Ante Aikio, p.c.). From this point of view the Indo-Iranian etymology by T. Sköld is an interesting possibility, although this etymology also has its problems.

Semantically the etymology would be convincing. However, formally it is difficult to explain the Finnic word from this Indo-Iranian form. The Finnic long vowel has to be secondary, so it cannot reflect the Indo-Iranian long  $\bar{a}$ . The ending *-as* of the Finnic word fits the Indo-Iranian original poorly.

It is more probable that even though the precise derivational process is unclear, the word is somehow connected to the Uralic *\*näji* words, and the resemblance to the Indo-Iranian word is simply accidental.

As a side note, (Katz 1985: 208–209) has suggested that also the Uralic stem for woman, *\*näji* or *\*näyi* is borrowed from Indo-Iranian *\*gnā-*, and Koivulehto (1991) has assumed that the Uralic word is borrowed from Indo-European. Katz has suggested that the Uralic word is borrowed from Indo-Iranian. A rather similar idea is suggested by Koivulehto (1991), who however argues that the word is a Proto-Indo-European loanword. Although the idea that the Uralic word ‘woman’ is borrowed from an Indo-Iranian or Indo-European word with the same meaning is of course semantically plausible, the etymology is complicated because of the uncertain phonological reconstruction of the Uralic word. It is not at all certain that a noun *näji* can be reconstructed for Proto-Uralic, and some of the potential reflexes might actually continue words that have etymologically nothing to do with each other. Sammallahti (1988) considers the comparison uncertain. The idea that the word is from Indo-Iranian would be phonologically more plausible, as from *\*gnā-* the Uralic word-initial *n-* could be more easily explained than from a PIE cluster *\*g<sup>w</sup>n-* (there are no parallel examples so we do not know how this cluster would have been simplified, but this could have also resulted in Uralic *\*kuna-*).

The idea of a Proto-Indo-European etymology for the Uralic word has been criticized by Lindeman (1993: 311–312), mainly because Lindeman finds Koivulehto’s vowel substitution unconvincing (Koivulehto assumes that Uralic *\*ä* here reflects PIE *\*æ*, a supposed intermediary between *\*e* and later *\*a* that was caused by the *\*h<sub>2</sub>*).

Not an Indo-Iranian etymology

**Mansi South *nāt-*, West *nōt-*, North *nūt-* ‘helfen’ (< PM *\*nāt-*); Khanty South *nōt-*, North. *nūt-* id. (< PKh *\*nāt-*)**

< POUG *\*nātV-* ‘to help’ (Zhivlov 2006: 138)

← ? Iranian *\*nād-*, compare Avestan *nāidyah-* ‘der schwächere’, OI *nādh-* ‘sich in Not befinden, in Bedrängnis sein’ (EWAia II: 35, s.v. *nādh*)

(Korenchy 1972: 61–62, No. 15; Joki 1973: 293; Sipos 2002: 40; Helimski 1997b: 121; Katz 1985: 194; 2003)

The Ob-Ugric cognates are regular (they show the same vowel correspondences as *\*wāta* ‘wind’, see the entry), but the etymology is considered uncertain by Korenchy (and Helimski [1997: 121]): the palatalized *\*ń* is not an expected outcome of Iranian *\*n*, and the semantic connection is also fuzzy.

Instead of Korenchy’s etymology, one can also assume that the Ob-Ugric word is borrowed from another Indo-Iranian verb for help, namely OI *nāthá-* m. ‘refuge, help’ < *\*nēt-h<sub>2i</sub>-o-* (Kroonen 2013: 388). This does not solve the problem of the palatal nasal, but semantically this is a better option. Here the Indo-Iranian verbal root is reconstructed as *\*HnaH* ‘to help’, from PIE *\*h<sub>3</sub>nah<sub>2</sub>* by Werba (s.v. *\*nā*; EWAia II: 33–34). Joki (1973: 293) actually mentions OI verbal forms like *nāthate* ‘searches for help’ in his commentary of the Ob-Ugric etymology, but he lists them after the Avestan forms that Korenchy mentions, not making it clear that these Indic forms are unrelated to Avestan *nāidyah-*.

To sum up, Korenchy’s first etymology can be rejected with justification, and the other etymology remains a promising possibility, even if the palatal *ń* cannot be regularly explained. The Indo-Iranian root ‘to help’ is not attested in Ossetic (and no trace is found in the “Scythian” sources), making it unlikely that the Ob-Ugric words would be borrowed from Alanic.

An Indo-Iranian etymology (← *\*nātHa-*)

**Mansi East *n̄χ*, South *nāχ*, West *naχ*, Pl. *naŋkət*, etc. ‘Schneller, Hahn, Drücke, Stellnagel’ (< PMs ? *\*nāŋk*); Kha North *nekī*, East (J) *niki*, South *neka* ‘Band; Pflock zu Befestigung’ (< PKh ? *\*nVkV*)**

< POUG ? *nVkV* (Honti 1982: 202, No. 782)

← PII *\*nakHa-* > OI *nákha-m* ‘finger, toe nail, claw’, Oss *nyx*, *nix* (EWAia II s.v. *nákha-*)

(Blažek 1990: 40)

The Mansi velar nasal makes this Indo-Iranian etymology unlikely (it cannot be explained on the basis of the Indo-Iranian form, as *\*ŋ* cannot reflect earlier *\*k* or *\*γ*). Also the primary meaning of the Mansi and Khanty words is not ‘nail’, which makes the semantic connection between the Mansi and Indo-Iranian words far-fetched. The Indo-Iranian word has a good Indo-European etymology: < PIE *\*h<sub>3</sub>engh-/\*h<sub>3</sub>negh-* (> Latin *unguis* id.). The

*ngh* cluster of the Proto-Indo-European form could explain Mansi *ŋk* but the Indo-Iranian words are derived from *\*h<sub>3</sub>negh-* and do not show reflexes of a form with *ngh-* (PIE *\*ŋ* has been regularly vocalized to *\*a*).

The Khanty and Mansi words are not mentioned by Zhivlov (2006). Honti lists them at the end of his list of Ob-Ugric etymologies, among the words that are not regular cognates in Khanty and Mansi. In light of this it seems unlikely that the words are inherited from an earlier proto-language. They are probably parallel borrowings from somewhere, but the Iranian etymology is not likely.

Not an Indo-Iranian etymology

### **Fi *niska* ‘neck’ (cognates in Ka, Lu *nišk[e]*, Ve *nišk*, Vo, Lv) < PFI *\*niska***

< ? Pre-Fi *\*niska*

← PIA (?) *\*niška-* > OI *nišká-ḥ* ‘a golden ornament for the neck’ (EWAia II: 48 s.v. *nišká-*)

(Blažek 1990: 41; Parpola 2005: 47)

The Finnic word has no competing loan etymology. However, the etymology that has been suggested independently by Blažek and Parpola is not very convincing. First of all, the Indo-Iranian background of the Old Indic word is unclear (EWAia), which casts doubts on the possibility that it could have been borrowed from some early form of Indo-Iranian into Uralic. This uncertainty might be the reason that Blažek considers the word a specifically Indo-Aryan borrowing. The substitution of Indo-Aryan/Indo-Iranian *\*š* (> OI *ṣ*) is problematic, as one would certainly expect *\*š*, which is reflected by most early borrowings. Parpola (2005: 47, footnote 313) is aware of this weakness and actually notes that Jorma Koivulehto had rejected the etymology in personal communication because of this phonological problem. The sibilant *š* in Veps and Ludic has to be secondary from earlier *\*s*.

Also semantically the etymology is suspicious, although Blažek tries to explain the semantics through metonymy. In addition to the basic meaning ‘neck’, the Finnic word also refers to ‘Tragbalken; Stelle oberhalb einer Stromschnelle’, but these meanings are also not very close to the meaning ‘ornament’ that is found in Indo-Aryan.

Not an Indo-Iranian etymology

### **Mansi North *nul* ‘Flutwasser’; Khanty South (Ko) *nut-yŋk* ‘Flut, Fluss; river’**

< ? POUg *\*nud-* ‘river’

← PII *\*nadī-* > OI *nadī-* ‘river’, ? OAv *anaiḍīm*, root *\*nad-* ‘rauschen’ < PIE *\*ned-* (EWAia II s.v. *nadī-*)

(Blažek 1990: 41)

The Khanty and Mansi words are not regular cognates, and the etymology is not mentioned by Honti (1982). The very scarce distribution of the word in both Khanty and Mansi raises suspicions about its old age. Blažek gives a Proto-Ob-Ugric reconstruction *\*nud-*, but it is impossible to derive the Khanty and Mansi words from this. The Mansi and Khanty consonant correspondence is irregular, as Khanty *t* cannot be derived from earlier *\*d* (cf. PU *\*padi* ‘dam’ > Kh *\*p̄il* > Kh East *p̄l*). The word is certainly not an old Indo-Iranian loan into Proto-Ob-Ugric or Proto-Ugric. Semantically the etymology would be possible. It is probable that the resemblance of the Indic, Khanty and Mansi words is merely accidental, even though the Indo-Iranian word has a secure Indo-European background according to EWAia.

Not an Indo-Iranian etymology

**Fi *ohra* ‘barley’, dial. *otra*, Karelian *osra* (cognates also in all Finnic languages) < PFI *\*očra***

← PI *\*Hatsra-* or PII *\*Hačra-* ‘sharp’, from root *\*Hat-* ‘sharp’, *\*astra-*, *\*atsri-*, compare OI *ásri-* ‘scharfe Kante, Ecke’, *ašra-* in compounds < PIE *\*h<sub>2</sub>ek-* ‘scharf sein/werden/machen’ (LIV<sup>2</sup>: 261; EWAia I s.v. *ásri-*) (Kallio 2012b: 231, footnote 9)

(Paasonen 1917: 3–8; Jacobsohn 1922: 142–143; Kalima 1936a: 25; FUV: 133; Koivulehto 1979: 67–71; Kallio 2012b: 231, footnote 9)

An Indo-Iranian origin for the Finnic word was first presented by Paasonen, although it differed from the etymology discussed here. Since then this word has often been considered a Baltic loanword (cf. Baltic *\*aštra-* ‘sharp’; Koivulehto 1979), but Kallio (2012b) argues that an Iranian origin is more probable, because the consonant clusters of the modern Finnic forms can be explained from Proto-Finnic *\*-čr-*. A Baltic borrowing would have resulted in a three-consonant cluster *\*-str-*, which was probably not permitted according to Early and Middle Proto-Finnic phonotactics.

Paasonen connected the Finnic word etymologically with Hungarian *ostor* and Mansi *aštər* ‘whip’, and this explanation is still found in FUV, but Koivulehto, Korenchy (1972) and Joki (1973: 92–93) rightly criticize this. Paasonen’s idea was to derive the word from PII PII *\*(H)āčtra-* ‘whip’; as is remarked by K. Häkkinen & Lempiäinen (op. cit. 153), this noun is derived from the root *\*Haj-* ‘to drive’, not from *\*Hač-* ‘sharp’ (see EWAia s.v. *āštra-* and the entry *\*očtVrV*), and the loan etymology is semantically unlikely. Koivulehto notes that because *\*-str-* appears only in loanwords, this word has to be a loan, and this view is certainly correct. The Baltic etymology is supported by Rédei (1986c: 79) and K. Häkkinen & Lempiäinen (1996: 153–154).

Kallio’s newer Indo-Iranian etymology derives the Finnic word from the cognate of the Baltic word that means ‘sharp’. Kallio’s etymology is semantically just as convincing as the Baltic etymology, as both the Indo-Iranian and Baltic words denote ‘sharp’. The Indo-Iranian or Baltic origin of the Finnic word depends on Finnic phonotactics rather than characteristics of the Indo-Iranian or Baltic originals. The meaning of ‘barley’ is not attested anywhere in Indo-European.

Junttila (2012: 73; 2015: 103) lists the Baltic etymology among the “dubious Baltic etymologies”. Junttila (2015: 204) also notes that Kalima had rejected the Baltic etymology because of untenable reasons, namely because the Pre-Finnic speakers must have known barley before the contacts with Baltic started (yet the Pre-Finnic speakers also knew ‘tooth’, and still they borrowed \*šampas from Baltic \*žambas).

Semantically the etymology is plausible: Koivulehto remarks that in many Indo-European languages the term for ‘barley’ is derived from the meaning ‘sharp’.

This word is interesting, as if it is indeed borrowed from Iranian and shows the vowel substitution \*a → \*o, then it is evidence that this substitution occurred in both Proto-Indo-Iranian and Proto-Iranian loanwords.

Koivulehto is also right in criticizing KESK (1975) for connecting the Finnic word etymologically to Komi eč, which also means ‘barley’. Koivulehto considers this as a parallel borrowing, ultimately stemming from the same Iranian root \*Hats-. See the respective entry.

An Indo-Iranian etymology

**Hu *ostor* ‘whip’ (Old Hungarian *Ystwr*); Mansi South *ašt̄ar*, East *ōšt̄ar*, North *āšt̄ar* (< PMs \*āct̄ar; Zhivlov 2013: 219)**

**< PU/PUG ? \*oćtVrV ‘whip’ (Korenchy \*oćt̄r̄)**

← ? PII (H)aćtrā- ~ (H)ažtrā- > OI *aštrā-*, Av *aštrā-* ‘whip; Geissel, Peitsche, Knute’, MP *aštr* id. (EWAia I: 143; Werba 333 No. 378 < PII \*Haj’ (an)treiben’)

(Munkácsi 1901: 494–495; Jacobsohn 1922: 52; FUV: 134; Korenchy 1972: 56–57. No. 8; Joki 1973: 299; Ligeti 1986: 135; UEW s.v. *oćt̄r̄-r̄*; Katz 1985: 231; 2003: 187–188; Lushnikova 1990: 108; Zhivlov 2013: 220)

The Indo-Iranian origin of the Hungarian *ostor* and Mansi \*āct̄ar is an old and widely accepted idea. Semantically the connection of these Ugric and Iranian words is plausible. The points which have raised questions about this etymology are 1) possible cognates in other Uralic languages: these are considered obsolete these days, and for good reason; 2) the exact source language of the Ugric words (Proto-Iranian, “Andronovo Aryan”, Proto-Indo-Iranian); 3) the dating of the borrowing and connected to this 4) the relation of the Hungarian and Mansi words (whether they are regular cognates inherited from Proto-Ugric or parallel borrowings).

Point 1) will not be considered here in more detail, as the reader can find arguments against the inclusion of the other Uralic words in the works of Korenchy and Joki who argue that Mordvin E *ukšt̄or*, M *ušt̄ar* and Mari *wašt̄ar* cannot be regular cognates of the Ugric words (they are usually considered to be reflexes of \**wašt̄Vra* > Fi. *vaahtera* ‘maple’; Finnic \**oćra*, which has sometimes been connected here, is analysed in a separate entry in this work). In order to be fully able to comment on the points 2) and 3), we will start our analysis on point 4).

The Proto-Mansi form is reconstructed with  $*\bar{a}$  by Zhivlov (2013). According to Zhivlov & Reshetnikov (2010), this would regularly reflect Proto-Uralic  $*a$  in words that consist of more than two syllables; so it seems probable that the Pre-Mansi form of this word would be  $*a\acute{c}tVrV$  (?  $*a\acute{c}tara$ ). However, there is no possibility to derive Hungarian *ostor* from a similar source form: Hungarian *o* often reflects Proto-Uralic  $*o$  or  $*u$  (for example  $*tulka > toll$  ‘feather’, UEW s.v. *tulka*). PU  $*a$  and  $*o$  are not usually reflected by *o* in Hungarian. In Aikio’s (2015b) word-list the only examples of PU  $*a > Hu o$  are  $*ammi-$  ‘old’  $> Hu \acute{o}$ , where the labial vowel can be explained due to the influence of  $*m$  and the contraction that its loss has caused, and  $*a\eta ta-$  ‘to open’  $> old$  ‘to solve’ (see also Aikio 2018). In the latter case the reason for the development  $*a > o$  are unclear, but it might have something to do with the complicated developments of the consonant-cluster  $*-\eta t-$ . Among the Uralic  $*o-a$ -stems listed by Aikio, none is reflected by Hungarian *o*.

Actually, the word *ostor* is attested in Old Hungarian also in the form *Ystur* (here *y* most probably reflects /u/). This form is presented by Joki (1973), who does not comment on it in any way. Although the Old Hungarian sources are not always consistent in representing the vowels, a form like  $**ustur$  could well be the predecessor of the modern Hungarian form (the change  $*u > *o$  seems to be quite late, as numerous modern Hungarian *o*-forms are attested with *u* in mediaeval texts). However, this kind of vocalism would not suit the Indo-Iranian source very well:  $*a \rightarrow *o$  is a well-known and plausible substitution, but  $*a \rightarrow *u$  is not.

Whatever be the origin of the Hungarian vowel, it is clear that it cannot origin from a reconstruction like  $*a\acute{c}tVrV$ . This points rather clearly to the assumption that Hungarian *ostor* and its Mansi “cognate” are parallel loans from the same Indo-Iranian source. This is no surprise, as numerous (Indo-)Iranian borrowings in the Ugric languages include phonetic irregularities, and it has been assumed by several scholars that at least a portion of them are parallel borrowings from a time when the Ugric languages were still in contact with each other but could no longer be considered variants of the same proto-language.

What is the exact source language, then? This is a problem which is connected to the other early borrowings in Ugric, too. Traditionally it has been assumed that the words are “Iranian” borrowings (Korenchy 1972), but this is not necessarily the case. Joki’s idea that the word is a “Middle Iranian” borrowing is not convincing for obvious chronological reasons. Katz (1985) considers the Ugric word to be a “frühurisch” loanword (“frühurisch  $a\acute{c}tr\bar{a}$ ”), and this explanation is also mentioned by EWAia.

Zhivlov (2013) has considered this particular word one of the “Andronovo Aryan” borrowings. Zhivlov’s idea clearly implies that the Mansi sibilant in this word reflects a PU palatal affricate (~ sibilant)  $*\acute{c}/*\acute{s}$ , which cannot reflect Proto-Iranian  $*\acute{s}$ . Hungarian *s* [š] is usually the reflex of PU  $*\acute{c}$ , but in some consonant clusters it seems possible that  $*\acute{c}$  is reflected by *s* in Hungarian. The examples are not many, nor are they good, but at least  $*waški$  (?  $*w\acute{a}šk\bar{a}$ )  $> *vas$  seems to be such a case. The Mansi form neatly suits Proto-Indo-Iranian (“Andronovo Aryan” according to Zhivlov)  $*a\acute{c}tr\bar{a}$ - which has a palatal affricate. However, since the Indo-Iranian noun is probably derived from the root  $*Ha\acute{z}$ - ‘to drive’, the Proto-Indo-Iranian form should rather be reconstructed as  $(H)a\acute{z}tr\bar{a}$ .

In the traditional reconstructions of Mansi historical phonology (such as Honti 1999), Proto-Mansi  $*\acute{c}$  can be derived from PU  $*\acute{c}$ . If this sound law is supported, we can simply assume that the Mansi word is an old loan and  $*\acute{c}$  is regularly retained in Proto-Mansi. The

affricates of Hungarian and Mansi still require further study, but the Indo-Iranian etymology of both words is convincing.

An Indo-Iranian etymology (Hungarian and Mansi words are parallel loans)

### **Ko eč ‘Korn, Samen’; ? Ud *waž, waž, waž* ‘Spelt, Dinkel’**

< ?? PP \*oč (Csúcs 2005: 393)

← PI \*(H)ats- ‘sharp’ (< PIE \*h<sub>2</sub>ek<sup>ʰ</sup>-), attested in \*astra-, \*atsri-, compare OI *ásri-* ‘scharfe Kante, Ecke’, *ašra-* in compounds (EWAia I s.v. *ásri-*)

(KESK: 91–93; Koivulehto 1979: 67–71; Rédei 1986c: 79; Csúcs 2005: 393 Kallio 2012b: 231, footnote 9)

This is one of the few etymologies confined to Permic that show an affricate reflex of the Proto-Iranian sibilant \*ts. Note that Rédei’s notation “uriran. \*ač-” is misleading, as the affricate č is often used to mark the result of the second palatalization (< PIE \*k, not the “primary” affricate that reflects PIE \*k<sup>ʰ</sup>). Finnic \*ocra ‘barley’ (see the entry) is a parallel borrowing from Iranian \*atsra-, which is derived from the same IE root meaning ‘sharp’. Helimski (1997b: 121) lists this among the doubtful etymologies of Rédei.

Interestingly, here the semantic development from ‘sharp’ to ‘barley’ or ‘corn’ is almost similar as in Finnic \*ocra. One could wonder whether the unattested Iranian donor language had this development already, or whether this is independent semantic development in Finnic and Permic. As the donor forms of the two words are not formally identical, it is safer to assume the independent development in Finnic and Permic. Names for barley and other cereals are often connected to ‘sharpness’. Latin *acus* ‘awn’ is derived from the same IE root \*h<sub>2</sub>ek<sup>ʰ</sup>- ‘sharp’.

On the other hand, Kallio (2012b: 231, footnote 9) suggests that the Komi word is borrowed from a Finnic language which still preserved the Proto-Finnic affricate. This is also possible, and it would explain the semantics of the Finnic and Komi words.

Whether the donor form was Iranian or Finnic, the Permic word with simple \*č has to somehow reflect an earlier cluster or \*-čr-; either the *r* has been dropped when the word was borrowed, or it has been borrowed as \*očra and the *r* has become lost with the apocope of the last syllable. The Iranian form that Rédei cites (this would be \*(H)ats- in the modern reconstruction) is simply the root that denotes ‘sharp’, and there is no independent word (= root noun) \*\*ats- in Iranian. Rédei’s formulation is thus misleading. The invalidity of Rédei’s reconstruction is noted also by Szemerényi (1988). Csúcs (2005) still cites Rédei’s problematic reconstruction \*ač, considering the borrowing as uncertain.

Rédei connected Udmurt *waž, waž, waž* ‘Spelt, Dinkel’ to the Komi word and derived both from Proto-Permic \*oč. Kallio only suggests that the Komi word is borrowed from Finnic and he does not comment on the relationship between the Udmurt and the Komi words. The relationship between Komi and Udmurt words shows irregular vocalism, and there is little reason to assume that they reflect the same Proto-Permic word (the Proto-Permic reconstruction given by Csúcs could yield the Komi form but not the Udmurt). It is

possible that the Komi word is borrowed either from Finnic or directly from Iranian, and the Udmurt word is from some completely different source.

? Indo-Iranian etymology

**Est *oide* ‘Graswurzeln’, dial. *õidad*, *uit* (no cognates in other Finnic languages)**

← PII ? *\*waida-* > OI *vedá-h* ‘a bunch of grass’ (EWAia II: 581)

or

← PII *\*waita-* > OI *vetása-h* ‘Calamus Rotang or a similar cane’, Av *vaēiti-* ‘willow’, Oss *wīdag*, *wedagæ* ‘root’ (EWAia II: 578–579)

(Blažek 1990: 41)

This word is attested only in Estonian, making it an unlikely candidate for an early borrowing from Indo-Iranian. The word is also found in only certain Estonian dialects (VMS s.v. *oide*), where the vowel correspondences are irregular, and it is not even listed in EES. Of Blažek’s two alternative Indo-Iranian sources, the first one has an unclear etymology, making it an improbable source for borrowing.

The other Indo-Aryan word (*vetásah*) also has cognates in Iranian and it reflects PIE *\*wey(H)-t-* (from the root *\*wey(H)-* ‘to bind, to twist’), which is reflected also by Germanic words for willow, such as German *Weide* ‘willow’, Old High German *wīda* and Old Norse *vīðir* (EWAia; Kluge 2012: s.v. *Weide*). The more credible Indo-European etymology of this Indo-Aryan word means that the word existed in Proto-Indo-Iranian already, and assuming that this word was borrowed into some early form of Finnic is less troubling, although the limited distribution in Estonian makes also this etymology unlikely, and because of the phonological difficulties mentioned above, the word is probably not an Indo-Iranian loan. The Germanic cognates of the Indo-Iranian forms reflect zero-grade forms of the IE root (PG *\*wīþja/ō*, *\*wīþig*), making also a loanword from Germanic to Finnic unlikely (the Germanic words are also semantically rather far from the Estonian word).

Not an Indo-Iranian etymology

**Fi *oja*, *ojas* ‘shaft of plough’ (cognates in Lu and Ve)**

← PII *\*ojas* ‘shaft’ [sic!] (see *\*ajša* for references)

(Toivonen 1917–20; Mayrhofer 1952: 13; Joki 1973: 293; SSA II s.v. *ojas*; Koivulehto 1999a: 227; 2016: 229)

The etymological dictionaries, especially SSA, present this etymology in a very confusing way. SSA gives a reconstruction that is neither Proto-Indo-Iranian nor can give the attested Old Indic form. The etymology itself is probably not correct either.

SSA II s.v. *ojas* gives a PII reconstruction *\*ojas*, which is impossible but stems from an old idea by Toivonen. The latter scholar suggested that the Finnic word was borrowed from Pre-Indo-Iranian *\*oyas*, with retained *o* in the initial but not in the second syllable. This is based on views that are now outdated, as *\*ojos* would have been impossible in Pre-Indo-Iranian, and there is no need to argue that the donor form would show both retained *\*o* and already *\*a* from PIE *\*e*. Mayrhofer (1952: 13) accepts Toivonen’s etymology and ideas, even though the hypothetical Pre-Indo-Iranian form *\*oyas* does not lead to any of the attested Indo-Iranian words. Also Joki considers the word a possible “frühurarisches” borrowing, giving a reconstruction “*\*oyas* ~ *\*oyos*”.

Joki mentions that K. Vilkuna has attempted to derive the Finnic word from mediaeval Russian, but stating that the word is not “tatsächlich belegt” in Russian. However, it seems indeed likely that the Finnic word is a late borrowing from Slavic *\*oje* ‘thill’, the Finnish, Veps and Lude words probably being parallel borrowings. The word is present in dialects of Russian as *оё* (Derksen 2015: s.v. *\*oje*), and the distribution in eastern North Finnic means that a Slavic borrowing is a likely option, even though the details require more research.

Not an Indo-Iranian etymology

## Karelian, Veps *ola* ‘flint’

< Pre-Fi *\*ola* ?

← ? PII, cf. OI *aṣṭhīlā* ‘kuula, pallo, pyöreä kivi, piikivi’ (EWAia III: 19)

(Lidén 1897: 83–85; Vilkuna 1933: 160–162; Joki 1973: 294; Burrow 1976: 61; Mayrhofer 1952: 13; SSA s.v. *ola*)

The etymology was suggested by Lidén and is still considered an Indo-Iranian loanword by Mayrhofer (1952: 13). Joki (1973: 294) is doubtful and assumes that the word was borrowed from Baltic instead. Also Koivulehto (2000a: 118) assumes a Baltic origin. There is not much substance in the Indo-Iranian etymology, since the Indo-Iranian word has an unclear background. Lidén (1897: 83–85) assumed that the Sanskrit word is an old compound/derivation from *\*al-s-thī-*, and the first syllable would correspond to Baltic *\*ōla*. This explanation is not mentioned by EWAia, but in KEWA Mayrhofer considered the etymology unlikely, as is cited by Joki (1973: 294). Also Burrow (1976: 61) considers the etymology unlikely.

Within Indo-Iranian, the Sanskrit word is isolated. The competing Baltic etymology is more convincing: from Proto-Baltic *\*ōla*, cf. Latvian *uola* ‘a small stone; egg’. However, it is important to note that Derksen (2015: 481) does not provide any Indo-European cognates for the Baltic word, and he seems to doubt even the Proto-Baltic origin of the Baltic word. The borrowing of the word from Baltic to Finnic and Mordvin would, of

course, support its presumed existence in Proto-Baltic, if the Uralic and Baltic words are not parallel borrowings from a third unknown source.

Kildin Saami *vueʹll* has often been connected to the Karelian and Veps words, but according to SSA it is a borrowing from Karelian. This is probably true, as the word has no cognates in other Saami languages.

The Baltic word was borrowed also into Mordvin separately as *al* ‘egg’. Grünthal (2012: 312–313) assumes that the Finnic and Mordvin words could reflect an inherited Proto-West Uralic form, but this is impossible because the Mordvin word could reflect an earlier *\*a-a*, *\*i-i* or *\*i-a*-stem, whereas the Finnic word can only reflect earlier *\*o-a*.

Not an Indo-Iranian etymology

**? Hu *odú*, *odor* ‘Höhle’; Ms North, East *ānter*, West *oanter*; Kh South *unt*, *unter*, East *ont***

< PUG? *\*omtVrV* (UEW s.v. *omte*, *oŋte*; Sammallahti 1988: 542; Abondolo 1996: 74)

← PII or PI *\*(H)antara-* ‘innerer’, OI *ántara-*, YAv *aŋtara-* ‘innerer, inner befindlich’ < PIE *\*h<sub>1</sub>en-ter-o-* (EWAia I s.v. *ántara-*)

(Katz 1985: 156; 2003: 129, footnote 24; Korenchy 1972: 46–47; Lushnikova 1990; Helinski 1991: 220, 221; Abondolo 1996: 74)

This is a very difficult etymology, because the Uralic forms that have sometimes been listed under one reconstruction clearly point to different forms. In addition to the Ugric forms mentioned above, Finnic *onsi* and Saami *vuowda* have been considered as cognates. As Korenchy remarks, the Saami word *vuowda* points to a cluster *\*-mt-*, which would be difficult to explain from Indo-Iranian *-nt-*. Semantically the Indo-Iranian (or other Indo-European) etymology would make sense.

Sammallahti (1988: 542) reconstructs both *\*omti* and *\*oŋti* (both in the meaning of ‘cavity’). Also Abondolo (1996: 74) gives *\*amti* ~ *\*umti*, and mentions also that a “near-synonym” with *\*ŋ* can be reconstructed. Because the Ob-Ugric languages show both forms with and without *-r*, it seems that the *-r* in the Ob-Ugric languages is a derivational suffix, as is probably *-r* in Hungarian *odor*, so the resemblance of the Uralic words and the Indo-Iranian noun *\*antara-* is probably simply accidental.

Katz has assumed that the Mansi word is borrowed from Indo-Iranian, and the rest of the words can be explained as an inheritance from PU *\*omti*. It remains uncertain why the Mansi word could not also be connected to this Uralic word, and the idea of an Indo-Iranian loan remains highly uncertain.

Not an Indo-Iranian etymology

**Fi *onki* ‘fish-hook’; SaN *vuog’ga* (< PSa *\*vuonŋe*); Mari *oŋgo*, *oŋgə* id < PMA *\*oŋgə***

< PU *\*oŋki* (UEW s.v. *\*oŋke*)

← PII *\*(H)anka-* > OI *ańká-* ‘Haken’, < PIE *\*h<sub>2</sub>enk-*, ON *angi* ‘brodd’, Icelandic, Faroese *angi*, MNI *anghe*, OE *anga*, OHG *ango* ‘brodd, hengsel’, MHG *ange* ‘fiskekrok, hengsel’ (EWAia I s.v. *ańká-*, *AÑC*; Bjorvand & Lindeman 2000: 34–36 s.v. *angel*; Kroonen 2013: 28 s.v. *\*angan-*)

(FUV: 134; Joki 1973: 295–296; Katz 1985: 343–344; 2003: 273–274; Rédei 1986c: 54; Lushnikova 1990: 176)

The Indo-Iranian etymology for this word is an old idea, even though other Indo-European sources are also possible. According to LÄGLOS; the Finnic word is possibly borrowed from Germanic. This is a convincing explanation, and the Indo-Iranian etymology can be rejected. Semantically the Germanic etymology is better. This is probably true also of the Saami word, which can be a cognate of the Finnic word or a separate borrowing from Germanic. Here Finnic/Uralic *\*o* from Germanic *\*a* can be explained with the substitution rule *\*o* ← *\*a* before a nasal that is found in many Germanic loans.

The Mari word can hardly be a Germanic loanword, so it remains possible that it is a separate borrowing from Indo-Iranian (or some other branch of Indo-European). Formally the Mari, Saami and Finnic words could be cognates and reflect a Uralic *\*o-i*-stem.

Katz (1985: 343–344; 2003: 273–274) considers Fi *onki* and Mari *ońgo* as parallel loans from Indo-Iranian. Katz likewise mentions Udmurt *ugi* ‘Ohrring’ and Komi *vugir* ‘Angel’, Ud *vug* ‘Bogen, Henkel’ as parallel loans. The latter Permic words are also mentioned by FUV (141) as parallel borrowings. Udmurt *ugi* could hardly be derived from earlier *\*o* but it can reflect an old *\*a-a*-stem. The word can perhaps be a loan from Iranian *\*anka-*. Komi *vugir* and Ud *vug* probably do not belong here because of the word-initial *v-*. They could be borrowed from Iranian *\*wanka-* (see the entry *\*wanjka*).

Not an Indo-Iranian etymology

**Mari *orža*, *arža* ‘Mähne (eines Pferdes); Landrücken’ (TschWB: 47) < Proto-Mari *\*arža***

← ? PII *\*warća-* ~ PI *\*wartsa-* ~ later Iranian *\*warsa-* > Av *varəsa-* ‘Haar (besonders Kopfhaar) des Menschen und Tiere’, Middle Persian *wars* ‘Haar’, OI *válsa-* ‘Schössling, Zweig’ (EWAia II: 545, s.v. *válsa-*)

(Jacobssohn 1922: 44, footnote 2; Setälä 1928: 306–307; Joki 1973: 251; Koivulehto 1999a:)

This etymology is old but it has not received attention in recent research on loanwords. According to Joki, the Mari word has been borrowed from Chuvash *orža*, *urža* ‘Widerrist’, *orža* ‘Mähne’, but this explanation is problematic because the Chuvash word has no cognates in the other Turkic languages.

Even if the Mari word is borrowed from Chuvash and not from Indo-Iranian, it is unlikely that the resemblance of the Iranian, Mari and Chuvash forms is accidental. Also the Chuvash word could be borrowed from Iranian. It also remains unclear whether all the three Chuvash words mentioned here can really be derived from a single proto-form (the

sibilant correspondences are not regular), and it indeed seems that the Chuvash words are relatively late loanwords from Mari.<sup>28</sup>

Because of the relationship *o* : *a* between Mari dialects, the word can be reconstructed for Proto-Mari. The Pre-Mari vowel would be \**ā*, which can reflect Pre-Mari/Uralic \**a*, cf. Mari *pat*, *pot* ‘pot’ < Proto-Mari \**pāt* < PU \**pata* (Zhivlov 2014: 128). This would fit well with the substitution of PII/PI \**a*.

The only problem with the etymology seems to be that the word-initial *w*- should be retained in Mari (cf. Mari *wanžem*, *wońčem* ‘to open’ < PU \**wanča*-, Mari *wurđo*, *wurđā* < PU \**warti* ‘stalk’, Zhivlov 2014: 128, 132). Because of this and the unclear relationship of the Mari and Chuvash words, the etymology cannot be considered certain. The Mari word is probably ultimately from Indo-Iranian, but it is more difficult to say whether it is a direct borrowing or not.

The Khanty (PKh) \**wār(ə)sV* ‘Roßhaar’ (Korenchy 1972: 82–83 No. 42; Joki 1973: 336; Zhivlov 2013: 11) is borrowed independently from the same Indo-Iranian source (it would be impossible to derive the Mari and Khanty words from the same proto-form). Possibly Fi *verso* and its Finnic cognates also go back to a separate borrowing of the same Indo-Iranian word (see the entry).

? Indo-Iranian etymology

**Fi *ora* ‘awl’ (cognates in all Finnic languages); SaI *oari* ‘awl’ (< PSa \**oarē* cognates in Ko, Kld, T, attested by Friis in SaN); Md E *uro* M *ura*; Hu *ár* ‘awl’**

< PU \**ora* (UEW s.v. *ora*; Sammallahti 1988: 542; SSA s.v. *ora*; Zhivlov 2014: 138; Aikio 2015b: 61)

← PII \**HārāH*- > OI *ārā*- ‘Treibstachel’; < PIE \**h<sub>1</sub>e/oh<sub>1</sub>l-eh<sub>2</sub>*-, cf. ON *alr* ‘awl’ (< ? \**h<sub>1</sub>h<sub>1</sub>l*-) (EWAia I s.v. *ārā*-; Lubotsky 1988: 104; RIVELEX II s.v. *ā rā*-)

(Setälä 1900: 32; Munkácsi 1901: 644; Jacobsohn 1922: 51–52; FUV: 134; Joki 1973: 296; Gamkrelidze & Ivanov 1984: 923; Rédei 1986c: 46–47; Gyula 1992; 2003: 81–82; Koivulehto 2001b: 248 = 2016: 290; Katz 1986; 2003: 262; SSA II s.v. *ora*; Parpola 2015: 63)

This is a well-known and convincing loan etymology, almost universally accepted (SSA mentions the Indo-Iranian etymology with a question mark). The loan was clearly acquired already by Proto-Uralic, as its distribution stretches from Finnic to Hungarian. What is more, as the cognates are regular in all the branches where the word appears, there is no need to suppose a parallel borrowing.

\**ora* is a well-known example of an “*o*-loan”. Here \**o* substitutes the long vowel \**ā*, but this does not tell us much, as there seems to be no difference in substitution of the long and short Indo-Iranian vowels in Proto-Uralic (only short vowels were allowed). Koivulehto has argued that this word is evidence of the true substitution of \**a* by \**o*, as the Indo-Iranian word reflects earlier \**ēlā*. However, this is not an universally accepted

---

<sup>28</sup> I am grateful to Christopher Culver for clarifying the Chuvash data out to me.

reconstruction. EWAia, which accepts the etymology of the Uralic word, derived the Indo-Iranian word from PIE *\*(H)oléh<sub>2</sub>-*. Lubotsky (1988: 104) reconstructs the Indo-European word as *\*h<sub>1</sub>e/oh<sub>1</sub>l-eh<sub>2</sub>-*. Because of this uncertainty, it is not possible to state for certain that the donor form did not have *\*(H)o-* (or *\*(H)ō-* from Brugmann's law) instead of later *\*(H)ā-*, but because there are many other examples of the *\*o*-substitution, there is no need to postulate a very archaic donor form instead of Proto-Indo-Iranian *\*HārāH-*. The word has to be an Indo-Iranian loan in any case, as the only securely related forms are found in Germanic, which has retained PIE *\*l*.

Gulya (2003: 81–82) argues that the Ugric (Hungarian) word is a separate loan, and that Hu *á* here reflects PUG *\*a*, not *\*o*. However, since the Hungarian vowel *á* regularly reflects both *\*o* and *\*a*, there is no reason to assume that it reflects a separate borrowing.

Gamkrelidze & Ivanov (1984: 923) argue that this etymology is an example of the *\*o*-substitution in front of *\*r*. Here it is interesting that the Indo-Iranian long *\*ā* was also substituted by *\*o* in similar environments as the short *\*a*.

Parpola (2015: 63) has noted that *\*ora* is borrowed from Proto-Indo-Aryan because the word is not found in Iranian. As the Indo-Aryan word has a likely Indo-European etymology, its absence in Iranian does not count for much, especially as the Old Iranian corpus is in any case much smaller than the corpus of Indo-Aryan texts, and in any case the word could have simply disappeared from the Iranian branch. It is safe to consider this word as a Proto-Indo-Iranian loanword, as has usually been done.

An Indo-Iranian etymology

**Fi *orja* 'slave'; SaN *oarji* 'south; west' (cognates in all Saami languages < PSa *\*oarjē*); Md E *urē*, M *urä* 'slave'; Ud *war* 'slave, servant'**

< PU *\*orja* 'slave' (UEW s.v. *orja*; Zhivlov 2014: 138; Aikio 2015b: 61)

← PII *\*ārya-*, > OI *ārya-*, Av *airiia-*, OP *ariya-* (Thieme 1938; EWAia I: 174–75, s.v. *ārya*; Schneider 2010: 102–103)

(Paasonen 1896: 49; Joki 1973: 297; Burrow 1976: 61; Katz 1985: 204–207; 2003: 168; Gamkrelidze & Ivanov 1984: 924; Rédei 1986c: 54–55; Katz 1987: 452; Szemerényi 1988: 174–175; Lushnikova 1990; Parpola 1999: 196–197; Sammallahti 2001: 408; Holopainen 2018b: 153–154; Kümmel 2018a)

Paasonen's etymology has been widely accepted in most later sources, although the semantic development has repeatedly raised problems (these are raised especially by Joki 1973), and some authors have been cautious to connect the Indo-Iranian and Uralic words on the one hand, and the Uralic words themselves on the other hand. Phonologically the etymology is completely plausible: Uralic *\*o* as the substitute of Indo-Iranian *\*a* or *\*ā* appears in numerous loanwords. Gamkrelidze & Ivanov (1984: 924) have noted that this is one of the examples that manifest the rule that *\*o* as the substitute of *a*, *ā* appears in front of *\*r*. Lushnikova, on the other hand, argues that this shows that *\*o* was the regular substitute in the absolute word-initial position. Although some of the *o*-loans can be explained as borrowings from a Pre-Indo-Iranian stage with retained *\*o* or as borrowed

from some other branch of Indo-European with retained *\*o* (see *\*ońcá*, *\*orpV*, *\*poréas*), *\*orja* can only be explained as an Indo-Iranian loan, because the word has no cognates in other languages of the Indo-European family.

The more precise origin of the Indo-Iranian word is not well-known; it is possible that it is derived from *\*ari-* ‘stranger’ (Thieme 1938), but Schneider (2010: 102–103) notes that Indo-Aryan *arí-* is an isolated word which does not have a secure Indo-European or Indo-Iranian etymology. According to Burrow (1976: 61) the Uralic etymology can only be accepted if *ārya-* is indeed derived from *arí-* ‘stranger’, but I see no obstacle in accepting the etymology even if it is just borrowed from the ethnonym *ārya-*, as the meaning ‘foreigner’ (from which later ‘slave’ can be explained) can still be the original meaning in the Uralic side. Whatever the source of the Indo-Iranian word may be, it does not influence the Uralic etymology much, as the Uralic word *\*orja* can be derived from Proto-Indo-Iranian (or Proto-Iranian) *\*āryá-* without problems. The semantic development from an ethnonym of a foreign group into ‘slave’ has many parallels in the world’s languages, such as Old English *wealh* ‘foreigner; Welshman’ > ‘slave’ and Amharic *barya* ‘name of a tribe in Central Eritrea’ > ‘slave’ (DatSemShift s.v. slave), as well as English *slave* (← Lat *sclavus* ← Gr σκλάβος, originally from the ethnonym of the Slavic peoples).

Szemerényi (1977b; 1988: 174) has argued that the Indo-Iranian ethnonym *\*āryá-* is borrowed from Semitic, but this has been criticized by Kümmel (2018a). Szemerényi (1988: 174) is also critical of the Indo-Iranian etymology of the Uralic word because of the semantic development, but his criticism is not very precise: he points out that Rédei assumes that the word is borrowed from the word family *ari-*, *āryá-*, with the original meaning ‘foreign’, and that a word for slave could hardly be derived from this. But as shown above, parallels for this exact semantic development can be pointed to, so Szemerényi’s argument can be rejected.

The question of whether the Saami word belongs here has been debated in Uralic linguistics. The UEW mentions the Saami forms with a question mark, and Zhivlov lists them among the regular reflexes of *\*orja*. Sammallahti (2001) mentions the Saami word as a parallel borrowing from Indo-Iranian. Aikio (2015b) does not mention the Saami word among the reflexes of PU *\*orja*. Pärpola (1999) has supported the explanation that the Saami meaning ‘south, west’ has developed from the reference to the location of more southerly Indo-Iranians. For the semantic development that is found in the Saami word, an interesting parallel can be found in East Slavic. Old Russian *упей* (*вьпей*) ‘a southern land to which birds of passage migrate, a fabled magical realm’ has been assumed to be an Indo-Iranian borrowing from this same ethnonym (Kim 2018: 13). Phonologically there is nothing to assume that the Saami word should be a parallel loan from Indo-Iranian.

Komi *ver* ‘husband, (Old Komi) servant’ has been connected into this word in many etymological works, such as the UEW (s.v. *orja*) and KESK (s.v. *ver*), and the Komi word is still mentioned as a cognate to Finnic *orja* by EES, even though *ver* cannot regularly reflect earlier *\*orja*. Already Munkácsi (1901) assumed that *ver* is borrowed from the Indo-Iranian word for man, *\*wiHra-* ‘man’. This explanation is supported by Katz (2003), and it seems indeed the mostly likely explanation for the origin of the Komi word. See the entry *\*wirä* for more detailed arguments on the Komi word. Udmurt *var* can be regularly derived from *\*orja* and its connection to the Finnic, Saami and Mordvin words seems certain. Joki (1973) considered also the connection of Udmurt *var* to this etymology irregular, but his doubts are not based on any convincing arguments.

Koivulehto (1999b: 330–331) has also suggested an alternative Indo-European etymology for the Uralic word, deriving it from PIE *\*worg<sup>h</sup>o-* (attested for example in Doric Greek δαμιόργος ‘Fachmann’). EES mentions this as one possibility alongside the Indo-Iranian etymology. This etymology is semantically suitable too, but since there are only a handful of tentative etymologies with PIE *\*g<sup>h</sup>* being substituted by Uralic *\*j* (Holopainen 2018a), the Indo-Iranian source is more compelling. Also, as the word-initial *\*w-* of the Indo-European word has no reflexes in Uralic, the idea of deriving *\*orja* from *\*worg<sup>h</sup>o-* is not very convincing. From an original meaning ‘worker’ it would also be more difficult to explain the semantics of the Saami word.

An Indo-Iranian etymology

**Fi *orpo* ‘orphan’ (reflexes in most Finnic languages; Est *orb*: *orvu* borrowed from Finnish); SaN *oarbīs* ‘orphan; lonely and deserted’; Mo E *urus*, M *oros* ‘orphan’; Khanty *-urī* in the compound *jəŋg-urī* ‘orphan’ (< PKh *\*ōrpī* or *ūrṗī*, with a suffix *\*ṽ*)**

< PU *\*orpV* (UEW s.v. *\*orpa*; Sammallahti 1988: 542; Zhivlov 2014: 138; Aikio 2015b: 61)

← PII *\*Hár<sup>h</sup>a-* > OI *ár<sup>h</sup>a-h* ‘small, weak, young’ < PIE *\*h<sub>3</sub>órb<sup>h</sup>o-*, cf. also Gr ὀρφανός, Lat *orbus*, Arm *orb* ‘orphan’, Goth *arbi* ‘inheritance’, root *\*h<sub>3</sub>erb<sup>h</sup>-* ‘to separate’, > Hittite *harp-* ‘to separate oneself’ (EWAia I s.v. *ár<sup>h</sup>a-*; Lubotsky 1988: 75)

(Setälä 1900: 31, footnote 2; FUV: 134; Korenchy 1972: 54–55 No. 5.; Joki 1973: 297–298; Katz 1985: 218, 220; 1987: 454; 2003: 178–179; Rédei 1986c: 46; Lushnikova 1990: 177–178; Koivulehto 1991: 106; 1999: 308; SSA s.v. *orpo*; UEW No. 1472 s.v. *\*orpa*; Bereczki 2013 s.v. *rweze*; EES s.v. *orb*; Milanova, Holopainen & Bradley 2018 [forthcoming])

The Indo-European origin of the Uralic word for ‘orphan’ is a widely accepted idea, but it is not certain that the word is borrowed from Indo-Iranian. Already Setälä (1900: 31, footnote 2)<sup>29</sup> has considered this word to represent an early loan with retained *\*o*, rather than Indo-Iranian, and this view was later echoed by FUV where the word is mentioned as a Proto-Indo-European loan. Koivulehto (1991: 106) considers the word to be either a Proto-Indo-European or early Proto-Indo-Iranian loan. The problem here is that *\*o* is a frequent substitution for PII *\*a*, so it is very difficult to determine whether the Uralic vowel reflects *\*a* or earlier *\*o*, especially as the factors causing the *o*-substitution are debated. It has been assumed that the environment before *\*r* is typical for the *\*o* loans where PII *\*a* appears as *\*o* in Uralic (Gamkrelidze & Ivanov 1984: 924).

The idea of Rédei (1986c: 46) that *\*orpa* is an Indo-Iranian loan is connected to his idea that all the early loanwords can be derived from various stages of Indo-Iranian, and he

<sup>29</sup> In addition to an early form of Indo-Iranian Setälä also mentioned “armeno-phrygo-thracish” as a possible donor language, but this branch of Indo-European does not exist according to modern views on Indo-European taxonomy.

does not substantiate the Indo-Iranian etymology with any convincing arguments. Likewise, EWUNG's statement that Uralic *\*orpa* is a loan from Proto-Iranian (!) seems to be a hasty conclusion that is not based on actual evidence, as this word has no PI phonological features, and there seem to be no PI loans which were borrowed already into "Proto-Uralic (or Proto-Finno-Ugric) proper".

Korenchy considers this a Pre-II loan because of the *o*, noting that the word serves as a counter-example of her (and Harmatta's) idea of the earliest loanwords being from Iranian. Korenchy reconstructs the Uralic word as *\*orpas*, arguing that the ending *-as* points to an early borrowing, although this alone is usually not considered an obstacle to the Iranian nature of the loans by Korenchy (see the entry on *\*pakas*).

Because of the ambiguous vowel substitution, for this etymology semantics can give more decisive arguments than phonology. Burrow is sceptical of the etymology, because 'orphan' is not an attested meaning of the OI word *árbha-*, *árbhaka-* 'small, small child', and he assumes that the Uralic word is a loanword from some other branch of Indo-European, where the meaning 'orphan' would be attested.

Dolgopolsky (1989: 8, 14) mentions that the same Indo-European word has possibly been borrowed also into Semitic as *\*Harb-* or *\*arb-* (> Akkadian *arb-* 'fugitive, person without family'), as well as to Kartvelian *\*obol-* (> Georgian *obol-* 'orphan') and into some other languages of the Caucasus as well. Dolgopolsky is uncertain of the exact directions of these borrowings. Further, he does not specify the Indo-European sources of these words, but at least the borrowing into Semitic has to be quite early. This is interesting from the point of view of the semantics of the IE word, as 'orphan' should then be reconstructed for the stage of Indo-European that provided the loans into Semitic and Kartvelian (if these etymologies are correct).

Simon (in press) notes that the meaning 'orphan' developed later in the individual branches of Indo-European, where this meaning is attested (such as Italic and Greek), and it cannot be reconstructed for the IE noun. This means that the word is probably not borrowed from Proto-Indo-European, but from some other early branch. It has been assumed that there are loanwords from Proto-Northwest-Indo-European in the Uralic languages (Koivulehto 2001b: 239–245), and this word could perhaps belong to this loanword layer, if we assume that the shift in meaning was present at that stage already. Balto-Slavic could be an obvious choice of borrowing, but there no meaning 'orphan' is attested – in Slavic (for example, Russian *ребёнок*) the reflex means 'child'.

Here no reflex of the Indo-European laryngeal *\*h<sub>3</sub>* is found on the Uralic side. This is also noted by Koivulehto (1991), who assumes that the early Indo-European loans that do not show laryngeal reflexes are borrowed from a different Indo-European dialect than the laryngeal loans. This is a possible explanation, but here one should also pay attention to the fact that the laryngeal in this etymology is *\*h<sub>3</sub>*, which is known for its labializing effect on PIE *\*e* (*\*h<sub>3</sub>e* > *\*h<sub>3</sub>o*), and it is possible that if the sound indeed was some kind of labial fricative, it would have been alien to the Uralic sound system, and the fact that it is not substituted is phonetically expected. Note also that *\*h<sub>3</sub>* is the laryngeal that has left no consonantal reflexes in Anatolian. According to Kümmel (2018b: 169: 170), *\*h<sub>3</sub>* was eventually lost in Indo-Iranian, and might have been a velar, uvular or pharyngeal fricative or glide in the Proto-Indo-Iranian stage.

This word has a wide distribution in Uralic, which fits well with the idea that it is indeed a very early Indo-European borrowing. The Uralic cognates are regular, but the

reconstruction of the non-initial syllable is more problematic. Korenchy assumed that the *\*-as* which is found in Saami means that the word was borrowed as *\*orpas*, with *-as* as the substitution of the Indo-European nominative ending (this substitution appears in many loans, see for example *\*tajwas*, *\*poréas*). Also Zhivlov (2014) reconstructs the Uralic word as *\*orpas*. The Finnic word with a labial vowel is clearly a derivative, and the old suffix *\*-j* is still visible in various Finnic forms like Karelian *orpoi* (see *\*repä* for a similar case). The Hungarian and Khanty words do not help much here, as the last syllable was regularly lost in both branches. Hungarian *-a* has to be a generalized possessive suffix *-a* (UEW s.v. *\*orpa*).

Katz (1987: 454) argues that the Khanty and Hungarian words continue a separate loan into Proto-Ugric. According to Katz, *ó* would be the expected outcome of PU *\*o* in Hungarian, but this is not correct, as examples for PU *\*o* > Hu *á* abound, cf. the well-known etymologies PU *\*kota* > Hu *ház* ‘house’, *\*wolka* > Hu *váll* ‘shoulder’. At least the Hungarian word can regularly be inherited from earlier *\*orpa*. The East Khanty word is a more difficult case, as Zhivlov (2014: 118) notes that it is the only old *\*o-a*-stem which shows *\*ō* or *\*ū* in Proto-Khanty. There is thus a possibility that the Khanty word was borrowed separately, but as Zhivlov remarks that there in total rather few examples of Uralic *\*o-a*-stems in Khanty, it can also be assumed that we simply do not know the development leading to Proto-Khanty *\*ōrpī* or *\*ūrpī* well enough.

A separate borrowing into Proto-Ugric and Proto-Finno-Permic has also been suggested by Gyula (2003: 81–82), who argues that the Ugric and Finno-Permic forms point to different vowel substitutions.<sup>30</sup> However, at least the Hungarian word can be derived from *\*o* like the more westernly words, so there is no reason to support Gyula’s arguments.

It is uncertain whether Finnic *orpana* ‘relative’ and its cognates belong to this same Uralic word. The UEW and Zhivlov reconstruct two different Uralic words. See below (*\*orpa*) for a more detailed discussion.

? Indo-Iranian etymology (rather a loan from some other branch of Indo-European)

**Fi** *orpana* ‘cousin, pl. relatives (cf. *orvasta-* ‘regard as relative; visit relatives’); **SaN** (obsolete) *oar’ben* ‘sister’, *oab’ba* ?; ? **Mo** *E ur’va*, **M** *ar’ve* ‘daughter-in-law; wife’

< ? PU *\*orpV* ‘a kind of relative’ (FUV; UEW s.v. *\*orpa*(?); Zhivlov 2014: 138)

? ← PII *\*Hárb<sup>h</sup>a-* > OI *ár**b**ha-* ‘small, weak, young’ (see the entry *\*orpa* for references)

(FUV: 134; Nirvi 1952: 82–87; SSA s.v. *orpana*; Bereczki 2013 s.v. *rweze*; Milanova, Holopainen & Bradley [forthcoming])

<sup>30</sup> Gyula (1992) argued that both Ugric and Finno-Permic words can be derived from PU *\*o*, but later Gyula (2003) has rather argued that Ugric points to an *\*a* substitution, whereas Finno-Permic points to *\*o*. It is difficult to understand Gyula’s argumentation. The same concerns also *\*kota* and *\*ora*, but here too both the Ugric and Finno-Permic words can be derived from an unitary proto-form and there is no reason to assume parallel borrowings.

The possible etymological connection between Uralic *\*orpa* ‘orphan’ and the West Uralic homonymous *\*orpa* ‘relative’ and their Indo-European origin is an old problem, and to date no consensus exists. SSA considers the two words related, but their exact relationship remains open. The UEW reconstructs two different entries, and considers only *\*orpa* ‘orphan’ to be an Indo-European loanword. Also Nirvi (1952) and Zhivlov (2014: 138) consider the two words etymologically distinct. West Uralic *\*orpa* has been derived from the same Indo-European source as *\*orpo* ‘orphan’, namely PIE *\*orb<sup>h</sup>o-* (> OI *arbha-*), but semantically this etymology is much more difficult than the Indo-European origin of the Uralic ‘orphan’ word. Phonologically there are no difficulties: Uralic *o* can here too be a reflex of retained PIE *\*o*, or it can be a substitution for PII *\*a* (see the entry *\*orpa* above for discussion about the phonetic environments of the *o*-loans). As far as I know, the idea that Uralic *\*orpana* is a separate borrowing has not been presented as such, but as its relationship to the more secure borrowing *\*orpa* is uncertain, it is necessary to discuss this Uralic etymology here as well.

UEW connects Mari *rweze* ‘boy’ here with a question mark. The Mari word does not correspond regularly to the other Uralic forms mentioned here and can be a parallel loan from Iranian; this has been argued by Bereczki (2013 s.v.), although an exact donor-form is not presented. The semantics of the Mari word can be probably be derived from the meaning of the Indo-Iranian word ‘small, young’, especially as the meaning ‘small child’ is attested for the Sanskrit derivative *árbhaka-*. The origin of the suffix *-ze* in the Mari word is unclear.

However, also an alternative proposal has been suggested, deriving the Mari word from Alanic. Gordeev (1967: 195–196) argues that the Mari word is borrowed from the pre-form of Ossetic *ærvad* ‘brother’ (< PI *\*bratār-*).<sup>31</sup> The same origin has been suggested for Ugric *\*arV*, but this is not a very convincing etymology (see the respective entry). The Alanic origin for the Mari word is possible, although a semantic change has to be assumed here. The word would explain the cluster *-rw-* but the relation of Ossetic *d* to Mari *z* is not very clear, and the Alanic origin of the Mari word cannot be considered as certain. Hopefully future research can reveal more details about the origin of this Mari word.

FUV is uncertain of the connection of *\*orpa* and *\*orpana*, but notes that SaN *oab’ba* ‘little sister’ is possibly a hypochoristic diminutive of a form that is attested in the North Saami form *oar’ben* ‘sister’ that has been attested in the collections of the Norwegian Lappologist Friis.

A further problem is the origin of the *-na* suffix in the Finnic and Saami words. T. Itkonen (1983: 356) assumes that the word *\*orpana* belongs to the vocabulary that is inherited from the Proto-Finnic/Finno-Saamic proto-language, but he does not comment on the Mordvin word. T. Itkonen’s view is echoed by Hahmo (1994: 150–151), who notes that the Saami and Finnic forms with *-n* can both continue a similar derivative that was formed in the common proto-language of Finnic and Saami. Hahmo also mentions the possible connection of *orpo* and *orpana* but does not take a clear stance on the issue. Hahmo agrees with Nirvi that the Mari word is probably a separate borrowing but she supports the connection of Mordvin *urvä* to the Finnic and Saami forms. The Mordvin form *árwänä* is a later derivative, as *nä* is a productive diminutive suffix in Mordvin, and

---

<sup>31</sup> I am grateful to Christopher Culver for pointing this out to me.

also the underived form appears in Mordvin. Zhivlov does not mention the Mordvin word *uřvã* at all.

It remains unclear what the function of the Finno-Saamic *n*-suffix in the word *\*orpana* was, and whether the Finnic and Saami forms actually represent a common innovation, or if the Saami word is an early loan from Proto-Finnic. The word is attested in most Saami languages and it is reconstructed for Proto-Saami as *\*ōrpēnē* by Lehtiranta (2001: 94–95). However, in Finnic the form *orpana* is only attested in Finnish, which makes it doubtful that it goes back very far in time. Also a form *orvas* (< PFi *\*orpas*) exists in Finnish dialects, with meanings that are near those of *orpana*. It is of course possible that both underived *orpa* and the derived form *orpana* existed already in Proto-Finnic (or earlier), and the reflexes of *orpana* have later become lost in other Finnic languages.

SSA assumes that the Finnic suffix reflects a suffix of the Indo-European donor form, connecting the ending *-ana* to Greek *-avos* in ὀρφανός, but this is a very ill-founded idea. The Greek form ὀρφανός is a late derivative with the productive suffix *-avo-* (Chantraine 1933: 196–198; Beekes 2010 s.v.), and it is not possible that the Indo-European donor language of the Finnic/West Uralic word had a form akin to this.

Due to the semantic difference, it remains uncertain whether this word is a borrowing from Indo-Iranian or not.

? Indo-Iranian etymology

**Fi *osa* ‘part, luck’ (found in all Finnic languages); SaN *oažži* ‘flesh’ (< PSa *\*ōńćē*, cognates in all Saami languages except A); Mari *užas* ‘part’ (< PMA *\*užas*); Mansi North *ūńś* ‘buttocks’, East *uńś*, South *ońći*, West *wuńś* ‘das Hintere’ (< PMs *\*ūńć*); Hu *ágýek* ‘loins’**

< PU *\*ońća* (Aikio 2015b: 61)

← PIE ? *h<sub>2</sub>onk-o-* or ← PII *\*Hanća-*, > OI *ańśa-* M ‘share of a fortune, loot; troop, party; N pr. a son of Aditi (he who has many shares)’, root *ńás-* (*ańś-*) ‘erreichen, zuteilen’, Av *ąsa-* ‘Anteil’, a hapax legomenon (Y 31.2.); ? Greek ὄγκος ‘Masse, Gewicht’ (RIVELEX I: 1; Hintze 2000: 32–33; Mayrhofer 2005: 17)

(T. Sköld 1955–1957; Joki 1973: 298; Gamkrelidze & Ivanov 1984: 924; Katz 1985: 359; 2003: 285; Rédei 1986c; Lushnikova 1990: 172; Koivulehto 1991: 107; 1999c: 230 = 2016: 232)

This is a well-accepted etymology, but the Ugric comparanda in the Uralic set were only recently discovered by Aikio (2015b:61; the Hungarian and Mansi words were considered reflexes of a different Ugric word in earlier research such as the UEW, s.v. *anćV*). As the cognates are phonologically regular, this word has to have been borrowed already into Proto-Uralic.

Koivulehto (1988a: 292; 1991: 107) has also assumed that this word could be a Proto-Indo-European borrowing from a form with a retained palatal stop, but this is impossible to prove, and as all of the examples of a palatovelar reflected in the Uralic loans are more

or less problematic, a proper Indo-Iranian borrowing from a form that already had an affricate \*č is more probable.

It is more problematic to determine whether the donor form had \*o or \*a. In theory, PU \*o can be equally well explained from PIE \*o or PII \*a, and Gamkrelidze & Ivanov (1984: 924) and Lushnikova (1990: 172) have argued that this is an “\*o-loan” from PII because \*a is substituted by \*o in front of \*n. There are not very many many examples to support this rule among the Indo-Iranian loanwords. A similar rule has been suggested for Germanic loans in Finnic (Kallio 2012b: 230, footnote 8; LÄGLOS s.v. *onki*; see also the entry \**onki*), but in this loanword layer the rule is supported by much more examples.

Semantically the Indo-Iranian word would be suitable origin. According to Hintze (2000: 31), the Vedic word’s meaning is usually ‘Anteil (= part)’, but the word is also used three times in the Rigveda to describe ‘a good thing, a richness given by the God Indra’ (“der Gottheit (Indra) gegebenen Gut”).

The word-initial laryngeal is not reflected here, as is the case with almost all other convincing Indo-Iranian loanwords. If the word is an earlier Indo-European borrowing, the lack of a laryngeal is even more disturbing. Perhaps the word could be borrowed from Pre-Indo-Iranian with retained \*o: in the Pre-II \*e loans the reflexes of laryngeals are not found either (see \**ertä*). Koivulehto (1988b: 292) notes that it is possible that the Indo-European word had no laryngeal, but the IE form was \**onko-*.

? Indo-Iranian etymology (can also be an earlier loan)

**Md E paz, pas ‘Gott’, M pavas ‘Gott; Glück’ (< PMd \*pavas)**

**< Pre-Md \*pakas (~ \*payas) ‘god’**

← PII \**bhaga-s* or PI \**baga-s* ‘god, luck’, PL \**bagāsas*, > Av *baya-* ‘Gott’, OP *baga- id.*, OI *bhāga-* ‘Herr, Zuteiler, Beiwort von Göttern; Wohlstand, Glück, Besitz, Vermögen’, root OI *bhaj-* ‘(zu)teilen, (als) Anteil geben/bekommen’, Av *baz* < PII \**bhag/ž* < PIE \**bhag-* ‘als Anteil bekommen’ (Rudnyčkyj 1974; EWAia s.v. *bhāga-*; Werba 1997: 209–210 No. 127 s.v. *bhaj-*; LIV<sup>2</sup> s.v. *bhag-*; Hintze 2000: 38–44; Zimmer 2012: 136)

(FUV: 135; Korenchy 1972; Joki 1973: 301, No. 112; Katz 1981; Katz 1985: 223–224, 2003: 181; K. Häkkinen 1987: 245; Helimski 2000: 192; Koivulehto 1999a: 227 = 2016: 229; Blažek 2005: 166–167)

This is a convincing Indo-Iranian etymology, and the borrowing has to be ancient, although the word is found only in Mordvin. Rédei (1986c) considers Erzya *paz/pas* and Moksha *pavas* as two different borrowings from the same Indo-Iranian source, but it is more probable that they reflect one and the same Mordvin word, with loss of the original \*-v- and vowel contraction in *paz/pas*. It is interesting that another word for ‘god’ found in the neighbouring Western Uralic languages is also an Indo-Iranian borrowing, and it also has a petrified reflex in Mordvin (see \**juma*). Joki considers the word to have been borrowed from “ar. bzw. miran.”, but there is no reason to consider this a Middle Iranian borrowing (Joki’s idea was probably due to the limited distribution of the word). The non-

initial syllable *-as* could not result from a Middle Iranian language, as PI *\*-as* was vocalized at the Old Iranian stage already.

K. Häkkinen (1987: 254) criticizes Rédei's idea of dating this word to Proto-Finno-Permic. Rédei assumes (correctly) that the Mordvin word has to be a relatively early borrowing, but Häkkinen considers it as methodologically weak to assume that the word was borrowed into a Finno-Permic stage, as there is nothing in the distribution of the word to suggest this. Furthermore, as the Finno-Permic and Proto-Finno-Ugric stages are phonologically identical, the borrowing could just as easily be considered a Proto-Finno-Ugric (= Proto-Uralic) borrowing as a Proto-Finno-Permic one.

Korenchy considers the word an Iranian borrowing. She argues that the ending *\*-as* would still have been retained in Proto-Iranian. While this probably is correct, it does not automatically mean that the word has to be Iranian. If the word is confined to Mordvin, a later, Iranian borrowing is more probable than a really old borrowing. The fact that the words reflecting the ending *-as* often have a limited, western distribution could point to their Proto-Iranian origin.

The Mordvin semantic development is interesting, as both in Iranian and Mordvin the meanings of 'god' and 'luck' are found. In Iranian, the meaning of 'god' is secondary, as the Indo-European verbal root meant 'to distribute'. Zimmer (2012: 136) criticizes the idea of Katz (1981) that a PIE word *\*b<sup>h</sup>agos* 'God' could be reconstructed on the basis of Iranian *\*baga-* and Slavic *\*boǫǣ*. Zimmer also notes (op. cit., footnote 12) that the meaning 'luck' has developed later in Indic, and it cannot be reconstructed for Proto-Indo-European, even though this is sometimes done.<sup>32</sup> Rudnyčkyj (1974) notes that in Slavic a similar development as in Avestan has took place, as the Proto-Slavic word *\*divъ* 'god' has developed a meaning 'bad spirit', and a new word for 'god' has developed from *\*baga-*. This is in similar vein as in Zimmer's argument, and against the idea of Katz that a PIE word *\*b<sup>h</sup>agos* 'God' can be reconstructed on the basis of Slavic and Iranian. Kallio (p.c.) notes that because of the lack of Winter's law in the Slavic word, it cannot be cognate to the Iranian word but has to be a loan.

The Iranian word has also been borrowed into Tocharian, but in the meaning of 'share': PT *\*pāka* > Tocharian A *pāk* B *pāke* (Tremblay 2005a: 424).

Korenchy mentions (North) Mansi *puuŋ* 'Reichtum, Vieh; Reich' as a possible cognate of the Mordvin word. Although Korenchy is aware of the irregular relationship of the Mansi and Mordvin words, she argues that it might be possible to connect the two forms through hypothetical Proto-Mansi ablaut. As the word-initial consonantism does not match, this idea is impossible. The Mansi word also cannot be an independent loan from Iranian because of the *\*ŋ*, which cannot correspond to Iranian *\*g* or *\*γ*. This etymology has also been criticized by Liimola (1973). The origin of the Mansi word remains uncertain.

## An Indo-Iranian etymology

---

<sup>32</sup> A semantic parallel for the Indic development is provided by Uralic *\*ońća*, the reflexes of which mean 'part' in many languages, such as Finnish *osa*, but in Karelian *osa*, Olonets Karelian *ozu* denote 'luck'.

**Fi *paksu* ‘thick’ (also in Karelian, Veps, Ludic, Votic, Estonian) < PFi *\*paksu***

← PI *\*dbadzu-* ‘vastag, nagy’ > OAv *bəzuuant-* ‘thick’, der. *dəbqzah-*, Baluči *baz* ‘tense’; OI *báhu-* ‘much, large, thick’, root *baṃh*, Av *dbqz* < PII *\*d(h)bʰanʒʰ* < PIE *\*d(h)bʰngʰú-*, root *\*d(h)bʰengʰ-* ‘dick, fest machen’ (EWAia II: 220–221 s.v. *bahu-*; LIV<sup>2</sup>: 76; NIL 13–15; Werba s.v. *baṃh*; Pronk 2013: 1–7)

(Schindler 1963: 203; Koivulehto 1999a: 220–221; 2001b: 256 = 2016: 222–224; Holopainen, Junttila & Kuokkala 2017: 116, 119)

This old etymology by Jochem Schindler was revived by Koivulehto in the late 1990s. Schindler derived the Finnic word from a Pre-Indo-Iranian form *\*bhagʒʰu-*, which is impossible, but Koivulehto was more precise to argue that the word should have been borrowed from Proto-Iranian with a substitution *\*ks* ← *\*dz*.

Although the etymology seems convincing at first sight, it includes several phonological problems. This is the only genuine example where Uralic (Finnic) *-u* corresponds to Indo-Iranian *-u*: here the Finnic *u* cannot be explained away as a derivational suffix. However, it is dubious whether we can trust that *\*u*-stems were possible at such an early date based on such scanty evidence. Holopainen, Junttila and Kuokkala (2017) argue that *-u* could have been possible before the Early Proto-Finnic period, as there are some Baltic and Germanic borrowings which seem to support this, but here the problem is again that the examples are so few that no far-reaching conclusions can be made. It is also interesting that some other words, especially *\*počau*, seem to point to the opposite conclusion, that in non-initial syllables *\*-u* was not allowed and was substituted by a diphthong/sequence of an illabial vowel and a labial consonant.

As is shown by Old Avestan evidence, the Proto-Iranian word had a consonant cluster *\*db-* (Pronk 2013: 2), which is not reconstructed by Koivulehto. As loanwords into Uralic always simplified the word-initial consonant-clusters, *\*p-* is the expected outcome here.

The postulated substitution of PI *\*dz* (in the word-medial position) by *\*ks-* is also problematic, as it is supported by very few parallels; see the entries on *\*maks-* and *\*teksä*. However, the etymology is semantically too good to be completely rejected and one has to admit that the Indo-Iranian and Finnic words bear a striking resemblance to each other, but because of the phonological problems with the sound substitutions mentioned above, it remains uncertain.

? Indo-Iranian etymology

**Fi *peijas*, *peijainen*, Est *peied* ‘wedding, (dial.) funeral’, dial. *peig*, SEst *peie*, Li *peijed* ‘funeral’ (only in Finnic) < PFi *\*peije-* (SSA s.v. *peijas*)**

< Pre-Fi *\*peji-*

← Pre-II *\*paHyya-* > OI *-pāyya-*, OI *péya-* (masc.) ‘Trinkopfer’ (neutr.) ‘Getränk’ in compounds: *pū́ rva-pā́ ya* ‘first drink’, root PII *\*paH* > OI *pā/pī*, Av *pā* ‘to drink’ < PIE *\*poh<sub>3</sub>-* (Werba 1997: 303 No. 316 s.v. *pā/pī*)

(Koivulehto 2005: 329–331; LÄGLOS III s.v. *peijas*)

The etymology is semantically convincing, and as Koivulehto has convincingly shown, there is no other plausible etymology for this Finnic word. A Baltic origin from a hypothetical form *\*pa(i)yas* has been suggested by Ritter (1977), but the Indo-Iranian etymology is better, as the exact source form is not attested in Baltic. Because of the vowel *\*e*, this has to be a very early loan, even though Koivulehto argued that the word could be a loan from Proto-Indo-Iranian, with Finnic *-ei-* reflecting the sound change *\*ai > \*ei*. In the light of the modern understanding of Finnic historical phonology, this cannot be true, as this change has not affected Livonian and South Estonian, which means that this is not a Proto-Finnic sound change. Kallio (2014: 159–160) argued that this is a “Gulf of Finland Finnic” innovation, meaning that this change happened in the predecessor of all Finnic languages other than South Estonian and Livonian. Kallio (p.c.) also notes that because the change was more precisely *\*ai\_a > \*ei\_ä*, the expected outcome would be *\*pejäs*, not *\*peijas*. As Proto-Finnic *\*peijas* has a reflex in Livonian *pejēd* (pl.), Koivulehto’s assumption cannot be correct, and the word has to be borrowed from an Indo-Iranian form with *\*e*. This does not solve all the problems with the vocalism, as *\*ei* should have developed into long *\*ii* in Finnic. Also a reconstruction *\*päijäs* would not be regular, as although the change *\*äi > \*ei* is attested in both Livonian and in Gulf of Finland Finnic (the proto-language of all the Finnic languages except Livonian and South Estonian), the conditions of this change are not known, and it should not happen in this environment as words like *\*päjwä* and *\*äijä* have stayed intact (Kallio 2018a: 261–262).

The Finnic ending *-as* is typical of loanwords, and many Indo-Iranian loans were borrowed as Uralic *-as* stems. There are various views on when *as*-stems became possible in western Uralic (see Helimski 2000: 191–195) for discussion), but it is usually assumed that they were not possible in Proto-Uralic, as no *s*-stems with a Uralic distribution are found. If the word *\*peijas* is indeed a Pre-Indo-Iranian loanword, it shows that the *s*-stems must have been allowed quite early in Uralic, as it is not possible that the word was borrowed later into Proto-Finnic.

Also a Germanic etymology has been suggested, *\*faigjaz* (> ON *feigr* ‘doomed to ‘die’). LÄGLOS mentions both Germanic and Indo-Iranian etymologies, but Kallio (2013: 216) sees the Indo-Iranian origin as more convincing.

An Indo-Iranian etymology

**Ko** *pež-* ‘to bake, fry, roast, scorch (sun)’; **Ud** *pīži-* ‘to bake, fry, roast; scorch (sun); to get tanned; to burn, hurt’ < PP *\*pōži*

< Pre-P *\*poči-* or *\*poši-*

← PII/PI *\*pač-a-* ‘to boil’ > YAv *hqm.pacaiti* ‘boils’ < PIE *\*pek<sup>w-</sup>* (LIV<sup>2</sup> s.v. *\*pek<sup>w-</sup>*; Cheung 2007: 286–287)

(Hyllested 2014: 17–19; Metsäranta 2017: 230)

Hyllested (2014) and Metsäranta (2017) have offered new ideas for the Indo-European origin of this Uralic/Permic verbal stem. Koivulehto (1991) derived the Uralic word from

Proto-Indo-European, with \*š (as reconstructed by Koivulehto) as the substitution of PIE \*H. In Hyllested's survey of laryngeal etymologies he finds Koivulehto's explanation unlikely and instead argues that the word is borrowed from some satem language, and Uralic \*š would rather reflect the sibilant \*š than a laryngeal. This is phonetically certainly more convincing. Metsäranta digs deeper into this etymology and argues that precisely the Indo-Iranian verb \*pač- would be a fitting original (with \*č from PIE \*k through the second palatalization). Metsäranta notes that both \*š and \*č could be reflected by the modern Permic words, and it is possible to determine which one has the substitution of Iranian \*č here. The consonant \*č would be phonetically expected, but one has to note here that there are only very few examples of the substitution of \*č.

An Indo-Iranian etymology

**Fi *paras* 'best', *parempi* 'better'; SaN *buorri* 'good' < PSa \**pōre*; Mo E *paro*, M *para*; Ma *poro*, *porâ* 'good' (< PMa \**porə*-); Ko, Ud *bur* 'good, right' (< PP \**bur*)**

< PU \**pīra* or \**para* (UEW s.v. *para*; Sammallahti 1988: 553 \**pīra*; Zhivlov 2014: 127)  
 ← PII \**parHa-s* > OI *párah* 'ferner; jenseitig; früher; später; fremd, feindlich', etc. (EWAia II: 86, s.v. *pára*, 88, s.v. *parás*)  
 (Koivulehto 1999a: 229 = 2016: 231; ms.)

This is a convincing etymology. The same Indo-Iranian word has probably been borrowed as \**pora* with different vowel substitution and semantic development into Pre-Saami and Pre-Permic (see the entry \**pora(s)*). Both etymologies are convincing, and it is difficult to account for the different vowel substitution here. Semantically both 'good' and 'old' can be explained from the Indo-Iranian form.

The Uralic words are regular cognates, so the word can be an old loan into Proto-Uralic. Zhivlov reconstructs the word as \**para*<sub>1</sub>. Permic *b-* is due to secondary voicing, which is possibly caused by the medial voiced consonant, \**r* in this case. Sammallahti (1988) reconstructs the word as \**pīra*. This is noted also by Koivulehto (ms.), who argues that this is one of the loans where *i* is the substitute for PII \**a*. In the reconstruction of Aikio (2015b), the Permic vocalism Ko, Ud *u* can be regularly derived from both of the PU combinations \**i*-*a* and \**a*-*a*. Aikio does not mention this word in his list of back-vocalic Uralic words, however.

The Indo-Iranian form \**parHas* goes back to full-grade PIE form \**perHos*, but there are also zero-grade forms attested in both Indo-Iranian and other languages, as Koivulehto (1999a) notes, such as Old Indic *puráh* 'voran, vorn, davor' < \**pr̥Ho-*, Greek πάρος 'früher, voran, vor' < \**pr̥Hos* or \**parHos* (according to Beekes [2010: 1154 s.v. πάρος], the Greek word reflects the PIE zero-grade form \**pr̥Ho-*). Gotō (2013: 149) notes that adverbs like *puráh* look like old ablative forms of root nouns. As noted by Holopainen (2018b) it is possible to assume that the Pre-Saami word \**poras* reflects the zero-grade form of the PII or PIE word. However, it could also be assumed that \**or* in the Saami form is a substitution for PII \**a* in the labial environment, as this substitution has many parallels. The more widespread Uralic word \**pīra*, on the other hand, could then reflect the form with *r̥H*. The different substitutions could then be explained by loans from different ablaut grades.

This explanation is, however, semantically not ideal. The meanings of ‘good’ and ‘old’ can both be derived from the attested Indic forms, but there is no compelling reason to assume that the meaning ‘old’ would come from an adverb which means ‘before’ and ‘good’ from the one meaning earlier. The development postulated above thus remains hypothetical.

Bereczki (2013 s.v. *poro*) does not mention the Indo-Iranian etymology. EES mentions an alternative etymology which has been suggested by Katz (1985, 2003: 155), namely that the Uralic word is borrowed from Pre-II *\*bhadrom* (= PII *\*bhadram*) ‘glücklich’. This is phonetically and semantically less convincing than Koivulehto’s etymology, and one can only wonder why the editors of EES have decided to include Katz’s etymology here (the actual reference to Katz is missing, as EES does not provide references to individual etymologies).

An Indo-Iranian etymology

**Ko, Ud *purt* ‘knife’ < PP *\*purt*** (Csúcs 2005: 372)

< Pre-P *\*parta* or *\*pirta*

← PI *\*parata-* (?), cf. OI *paraśu-*, *parśu-* ‘Beil, Axt, Streitaxt’; Oss *færæt-* (Zimmer 1990: 8; EWAia II s.v. *paraśu-*; Cheung 2002: 186)

(Joki 1973: 305–6; Rédei 1986c: 76; Lushnikova 1990: 241; Csúcs 2005: 372)

This old etymology is, in principle, convincing – the similarity of the Indo-Iranian and Permic words could hardly be caused by accident, and a word for ‘knife’ is a typical cultural borrowing. However, the etymology of the Indo-Iranian words is not clear at all. Mallory (2015) has noted of Tocharian *\*paratwä*, the alleged cognate (?) of the Indic word that it is an “infamous Wanderwort”. It is often assumed that the Tocharian word was borrowed from Iranian (Tremblay 2005a: 423; Isebaert 1980; Peyrot 2019). These words in various Indo-European languages are often connected to the Semitic words like Akkadian *pilaqqu* ‘Stilett, Spindel’, and it looks probable that these words were ultimately borrowed from the Semitic languages of the Near East. Helimski (1997b: 121) assumes that also the Permic word is a Wanderwort.

Nevertheless, as there are large numbers of Indo-Iranian cultural loans in Permic, the most obvious source of the Permic word is indeed Indo-Iranian, and this makes sense also phonologically. Permic *t* could be explained either from Ossetic-type *t* or *θ* that is attested in some Iranian forms. Indo-Aryan *ś* in relation to Iranian *t* and *θ* is irregular, but whatever the origin of the Iranian consonantism is, it suits the Permic words better. The word is also not attested in the neighbouring Turkic languages, which would be another obvious source for a Permic cultural word. The more precise dating of the borrowing is difficult, but the word has to be older than Proto-Permic because of the vocalism. The word could theoretically be a late loan from Alanic: as Permic has no *f*, an Ossetic-type form could be the source, but the Alanic fronted vocalism would probably have become Pre-Permic *\*ä* rather than *\*a*.

The Indo-Iranian and other Indo-European words refer to ‘axe, hatchet’, but ‘knife’ is not very far from this semantically. It is also possible that the various meanings attested in Udmurt compounds reflect a more original usage of the word as a cutting-tool. A parallel development can be observed in the semantics of Evenki *purta* ‘knife’ which according to Rédei and Joki is borrowed from the same Iranian form as the Permic word, although it cannot be ruled out that this word is borrowed from Permic (see the discussion in Joki). Cincius (1975–1977 s.v. *purta*) only mentions the Udmurt and Komi words in connection with Evenki *purta*, but it remains unclear what is here considered as the exact relationship between the Evenki and Permic words, as it is not explicitly stated that the word is a loan. No mention of the Iranian words is made. The meaning ‘knife’ appears in some later Indo-Aryan forms that are listed by Joki, such as Sinhala *pihiya*, *pihāya*.

No other serious alternative to this Indo-Iranian etymology has been presented to explain the origin of these Permic words. The alternative etymology of Katz (1985: 334; 2003), from “frühurarisches” *\*k<sup>w</sup>rtim* is not convincing, as there are no reliable parallels for the substitution *\*p- ← \*k<sup>w</sup>*.

An Indo-Iranian etymology

**SaN *buohta* ‘opposite to, just in front of’, cognates found in all Saami languages < PSa *\*pōteḱ* (Lehtiranta 2001: 110–111)**

**< PSa *\*puoteḱ* < Pre-Sa *\*pati* or *\*poti***

← ? PI *\*pati*, > OAv. *paīti*, YAv *paiti*, OP *patiy* ‘gegen, entgegen’ (IEW: 842; EWAia II s.v. *prāti*; Dunkel 2014: 655ff. s.v. *\*prōti*)

(Sammallahti 2001: 404; Holopainen 2018b: 165)

This etymology is phonologically possible, and semantically the Saami preposition and the Iranian adverb are close to each other, but the limited distribution of the Saami word raises some questions. If the word really is a borrowing, it should be from Proto-Iranian or some later Iranian language, as the preposition *\*pati* does not have a formal parallel in Indic. Iranian *\*pāti* and Indic *prāti* ‘gegen, entgegen’ continue PIE *\*proti* (Dunkel 2014: 655, s.v. *\*prōti*), with an irregular dissimilation in Iranian. However, a similar dissimilation is found also in Greek dialects, where forms with both *\*pr-* and *\*p-* are attested (according to Dunkel, these forms are “isofunktional und räumlich komplementär verteilt”), as well as in Hittite, Latin and Lithuanian. It is uncertain whether this is an independent, parallel development or whether this difference between *\*proti* and *\*poti* existed already in PIE and was levelled in Indic. Derksen does not mention the Lithuanian enclitic particle *-pāt*. Theoretically a Baltic origin would be possible too, but as the word is better attested in Iranian as an individual adverb, it is more probable that the Saami word, if it is a loan, is borrowed from Iranian.

? Indo-Iranian etymology

Fi *piimä* ‘sour milk, Est *piim* ‘milk’, also in Karelian, Votic and Izhorian; ? Md E *ped’a-*, *päd’a-*, Mo *ped’a-* ‘seihen, durchseihen; melken’; ? Hu *fej* ‘to milk’

< ? PU *\*pejmä*

← Pre-II *\*peyHmṇ-* (> PII, PI *\*payHman-*) ‘thick fluid; milk’ > Av *paēman-* ‘mother’s milk’; derived from the PIE root *\*peyH-33* ‘to be thick; to swell’ (EWAia II: 83–84; Garnier et al. 2017: 300; LIV<sup>2</sup>: 464–465, s.v. *\*peḷH-*; Werba 1997: 356–457 No. 424 s.v. *pyā(/pyai)*)

(Munkácsi 1901: 263, 597; Joki 1973; 302; Lushnikova 1990: No. 121; Katz 1985: 237–238; 2003: 192; Helimski 1991: 221 SSA II s.v. *piimä*; EES s.v. *piim*; UEW s.v. *\*pije-* [378] & *\*päð’e-* [359])

Katz considers the Finnic, Mordvin and Hungarian words cognates, and assumes a Pre-Indo-Iranian borrowing: ← *\*peyma-*. From the borrowed Uralic word the verb *\*pej-* would then be a back formation: *-ma* would have been reanalysed as a suffix. Then the Mordvin verb *ped’a-* would have been further derived from the retrograded verbal stem with the frequentative (?) suffix *-tA-*. Hungarian would continue the bare stem *\*pey-*, and Finnic *piimä* would continue the original, borrowed noun.

The idea that the Finnic, Mordvin and Hungarian words mentioned above are borrowed from Indo-Iranian is an old idea, but because of problems with the vocalism of these words both the Indo-Iranian etymology and the mutual relationship of these words has remained highly uncertain. Aikio (2014: 90–91) has recently explained that the development *\*ej > \*ii* is regular in Finnic (cf. also Kallio 2018a: 262–263), and he has argued that this word can easily be borrowed from a Pre-Indo-Iranian form that is reflected by Avestan *paēman-* ‘mother’s milk’. Fi *tiine* is another etymology that shows a similar development and it has also been dealt with by Aikio (see the entry *\*tejniš* in this work).

Also a Baltic origin for the Finnic word has been suggested by Larsson (1984: 129–140), and his etymology has been supported by SSA and EES. Cognates of the Indo-Iranian root *\*payH-* are attested in Baltic: the noun *\*peinas* ‘milk’ is reflected as *pienas* in Lithuanian and *piens* in Latvian, and there is also a verb *pýti* ‘to get milk’ in Lithuanian. None of these forms would be a suitable origin for the Finnic word, but Larsson assumes that also a noun *\*piyimas* could have existed in Baltic, and Finnic *piimä* could be borrowed from this. Larsson also argues that the Avestan word *paēman-* reflects the PIE diphthong *\*oy-*, but according to the recent research of Garnier et al. this is not the case, rather the Avestan word can be safely derived from *\*peyHmṇ-*. The Baltic etymology has been doubted by Junttila (2012: 275) because the actual source is not attested. Although technically possible, Larsson’s etymology can now be considered wrong with good reason, as the Finnic vocalism can be regularly derived from *\*-ej-* and there is no reason to postulate a hypothetical Baltic origin. The remark in SSA that the hypothetical Baltic origin is better for phonological reasons is, of course, outdated now, and EES’s argument that the limited distribution of the word within Finnic would support the Baltic origin is untenable, as the

---

<sup>33</sup> According to Garnier (2015; Garnier et al. 2017: 301) the original PIE root is *\*(s)peh<sub>1</sub>-*, from which through analogy and metathesis *\*(s)peyh<sub>1</sub>-* developed.

word is in any case attested in several Finnic languages, and both the Baltic and Indo-Iranian loans have very variable distribution within Finnic.

Although the Indo-Iranian etymology of the Finnic word can be considered certain now, its relationship to the suggested Mordvin and Hungarian cognates is less clear. Deriving the Mordvin word verb from this Indo-Iranian source is more difficult because of its consonantism (see below), but the Hungarian verb *fej* ‘to milk’ could reflect the same proto-form as the Finnic word (the PU diphthong *\*ej* would be reflected as *ej* in Hungarian). However, the problem here is that the Finnic word is a noun and there is no trace of a verb derived from the same stem, and in Hungarian only a verb is attested and there are no signs of a noun *\*pejmä*. It is of course possible that *\*pejmä* was borrowed into Proto-Uralic as a *nomen verbum*, but it is very difficult to prove this kind of claim correct.

Recently Honti (2017: 95–97) and Róna-Tas (WOT; 2017: 62–63) have discussed the etymology of the Hungarian word. Honti has rejected the Iranian etymology, in my view without sufficient reason (both Honti and Róna-Tas seem to be unaware of Aikio’s recent work on the Finnic word). Honti considers possible the etymological connection of the Mordvin and Hungarian words, but not their relationship with Finnic. Róna-Tas thinks along the same lines with Honti, and further states that it is problematic that in Iranian the word is only attested as a noun, whereas the Hungarian word is simply a verb. While it is true that there is no exact Iranian correspondence to the Hungarian verb ‘to milk’, meanings that come rather close to it are attested in various Indo-Iranian forms, such as Vedic *payate* ‘oozes with milk’ and Avestan *pipiiuš-* ‘bringing milk’, cited by LIV<sup>2</sup>. Cheung (2007) even reconstructs the meaning of the verbal root *\*payH-* as ‘to lactate, (breast-)feed’ in Iranian. And a semantic parallel is also provided by Finnish *lypsää* ‘to milk’ and Mordvin *E lovso*, *lovco* *M lofca* ‘milk’, which are probably etymologically connected, although they are not regular cognates but borrowed from the same (unknown) source (Aikio 2015b: 46).

Honti has also wondered whether a word denoting ‘milking’ can have existed already at Proto-Uralic/Finno-Ugric times, but this argument against this etymology is not very efficient. Also hunter-gatherers could have practised milking, and if the word denoted simply ‘milk’ as a noun, it is natural that the people have had some kind of knowledge of milk, as they certainly hunted for mammals and were acquired with at least some sort of domesticated animals (like the dog), and naturally Proto-Uralic-speaking mothers produced milk too.

The UEW does not take a clear stance on the issue, as it tentatively reconstructs two stems, one based on the Finnic word, the other on the Mordvin, and the Hungarian word is listed under both entries. However, both of the UEW’s assumptions might be wrong, as all of the three words here might simply reflect parallel borrowings from (Indo-)Iranian at different times and different regions. The Finnic word obviously reflects the source of Avestan *paēman-* ‘mother’s milk’, as argued above, whereas a reflex of a different formation, PII *\*payHas-* (> Av *paiias-*, OI *payas-* ‘milk’ cf. EWAia; Garnier et al. 2017: 301–302) could be a fitting source for the Mordvin and the Hungarian words. Of course, the borrowing of this word has to have been rather early too because of the *e*-vocalism that is reflected by both Mordvin and Hungarian.

The Hungarian word could also be a later Iranian borrowing, acquired independently. Here Hu *e* could result from earlier *\*ä*, from a later, more palatal reflex of Iranian *\*a* that is found in some Middle Iranian languages. However, the Mordvin word cannot reflect *\*ä*, as

the correspondence Erzya *e*, *ä* ~ Moksha *e* points to an early *\*e-i*-stem. This means that the reconstruction *\*pãδ'e-* presented by the UEW has to be wrong.

There are also other arguments against this reconstruction by the UEW. In Mordvin, the *-j* of older diphthongs often disappears, cf. such cases as *sed'*, *säd'* 'bridge' from PU *\*sejti* and *vano-* 'to see' from Indo-Iranian *wayna-* (Koivulehto 1999a: 229). If one supposes that the *-d-* in the Mordvin word can be analysed as a part of a suffix, we could assume that the Mordvin forms would reflect Pre-Mordvin *\*pej-* which lost the *\*j* in the position before the suffix. A similar, although not entirely identical, case is provided by Mo *vad'ems* 'to smear', if this word can be derived from the PU word *\*waji* 'butter' (Aikio 2014a: 3). Here both Mordvin languages retain the reflex of the PU noun which actually has the diphthong intact (E *vaj*, M *oj*), but we could perhaps assume that in the verb *vad'ems* the *\*j* has been lost in the pre-suffixal position. Admittedly, the development of diphthongs in Mordvin still requires further research.

To sum up, it can be stated that there are no reasons to reject the Indo-Iranian etymology of the Finnic word. Also the Iranian origin of the Hungarian word is convincing, and deriving the Mordvin verb from the same Iranian source looks possible.

An Indo-Iranian etymology (possible parallel borrowings into Pre-Finnic, Pre-Mordvin and Pre-Hungarian)

**Ko *bord(j)*, Ud *burd* < PP *\*börd* (Csúcs 2005: 322)**

**< Pre-P *\*peträ* ~ *pertä* 'wing'**

**← Pre-II *\*petro-* 'wing' > PII *\*patra-*, root *\*path-* 'to fly, to fall' > OI *pát(t)ra-m* 'wing', root *pat<sup>i</sup>-* 'to fly' (EWAia II s.v. *pátra-*; Werba 1997: 301 No. 312)**

(Katz 1985: 260; Koivulehto 1988a: 43, 51; 2001b: 290)

Koivulehto's etymology for the Permic word is convincing. The metathesis is also a convincing explanation, as *\*-tr-* in early loans was substituted by different means because of Uralic phonotactics. It is interesting that in the borrowing *\*takra* ← *\*dātra-* the substitution is different – perhaps this is due to the fact that *\*takra* was borrowed from Proto-Indo-Iranian or Proto-Iranian, whereas *\*peträ* has to be a Pre-Indo-Iranian loan because of the vowel *\*e*. Technically the word could be borrowed from Proto-Indo-European or some other Indo-European language too. Of course Indo-Iranian would be the most obvious choice for Permic, but it is troubling that this word is found only in Permic despite being a very early borrowing. It is possible that the word for 'wing' has disappeared from the other languages. Theoretically it is also possible that the word was borrowed later from Iranian into Permic, and the Iranian vowel *\*a* has been substituted by some vowel that later gave the *o* ~ *u* correspondence (see *\*kertä* for similar arguments). Note that the archaic loans *śo* ~ *śu* 'hundred' and *ozir* ~ *uzir* 'rich' manifest the same vowel correspondence.

The geminate *-tt-* in some Indo-Aryan forms is secondary (EWAia). Interestingly, old *\*-rt-* clusters often show secondary germination in Permic (cf. *\*mertä*, *\*kertä*), but in this word the *-rd* cluster points clearly to *\*-rt-* and not *\*-rtt-*. It is unclear whether the words

with *-rd* and *-rt* in Permic belong to separate loanword layers. Both *\*ertä* and *\*pertä* point to retained Pre-II *\*e*, so they could be earlier loans than *\*mertä* and *\*kertä* which were probably borrowed from forms which had *\*r̥*.

The Komi word is a secondary *j*-stem (Metsäranta 2017: 229, footnote 26). Usually the *j*-stems reflect Uralic *\*i*-stems.

An Indo-Iranian etymology

### **Fi *perna* ‘spleen’, Est *põrn* (: *põrna*) id. and regular cognates in all the other Finnic languages**

< ? PFI *\*perna*

← PIR *\*spr̥dzan-* or later I *\*spr̥zan-*, > Av *sp̥ar̥azan-* ‘spleen’ ? < PII *\*spr̥žhán-*

? > OI *p̥līhán-* id. (EWAia II: 196–197, s.v. *p̥līhan-*; de Vaan 2008: 340)

? < PIE; cf. Greek *σπλήν*, Latin *lien*, OCS *slězena*, Middle Breton *felch*, ? Armenian *p’aycatn* id. (Steer 2015: 218; Beekes 2010: 1384–1385).

(Katz 1985: 156; 2003: 130; Koivulehto 2003, ms.; Häkkinen 2004; EES s.v. *põrn*)

Koivulehto’s etymology was not published in his lifetime as such but it is found in the handout of his 2003 presentation to the Finno-Ugrian Society, and also in the unpublished manuscript on Indo-Iranian loans. The etymology appears in printed form in the etymological dictionary of K. Häkkinen (2004). EES mentions the etymology too but does not refer to Koivulehto’s handout (the information was probably copied from Häkkinen’s dictionary).

Semantically the etymology is very good. The Finnic word is probably a loanword, as no cognates in the related Uralic languages are found. Besides the Indo-Iranian etymology, no compelling explanation for the origin of the Finnic word exists; Liukkonen (1999: 104–105) has attempted to derive the Finnic word form Baltic *\*sp̥ernā* ‘wing’, but this etymology is semantically very unconvincing.

However, the reconstruction of the Indo-Iranian word is difficult, and even the Sanskrit and Avestan words cannot regularly be derived from an unitary source, and there are problems with the relationship of the other Indo-European forms with similar meaning, as well. This makes the assessment of the etymology somewhat problematic, even if the Proto-Iranian or later Old Iranian forms which can be reconstructed on the basis of Avestan would be fitting originals for the Finnic word.

Precisely the Iranian words look like a more convincing source, as the prehistory of the Indic words remains obscure. De Vaan (2008: 340) reconstructs PII *\*spl̥žhán-* (which would be *\*spr̥žhán* if we assume PII *\*l > \*r*), and from this the Avestan word can be regularly derived, but the Sanskrit form remains mysterious, as the long *ī* cannot be regularly derived from this reconstruction. The apparent Proto-Italic cognate *\*lixēn-*, from which Latin *liēn*, is irregular according to de Vaan (he notes that this form has possibly influenced by taboo).

Recently Steer (2015: 187) has reconstructed the PIE form (on the basis of Latin and Old Indic) as *\*splig<sup>h</sup>ē n-* (with a question mark). Frisk (1954–1972: 769–770, s.v. σπλήν) is supportive of the Indo-European origin of the Greek word and the other mentioned forms, but he admits that a common reconstruction is impossible and large assumptions (“blosse Vermutungen”) have to be made about the development of the words in the daughter languages. The Indic and Latin forms continue similar stem formation and vocalization according to Frisk. Also Beekes (2010: 1384–1385) notes that the relationship of the various words in various Indo-European languages is unclear, that is, the words look similar but are too different for a unitary PIE form to be reconstructed. Martirosyan (s.v. *p’aycaln*) is more optimistic and argues in favour of reconstructing the word back to PIE. He notes that the Armenian word with *-y-* might correspond regularly to the Indic and Latin words which also reflect a form with PIE *\*i*.

Besides the problems of Indo-European, also the Finnic side poses some problems. Problematic is the Proto-Finnic vowel *\*e*, as this phoneme does not appear in inherited Uralic vocabulary but is probably a Finnic innovation. There are also some other tentative Iranian loanwords for which the vowel *\*e* has to be reconstructed such as *Fi verso*, *Est vörse* (< ? *Pfi vērša* ← *PII \*wr̥ká-*; see the entry). If the etymologies for *\*perna* and the other loans showing this vocalism are correct, it forces us to think that some loanwords were acquired into Proto-Finnic after the genesis of this vowel, or that this vowel has developed from some earlier Pre-Finnic vowel phoneme. Kallio (2014: 161) notes that Proto-Finnic *\*meļa* is probably from earlier *\*melä*, as the Mordvin and Saami cognates clearly point to the latter pre-form. This means that the vowel combination *\*e-a* does not exclude the possibility that the word existed already in Pre-Finnic.

Phonetically, there are no problems in the vowel substitution for Iranian syllabic *r̥* or from its later reflex *\*ər* by the *Pfi* central-vowel *\*e*. The substitutions of the other segments involve no problems: the cluster *\*-rtsn-* or *\*-rzn-* would have been simplified somehow on the Uralic side, so *\*-rn-* is a logical outcome. The word is a Finnic *\*a*-stem like most of the Iranian loanwords are. The Iranian *n*-stem is acquired without a substitution of the word-final consonant (cf. other examples such as *aćma* ← *\*(H)aćman-*). Indic *l* is obviously secondary (see the Introduction for arguments about the change *\*l > \*r* in Indo-Iranian). Also Katz (2003: 130) notes that *l* is secondary in this Indic word.

An Indo-Iranian etymology

**Khanty North, *pǎnt*, East *pěnt*, South *pěnt* ‘Weg (der Menschen oder Tiere)’ (< PKh ? *\*pint*)**

← Iranian *\*pantH-* : *\*paθ-*, compare Avestan *paθō* ‘Pfad, Weg’, OI *pánthā-*, *pathī-*, *path-* ‘Weg, Pfad, Bahn’ (EWAia II: 81–82 s.v. *pánthā-*), from PIE root *\*pent-* ‘to find a way’ (Pinault 2012: 405–406)

(Korenchy 1972: 62, No. 16; Katz 1985: 313; 2003: 250; UEW s.v. *pǎntV*; Lushnikova 1990: 241; Koivulehto ms.)

The Khanty word is most probably of Indo-Iranian (or other Indo-European) origin. Korenchy assumes that the loan is “Middle Iranian”. The Komi word *pad(-vež)* ‘crossroads’

presented as a possible cognate by Korenchy and the UEW is not a regular cognate of the Khanty word but could be a parallel loan from the same Iranian source (this option is also mentioned by Korenchy, UEW and Katz [2003]). The Iranian etymology is supported by Koivulehto (ms.). An alternative etymology for the Komi word is suggested in Appendix I.

Korenchy reconstructs the Proto-Khanty form as *\*pəntV*. The Khanty words seem to reflect Proto-Khanty *\*pint* if the system/reconstruction of Zhivlov (2006) is applied, which in turn usually goes back to Proto-Ob-Ugric *\*e* or *\*ä* in Zhivlov's system (see Zhivlov 2006: 113–116, 134–136). So this word could belong to the layer of loans where PII or PI *\*a* has been substituted by *\*ä* in Pre-Khanty/Ob-Ugric. As the *\*ä* substitution seems to have prevailed for a long period in the Ugric languages, it is difficult to determine whether the vowel here reflects the fronted *a* which is found in Ossetic *fændag*. The attested Middle Iranian forms, Khotanese *pande* and Middle Persian *pand* do not point to the fronting of the *a*.

An interesting parallel is Germanic *\*paþa-* 'path' which is borrowed from the Iranian languages of the steppe (Mayrhofer 1970; Kroonen 2013 s.v. *\*paþa-*). According to Mayrhofer, the "Nomadenpfad" was an important concept in the steppe, and it is possible that the word borrowed into Khanty is from the time when the linguistic ancestors of the Ob-Ugrians came under the cultural and linguistic influence of the Iranian groups living in the Eurasian steppe.

An Indo-Iranian etymology

**Mansi North *pānt*, East *peñt*, South *pāntē*, West *pəñt* 'der Mann einer älteren Schwester' < ? PMs *\*pñtV***

← Iranian *\*band-*, compare Avestan *band-* 'to bind', *banda-* 'Bande, Fessel (bond, fetter [Kroonen: < PIE *\*bʰondʰo-*])' OP *baṛdaka-* 'Vasall, Gefolgsmann', MP *band-/bāstan* 'to bind', *band* 'Verbindung, Bindeglied', *bandag* 'Diener', Khotanese *bañ/basta-* 'to bind', Sogdian *βynd/βyst-*, Parthian *bynd/bst-*, Chwarezmian *bncj̄/bstyk-* id.; OI *bandh-* 'to bind', *bāndhuḥ* 'Verbindung, Verwandtschaft etc.>'; root PII *\*bʰandʰ* < PIE *\*bʰendʰ-* (EWAia II: 208 s.v. *BANDH*; Werba 1997: 359–360, No. 430 s.v. *bandh*)

(Munkácsi 1901: 153–154; Korenchy 1972: 63, No. 17; Katz 1985: 222; 2003: 180; Lushnikova 1990: 190; Helimski 1991: 221; 1997)

Korenchy reconstructs "Middle Iranian" *\*band-* as the original. From an Indo-Iranian point of view even Proto-Iranian would be phonologically suitable, and it is difficult to determine the age of the borrowing in this respect. The attested Middle Iranian languages show both *a* vocalism and front-vowels (such as Sogdian *βynd*). The Mansi vowels seems to continue Proto-Mansi *\*ĩ*. A similar substitution is found in the etymology for *\*wĩsəy* 'calf'. Korenchy explains the relationship of the vocalism of the Mansi forms through hypothetical Proto-Mansi ablaut.

The semantic connection between the Mansi word and the Indo-European word family is reasonable. The meaning 'son-in-law' is not attested in Indo-Iranian, but Korenchy provides parallels from other Indo-European languages where kinship or social terms have

been derived from the root *\*b<sup>h</sup>endh-* ‘to bind’: Greek *πενθέρος* ‘father-in-law’ and Lithuanian *beñdras* ‘comrade’. Also the words for ‘vassal’ and ‘servant’ attested in the Iranian languages are semantically not very far from terms for ‘in-law’, although one has to admit that specifically ‘elder sister’s husband’ is not exactly what one would expect to develop from ‘servant’ (‘son-in-law’ or the like would make more sense).

Phonetically the palatalized *\*ń* in the Mansi word weakens the etymology, and it cannot be considered certain. This was also noted by Korenchy herself, and later by Helimski (1991: 221; 1997), who considers the etymology to be unconvincing. There are some examples of palatal *\*ń* as the substitute for Indo-Iranian *\*n*, but in the most convincing case *\*sirańa* this can be explained by the Indo-Iranian cluster *\*ny*.

An Indo-Iranian etymology

**SaN *boazu* ‘reindeer’ (< PSa *\*poacōj*; cognates in all Sa languages), Mari *püčə*, *pučə* ‘reindeer’ (< PMA *pučə*); Ud *pužej* ‘reindeer’ (< PP *\*pužV*)**

< PU *\*počaw*

← PI *\*patsu-* > Avestan *pasu-* ‘livestock’ < PII *\*pácu-* ‘cattle, animal’ < PIE *\*pk<sup>h</sup>-eu/\*pk<sup>h</sup>-u*, root *\*pek<sup>h</sup>-* ‘(Wolle oder Haare) rupfen, zausen’ (EWAia s.v. *pásu-*; LIV<sup>2</sup> s.v. *\*pek<sup>h</sup>-*; Bjorvand & Lindeman 2000: 210–211 s.v. *fe*; Pinault 2012: 400; Kölligan 2018: 2262)

(Koivulehto 2007: 251– = 2016: 382ff.)

This is a convincing Proto-Iranian etymology, showing that Koivulehto’s palatal criterion (U *\*č* ← PI *\*ts*) works well at least in the word-internal position. Semantically the etymology is convincing: the Uralic words refer to ‘reindeer’, but this can be explained from the Indo-Iranian meaning ‘cattle’ without difficulties. The Indo-European *u*-stem noun originally denoted ‘sheep’<sup>34</sup>, but the meaning ‘livestock, cattle’ is found in both Indic and Avestan. Häkkinen (1999: 167) has noted that it is uncertain which animal the Uralic word exactly referred to, and how tamed this animal was. Based on the attested meaning of the present-day Uralic languages the meaning ‘reindeer’ could be reconstructed for the proto-language already.

The Uralic words can be derived from an old *\*o*-*a*-stem, so there is no reason to assume parallel borrowings. There are slight irregularities in the Mari word. In most etymological sources (Bereczki 2013 s.v. *püčə*), the front-vowel *ü* is considered secondary, and this explanation is also mentioned by Koivulehto. The Indo-Iranian etymology is not mentioned by Bereczki (2013). The Saami and Udmurt words could also reflect an earlier *\*a*-*i*-stem, but the Mari *u* vocalism points rather to Uralic *\*o*, as there are not many examples of PU *\*a* > Mari *ü*.

Finnish and Karelian *poro* ‘reindeer’ is mentioned as a possible cognate by the UEW, and this is supported by Koivulehto, who argues that in *poro* the weak-grade of *\*t*, *r* in

<sup>34</sup>According to Pinault (2012: 400), the IE noun was an original hysterokinetic collective. For arguments against alternative etymologies of the PIE noun, see Bjorvand & Lindeman (2000: 210–211, s.v. *fe*).

some western dialects of Finnish, has become analogically present in the nominative form too. Koivulehto supposes that the word has spread from *r*-dialects to other dialects of Northern Finnic. However, this is probably not a correct scenario, as Kuokkala (2012) has shown. Kuokkala argues that the Finnic word is borrowed from Saami, and that the *r* is simply a substitute for the Saami affricate.

The non-initial syllable is particularly interesting, as here a Uralic sequence *\*-aw* can be reconstructed, resulting in the labial vowel of Saami. As the Iranian form ended in *-u*, and *-aw* in some inflected forms, it seems that the speakers of Uralic substituted the labial element of the Iranian word by a sequence of a vowel and a labial semi-vowel, either because the labial vowels were not allowed, or because the sequence *\*-aw* reflects the Iranian *\*-aw* of the inflected forms such as the Avestan genitive *pasəuš*. The origin of the Udmurt *-ej* is unclear, but due to regular Permic apocope it seems unlikely that this reflects the PU suffix.

The Ob-Ugric words for ‘reindeer’ are parallel loans from Iranian (see the entry *\*pāččäy*) according to Koivulehto. In the UEW these words were listed as possible cognates, but this is impossible as the Ob-Ugric words cannot be derived from an old *\*o-a*-stem. Abondolo (1996: 72–73) assumes that the words in Saami, Finnic, Mari and the Ob-Ugric languages are cognates. He bases this idea in his assumption that there was paradigmatic vowel alternation in Uralic, and the Ob-Ugric forms descend from a “light” isotope, whereas the rest reflect a “heavy” isotope. The Permic words are not included in Abondolo’s set, for some reason (he does not comment on their absence). Abondolo’s idea is based on a theory that is not accepted usually, and Koivulehto’s argument that the words reflect parallel borrowings is much more convincing.

UEW also lists similar words from other language families in Northern Eurasia (Yukaghir *pədzə*, *pəd’e* ‘Elentier’; Yakut *büčən*, *bičən* ‘кабарка мускусная (Moschus moschiferus)’, Evenki *bičən*, Even *büčəke*, *büčən* ‘косуля’, Ulch *bočan*, Oroch *bučan* ‘изюбр’), so we might be dealing with a Wanderwort here, and although the Iranian etymology is phonologically convincing, it cannot be ruled out that the Uralic word is borrowed from some other unknown source which could also explain the forms in the Siberian languages.

An Indo-Iranian etymology

**SaN *buohča* ‘penis of a man or a horse’ (cognates in East Saami and SaLu; < PSa *\*puoče*), Hu *fasz* ‘penis’**

< *\*piši* or *\*poši* ‘penis’ (UEW s.v. *\*pačs*; Sammallahti 1988: 548; Zhivlov 2016b)

← PII *\*pásas-* ‘penis’ > OI *pásas-* < PIE *\*pés-os-*, > OI *pásas-* (EWAia II: 111, s.v. *pásas-*; Stüber 2002: 197–198; Clackson 1994: 101, footnote 76)

(Munkácsi 1901: 258–259; Takács 1997: 373; Koivulehto ms.)

Although the etymology was suggested already by Munkácsi, it is missing from all Hungarian etymological dictionaries. Likewise, sources on Indo-European loans in Saami do not mention the etymology (Sammallahti 1998, 1999, 2001). Takács (1997) has briefly

discussed the relationship between the Uralic and Indo-European words, assuming that the word in Uralic is probably an Indo-European borrowing, but he does not note the discrepancy of the sibilants. This etymology is a rare example of the PII sibilant \*s being substituted by something other than simply Uralic \*s. Hungarian and Saami both clearly point to palatal \*ś. A borrowing from another branch of Indo-European would not solve the problem of sibilants, and because of the \*e-vocalism of the PIE word, Indo-Iranian is the best source for the Uralic word. It is possible to assume that some kind of dissimilation took place in Uralic, but this is pure speculation. Koivulehto (ms.) suggests that affection is the reason for the irregular palatalization here, and this is probably the best explanation.

Otherwise the etymology is convincing. Semantically the etymology is plausible, and it would be strange if the similarity between the Uralic and Indo-Iranian words was mere coincidence.

The Indo-Iranian word has a plausible Indo-European etymology (the root \*pes- is widely attested, and the Greek s-stem πέος is an exact parallel to the Indo-Iranian word), so a borrowing in the reverse direction is impossible. According to Takács, there are similar words in several language families of Eurasia, and Takács assumes that all of them are borrowed from Indo-European.

The etymology has also been considered as a Nostratic inheritance (Illich-Svitych 1984: 96–97, No. 371; Zhivlov 2016b). Borrowing is a more likely option, as the Nostratic hypothesis is not very convincing in general, and it remains uncertain how could the Uralic and Indo-European forms be derived from the Proto-Nostratic reconstruction \*/pʰ/a/se.

The UEW's ambiguous reconstruction \*paćz is misleading. The Uralic word can be either an old \*i-i-stem or an \*o-i-stem. As \*o often appears after \*p-, this is more likely solution here.

Regarding the vowel substitution, the Uralic vocalism can be reconstructed as either \*e or \*o: in Hungarian, reflexes of \*o and \*a often merge as a or á (there are very few exceptions where Hungarian shows \*o > o instead), and in Saami \*e-i and \*o-i-stems have similar reflexes. The old reconstruction with a is impossible, as Aikio (2015b) has shown that \*a-i-stems become \*oa-e-stems in Saami. Although the vowel cannot be decided based on the Uralic evidence only, it is clear that in loans usually PII \*a becomes Uralic \*o after word-initial \*p-, meaning that this word should probably be reconstructed as \*pośi.

An Indo-Iranian etymology

**Md E *pango*, M *panga* ‘Pilz, Schwamm; Mari *pongo*, *pongâ* ‘Pilz, Fliegenpilz, Baumschwamm, Meltau; dumm’ (< PMA \**pongə*); Mansi East *pēŋk*, West *pēŋk*, North *pāŋx* ‘Fliegenpilz; Rausch’ (< PMs \**pŷŋkə*), South *pēŋl* ‘betrunken sein’ (< PMs (der.) \**pŷŋl*); Kh East *pāŋk*, South *paŋx* ‘Fliegenpilz’ (< PKh \**pāŋk*), East *paŋkəl* ‘singen, nachdem man Fliegenpilze gegessen hat’; Nganasan *hʷaŋkudʷa* ‘betrunken sein’ (< PS \**peŋkâ*-)**

< PU \**piŋka* ‘psychedelic mushroom’ (UEW s.v. *paŋka*; Zhivlov 2014: 119, 133; Aikio 2015b: 59)

← PII or PI \**b<sup>(h)</sup>anga-*, ? > MP *bang*, *mang* ‘Bilsenkraut’, ? Av *baŋha-* ‘name of a plant and a narcoticum that is made of it’ (EWAia II: 240–241, s.v. *b<sup>h</sup>aŋgá-*)

(Munkácsi 1903; Korenchy 1972: 64 [No. 19.]; Joki 1973: 300–301; UEW s.v. *paŋka*; Rédei 1986c: 74–75; Lushnikova 1990: 190–191; Bereczki 2013: 194, s.v. *poŋgo*)

This is an old etymology, but its acceptance has varied over the years, and the matter has been complicated by the uncertain relationship of the Uralic words that have been grouped under this proto-word. Also the etymology of the assumed Indo-Iranian donor form is highly uncertain (see below).

Proto-Indo-Iranian (and Proto-Iranian) *\*a* has been substituted by Proto-Uralic *\*i* in a number of loanwords. Mari *o* can regularly reflect Proto-Uralic *\*a*, but if the Ob-Ugric and Samoyedic words are cognates of the Mari and Mordvin forms, the Proto-Uralic word has to be reconstructed with the high back vowel *\*i*. However, it is not regular to derive Mari *o* from Proto-Uralic *\*i*; it is not totally clear what the regular reflex of the latter sound in Mari is, but there are numerous examples of Mari *ü* such as *šüdö* < *\*šita* and *üpš* < *\*ipsi*. Zhivlov (2014: 133) has suggested that Proto-Mari *\*o* is the regular reflex of Proto-Uralic *\*i* before *k* and *ŋ*, but there are not many examples of this: *\*soŋka* ‘old’, which is also a loan from Indo-Iranian, is one possible parallel (see the entry). The Mordvin vocalism is ambiguous in this respect, as *\*a* and *\*i* merged in *a*-stems.

Korenchy considers these words to be separate Iranian borrowings in individual Uralic languages (she derives the Khanty and Mansi words from Proto-Ob-Ugric). Bereczki connects the Ob-Ugric, Mari and Mordvin forms with a question mark and leaves the possible Permic and Samoyedic cognates out. Aikio (2013a: 11) has considered the Nganasan word a possible loan from Ob-Ugric; in his latest publication (2016: 49) he does not explicitly state this, although he considers the Samoyed form’s vocalism to be irregular. The Khanty and Mansi words are regular cognates, and thus probably inherited from some earlier proto-language; Zhivlov (2006: 151) reconstructs the Proto-Ob-Ugric form as *\*p̄iŋka* (> Proto-Mansi *\*p̄iŋkə*; *\*p̄iŋkl-*, Proto-Khanty *p̄aŋk*).

The Komi word family *pagalni* ‘die Besinnung verlieren, befangen werden’, *pagir* ‘sauer; scharf, durchdringend’ cannot be cognate to the rest of the words because Komi *a* cannot reflect PU *\*a–a* or *\*i–a* and it might be a later parallel loan from Iranian, as was suggested already by Korenchy. Helimski (1997b: 121) mentions that the Iranian origin of the Komi word is doubtful. Csúcs (2005: 365) lists the Komi word as a Permic word (PP *\*pag*) but does not mention the possible Finno-Permic/Uralic origin. The Iranian etymology is mentioned with two question marks: Csúcs gives “miran. *\*bang* or *\*pang*” as the donor forms without providing further references. The origin of the Komi word requires more research: the vowel *a* does not fit well with the idea that *\*g* could reflect earlier *\*ŋg* (from Indo-Iranian *\*ng*).

Zhivlov (2014: 133) has objected to the view that the Samoyed word could be a loan from Ob-Ugric. The Samoyed (Nganasan) verb is derived without the suffix *-l*, which is present in the Ob-Ugric verb forms that are derived from *\*p̄iŋka* and agree in meaning with the Samoyed verb. Zhivlov states that it is unlikely that the suffix would have been lost in Samoyed, if the word had been borrowed from the Ob-Ugric verb. However, the Samoyed vowel still remains unexplained, and it might be possible that the speakers of Samoyed would have borrowed the noun from Ob-Ugric and then derived the verb ‘to get intoxicated’ from it later.

If the etymology is correct and Zhivlov is right about the relationship of the Samoyed and Ob-Ugric words, this is one of the few Indo-Iranian loanwords which has a reflex also in the Samoyed branch. Finding a solid example of this would be quite remarkable, as the

distribution of Indo-Iranian loans has been one of the main arguments for the binary split of Finno-Ugrian and Samoyed.

The Indo-Iranian word has no known cognates in other Indo-European languages, which poses some problems for the etymology (EWAia II s.v. *bhaṅgá-*). The Sanskrit word *bhaṅgá-* is probably not related to Avestan *baṅha-*, nor to Middle Persian *bang*, *mang*. Middle Persian homonymous *bang* is borrowed from Indic. Avestan *baṅha-* could also reflect earlier *\*basa-*. The Indo-European words for ‘mushroom’, such as Latin *fungus*, which are mentioned by Korenchy and Bereczki do not belong here etymologically. However, if a form *\*bhaṅga-* in the meaning of narcoticum or the like existed in Indo-Iranian, the Uralic word can be borrowed from this, although this is far from certain.

Katz (1985: 172–173; 2003) has argued that the Uralic word was borrowed from Indo-Iranian because there are no cognates to the Indo-Iranian word in the related languages. While this possibility cannot be ruled out, one has to note that there are also other cases where Uralic has acquired an Indo-Iranian loanword which does not have any cognates (in some cases it is more obvious that this has to be the direction of the borrowing). Obviously the Proto-Indo-Iranians could have borrowed words from other sources by the time they came into contact with speakers of Uralic.

An Indo-Iranian etymology

**Fi *pohja* ‘bottom, ground; north’, Est *põhi* id., cognates in all Finnic languages**

< Pre-Fi *\*pošja* (SSA II s.v. *pohja*)

← PII *\*pasčā* > OI *paścā*, Av *pašča* ‘after, behind’ (EWAia II s.v. *paścā*)

(T. Sköld 1999; LÄGLOS III s.v. *pohja*; Kallio 2013: 216)

The Indo-Iranian etymology for this Finnic word is not convincing. Also two competing Germanic etymologies have been suggested (PG *\*bandsja-*, cf. *\*bandsa-* > ON *báss* ‘Kuhstand’ and PG *\*bosjōn-*, cf. PG *\*busjōn-* > Old Swedish *bysia* ‘straw’) although both are rejected by LÄGLOS in favour of the Indo-Iranian origin (it is noted in LÄGLOS that *\*bandsja-* is unattested and *\*bosjōn-* is formally impossible, as the Germanic word should be reconstructed as *\*busjōn-*). Kallio (2013: 216) is critical of the Indo-Iranian etymology, which fails to explain the word-internal consonantism. The Indo-Iranian etymology is also semantically not very good. The Indo-Iranian form reflects PIE *\*pos-ti-*, which is reflected by a number of forms in Indo-European languages, such as Lithuanian *pàs* ‘bei’ and Old Church Slavonic *pozde* ‘after’, but these Baltic and Slavic forms are phonetically and semantically likewise impossible origins for the Finnic word. A Proto-Indo-European source is also ruled out for the same reasons.

Kallio agrees with the criticism of the Germanic etymologies which is presented by LÄGLOS, and it seems that the origin of Finnic *pohja* remains unclear for now. Saami *buoššu*, which mentioned by SSA as a possible loan from Finnic, has to be a loan and cannot be a cognate because of the Saami š.

Not an Indo-Iranian etymology

**SaN *boaris* ‘old’ < PSa *\*boares*, cognates in all Saami languages; ? Ko *periš*, Ud *peres* ‘old’ < PP *\*pōriš***

< Pre-Sa *\*pora(s)* or *\*pari(s)* (UEW s.v. *\*por3*)

< PII *\*parHa-s* > Old Indic *pārah* ‘**ferner; jenseitig; früher; später; fremd, feindlich**’ (EWAia II: 86, s.v. *pāra*, 88, s.v. *parás*)

(Blažek 1990: 42; Koivulehto 1999a: 229 = 2016: 231; 1999b: 331; Sammallahti 2001: 408; Holopainen 2018b: 163–164)

Koivulehto argues that this word is borrowed from the same source as *\*para(s)* (see the entry). In his view the two borrowings do not reflect two distinct stages of Indo-Iranian, but they manifest different substitutions, *\*o* and *\*a*, both of which are known substitutions for PII *\*a*. Koivulehto only discusses the etymology of the Saami word *boaris*, but in the earlier research Komi *periš* and Ud *peres* ‘old’ have been connected here, and the UEW derives the Saami and Permic words from a Proto-Uralic (Finno-Permic) proto-form. As noted by Holopainen (2018b), the Permic words cannot be regular cognates of the Saami word, and it is better to consider them parallel loans. The Saami word can reflect either an earlier *\*o–a* or *\*a–i*-stem according to the new vowel reconstruction of Aikio (2015b). Here *\*o–a* is more probable, as *\*i*-stems are very rare among the early loans. The *\*o*-vocalism can be explained by the labializing influence of word-initial *\*p-*.

It is more difficult to determine what the pre-form of the Permic words is. According to the UEW, the Permic vocalism is due to delabialization after *\*p-*. The Permic words could also be derived from Pre-Permic *\*pārā* (Metsäranta: forthcoming). *\*ā* is the substitute for Iranian *\*a* in many loanwords in Permic, and it is probable that the Permic word is a parallel loan.

The Indo-Iranian etymology for these words meaning ‘old’ is convincing. The Indo-Iranian adverb *pāra-s* has a lot of polysemy, at least based on the Indic reflex *pārah*. However, the meaning ‘old’ can be easily derived from this, as a meaning ‘earlier’ is attested for the Indic word, and also the Greek cognate *πᾶρος* shows a similar meaning. There are actually several closely related forms in Indo-Iranian, and it is difficult to determine from which word the Saami and Permic words have actually been borrowed. Formally the Saami word could be derived from either *\*pr̥Ha-* or *\*parHa-*. See above the entry *\*pīra* for a more detailed discussion.

Saami *-s* can be either a reflex of Indo-Iranian *-as* or a Saami adjective-forming suffix. Komi *-iš*, Udmurt *-eš* have to be later adjectivizing suffixes, as is noted by the UEW.

Indo-Iranian etymology (Saami and Permic words parallel loans)

**Mansi North *por* ‘Name einer wog. Sippe’; Khanty East, South *pär-*, North *pör-*, *pär-* ‘Name einer der beiden Phratrien, in denen die nordostj. exogamenges od. Sippen zusammengefaßt sind’**

**< POUG ? \**por* ‘name of an Ob-Ugric phratry’**

**← PI \**parHa-* ‘farther, previous’ (see \**pīra* for references)**

(Katz 1985: 212; 2003: 172; Rédei 2004a: 130–131)

This etymology has been suggested independently by Katz and Rédei. Both argue that the Ob-Ugric name for the phratry *por* could be borrowed from the same Indo-Iranian donor form as the more widespread Uralic word \**pīra* ‘good’. The semantic development would be that the name of the phratry has developed from a more general meaning of ‘foreign, alien’. Rédei refers to the fact that of the two Khanty phratries, *por* is the one that is usually considered inferior, and it would be possible for the name of this phratry to develop from a negative adjective used for them. For the name of the other phratry, see the entry \**mańéV*.

The Ob-Ugric vowel correspondences are not regular, and the word is not found in the list of Ob-Ugric reconstructions of Honti (1982) or Zhivlov (2006). Whatever is the origin of the Ob-Ugric word, it is clearly a late loanword which has diffused from language and dialect to other.

Due to the fact that large part of the vocabulary of the Ob-Ugric languages is of unknown origin, and the semantic development that Rédei assumes is hypothetical and difficult to prove, the Indo-Iranian etymology remains uncertain.

? Indo-Iranian etymology

**Fi *porsas* ‘piglet’ (cognates in all Finnic languages); ?Md E *purtsos*, M *puŗts* id.; Ud *parś*, *parís*, Ko *porś* ‘pig’**

**< Pre-FI/PU \**porćas* ~ \**porśas* ‘pig’**

**← ? PII \**parća-s*/Pre-II \**porćo-s*, Av *pərəsō* (= \**parəsō*), Khotanese *pāsa-* (Zimmer 1990: 6–7; 1999: 36–37; Schmitt 2009: 18)**

or

**← PBSI \**porśo-s* (Kallio ms.)**

(Setälä 1900: 31–32; Paasonen 1908–1910b: 76; Jacobsohn 1922: 136–223; Benveniste 1949; Mayrhofer 1952: 13; KESK: 226; Joki 1959, 1973, 1988: 585; Burrow 1976b: 62; FUV: 135; UEW s.v. *porćas*; Mayrhofer 1984: 251, footnote 4; Gamkrelidze & Ivanov 1984: 921; Katz 1985: 255–256; 2003: 205–206; Lushnikova 1990: 242; Helimski 2000: 192; Koivulehto 2001b: 242 = 2016: 284; Csúcs 2005: 365; Hyllested 2014: 84–85; Kallio [ms.]; SSA II s.v. *porsas*)

Although this etymology was discovered already by Jacobsohn, and since then the ultimate (and transparent) Indo-European origin of the Uralic words has been widely accepted, there are a number of difficulties with this etymology, and until now there has not been consensus on whether this is a borrowing from Indo-Iranian or some other branch. This word has been traditionally considered an Indo-Iranian borrowing, but for some decades now it has been customary to suppose another IE (often PIE) origin for at least the Finnic word (see Koivulehto 1991). This is due to the (tentative) assumption/idea that PIE palatal *\*k'* can be substituted by *\*ć* in early loans in Uralic, first presented by Joki (1959: 52). Already Benveniste (1949: 87) assumed that the Uralic word could be derived from an earlier form of Indo-European rather than Indo-Iranian.

This idea has since found additional support from Koivulehto (1991, 2001b), but in the end this support is not as strong as it might seem: there are a number of problems in many of the etymologies with retained *\*k'*, and most of them can actually be explained as borrowings from satem languages. The second syllable vowel is rather ambiguous, as it is doubtful whether *\*-os* would have been a possible substitution here – at least there are no plausible examples of such substitution in PIE loans. The vowel in the first syllable tells little of the age of the word (*contra* Dolgopolsky 1989: 21), as *\*o* appears frequently as a substitute for PII *\*a*, especially in labial environments.

Gamkrelidze & Ivanov (1984: 921) argue that the word is a specifically Iranian borrowing, with *\*a* being substituted as *\*o* before *\*r*. The vowel substitution may be correct, but the word cannot be Iranian because of Uralic *\*ć*, as here one would expect Uralic *č* then. The *o* here can be explained otherwise, as the word could have been borrowed from a language which still retained PIE *\*o*, although it is true that *o* as the substitution of *\*a* quite often appears before *\*r*, and Gamkrelidze and Ivanov seem to be right with this rule. However, this rule mostly appears in earlier Proto-Indo-Iranian loanwords, and their idea that only the specifically Iranian loanwords manifest this substitution should be abandoned. Gamkrelidze & Ivanov do not comment on the phonological irregularities between the various Uralic forms. Lushnikova (1990: 32) mentions that the *o*-substitution can be due to the labial environment.

Mayrhofer (1952: 12, footnote 1) supports the possibility that the Uralic word is a Pre-Indo-Iranian loan, but he notes that the Uralic words should rather be borrowed from Baltic, where the word is well attested. The Uralic word cannot be a true Baltic borrowing because of the *\*ć* (*\*š* would be expected if the word was really borrowed from Baltic, and this can be seen in South Estonian *pahr* which is a later borrowing from Baltic *\*paršas*). Napol'skikh (2002) and Kallio (ms.) consider the Finnic word a Proto-Balto-Slavic borrowing. The word is well-attested in Balto-Slavic, and as mentioned above, the substitution of the sibilant/affricate (from PIE *\*k'*) is regular, as is the vowel substitution (regardless of whether *\*a* or *\*o* is reconstructed for the stage of Balto-Slavic from which this word was borrowed). Kallio also presents a number of other Balto-Slavic borrowings, which serve as parallel evidence.

Burrow (1976b: 62) was in favour of the Indo-Iranian etymology because of the wide attestation of the word within Uralic (he argues that the Baltic origin could not result in this). Burrow was obviously unaware of the possibility of early contacts with Proto-Balto-Slavic: an early Balto-Slavic borrowing could result in a similar, wide distribution. Regarding the *\*o* vocalism, Burrow argues that the borrowing is so early that *\*o* had been retained in the donor language, but his explanation differs from the mainstream view: he

argues that \*o, when standing in an ablaut relationship with \*e, did not merge with \*e and \*a but developed differently into long \*ā in Proto-Indo-Iranian (this is only partly correct in the light of modern studies of Indo-Iranian historical phonology, as \*o developed into \*ā in open syllables only).

Other satem branches cannot come into question as donor languages due to geographical reasons. The word is widely attested in centum branches, such as Italic (Latin *porcus*), Celtic (Old Irish *orc*) and Germanic (German *Ferkel* and its cognates), and Koivulehto's idea was that the Finnic word is borrowed from Northwest Indo-European, a predecessor of Germanic and other "northern" branches. However, as most of the Northwest Indo-European loans turn out to be rather Indo-Iranian or Balto-Slavic loanwords, and as the ideas of prehistorical contact zones have evolved since Koivulehto's time, the NWIE etymology looks less convincing now – at least there is no possibility to prove that a word like \**porćas* is borrowed from such a language, making the assumption too hypothetical.

Hyllested (2014: 84–85) has recently considered the pig words of Indo-European as borrowings (a "Central Asiatic culture word") from "Altaic" words denoting badger, such as Turkic \**borsuq* 'badger' (> Chuvash *porʒš* id.) and he assumes that the word \**pork*ós should not be reconstructed for PIE. There is not much substance to Hyllested's Altaic arguments, as both the semantics and phonology of the Altaic > Indo-European borrowing are difficult to explain. Hyllested argues that the use of the badger in the diet of the Altaic peoples would explain the semantic shift, but this is highly dubious. Phonologically the etymology is even more difficult: the palatal sibilant/affricate or earlier palatal *k*´ that has to be assumed for the most Indo-European forms probably would not have been the substitute for Turkic *s*. Also Indo-European \**p* from Turkic *b* looks dubious. At least de Vaan (2008 s.v. *porcus*) and Zimmer (1990: 7) consider the Indo-European term for the pig to be derived from the root \**perk*´ - 'to dig up earth',<sup>35</sup> so it would have an Indo-European etymology, and if this is correct, it is a further argument against the "Altaic" origin.

Regarding the Uralic forms, Hyllested assumes that at least the forms in Mordvin and Permic are borrowed from the same Turkic word. Unfortunately, this offers no solution for the intriguing consonantism in these languages, as the Mordvin affricate and the Permic *ś* would not be the expected reflexes of Proto-Turkic *s* or Chuvash *š*.

The Mordvin and Permic words are problematic: they certainly originate in some Indo-European form too, but are not regular cognates to the Finnic word. The Permic \**a* points to a later borrowing. The Mordvin affricate is irregular and disturbing, as noted by Koivulehto (2001b: 242), but this can perhaps be explained through affection.

The Mordvin word is irregular only regarding the affricate/sibilant. One would expect *ś*. There are some cases where PU \**ć* is reflected as affricate in Mordvin, such as the irregular reflex of Uralic \**icā* 'father' (see the entry), but the reasons for this are unknown. According to EES the Mordvin words could have been borrowed from Finnic, but this is an *ad hoc* assumption that does not explain the Mordvin affricate in any way. As there are

---

<sup>35</sup> Zimmer (1999: 36–37) notes that the IE noun is not necessarily from the root \**perk*´ 'to dig up earth', as it can also be derived from \**perk*´ 'to be speckled'. In any case, the noun \**pork*ós can be reconstructed for Proto-Indo-European.

some other cases where \*ć has been irregularly retained as \*ć in Mordvin, we simply have to assume that this belongs to these irregular cases, or that ć is reflected by Mordvin ć in certain environments that are yet to be determined. There are not many examples of the consonant cluster \*-rć- in the Uralic vocabulary, so one could assume that in this environment ć was retained as such, and did not develop into ś/z.

The vowel correspondences between the Permic forms (Ud *parś*, Ko *porś*) are also irregular, if it is assumed that the Udmurt and Komi words are derived from \**porśas*. One would expect Udmurt *u* and not *a*. The Permic word is probably a parallel borrowing too, but it is troubling that it clearly points to a palatal sibilant, which means that it cannot be borrowed from a very late Iranian form (*contra* Csúcs 2005: 365, who mentions Old or Middle Iranian \**parś* [sic!]) as the possible donor form for the Permic words). KESK (226 s.v.) notes that the irregularity in Udmurt vocalism could be due to the influence of \**r*.

? Indo-Iranian etymology (probably borrowed from Balto-Slavic)''

**SaN *boarhti* ‘bark vessel’ (< PSa \**poartē*, cognates in most Saami languages);  
? Ko *pert*, Ud *purti* ‘kettle’ (< ? PP \**porti*; Csúcs 2005: 372)**

< ? PU \**porta*

← PIE \**poh<sub>2</sub>tro-* or PII \**paHtra-* > Old Indic *pā́tra-* ‘Gefäß, Behälter, Schale, Geschirr, Becher’, Av *pāθra-uaunt-* ‘Schutz gewährend’ (IEW: 839; EWAia II: 119, s.v. *pātra-*)

(Katz 1985: 177–178; 2003: 146–147; Koivulehto 1988a: 46; 1999b: 331; Holopainen 2018b: 164–165)

The Indo-Iranian etymology has been independently suggested by both Katz and Koivulehto. This etymology is in principle convincing, but Koivulehto has also argued that the Saami word could be borrowed from Germanic. It may well be that the Saami word was borrowed from Germanic \**fōdra-* ‘fodder’ (> Gothic *fodr* ‘scabbard, sheath, casting’, see Kroonen 2013 s.v. \**fōdra-*) and the Permic words from its Indo-Iranian cognate. Semantically both the Germanic and the Indo-Iranian words fit well the Uralic forms.

The word has been reconstructed for Proto-Finno-Permic as \**portta* by Sammallahti (1988: 553), UEW (s.v. *portta*) and Csúcs (2005: 372), as well as Katz, but both the Permic and the Saami words show secondary gemination of \**rt* according to Kallio (2012b). See also \**kertä*, \**mertä*. It is then probable that the word has been borrowed as \**porta* into both branches.

Here the *o*-substitution can be due to the labial environment (word-initial \**p-*), but \**o* also frequently appears before \**r*. It is also theoretically possible that the *o* reflects Pre-II \**ō* here, but this is more unlikely because of the limited distribution of the word. If the Saami word is indeed borrowed from Germanic, the *o* is simply the substitute for Germanic \**ō*. It is also possible that the Saami form reflects earlier \**a-i*, as the reflexes of \**a-i* and \**o-a*-stems merged in Pre-Saami (Aikio 2015b), but in the light of loanword evidence the reconstruction of a \**a-i*-stem for this word is not very likely.

? Indo-Iranian etymology

**Ma *pūndaš*, *pāndaš* ‘Boden’ (< PMa *\*pūndaš*); Ko *pīdeš* ‘Boden’, Ud *pīdes* ‘Boden, Grund’ (< PP *\*pīdes*)**

< ? PU *\*puntaksi* <sup>36</sup> (UEW s.v. *punt3* (*punt3-ks3*))

← ? PII *\*b(h)ud<sup>h</sup>na-*, OI *budhná-* ‘ground, bottom’ < ? PIE *\*b<sup>h</sup>ud<sup>h</sup>-no-* or *\*b<sup>h</sup>ud<sup>h</sup>-mn-o-* (EWAia II s.v. *budhná-*)

(FUV: 135; KESK: 235; Joki 1973: 304–5; Rédei 1986c; Helimski 1997b: 122; 2000: 192; Katz 1985: 108–109; 2003: 94)

The Indo-Iranian etymology of this word is an old idea, accepted without hesitation by Rédei and Helimski. The etymology is also mentioned by EWAia. However, the etymology is not as unproblematic as it may seem. Semantically the etymology is convincing, but, the medial consonantism is difficult to explain from the Indo-Iranian forms. A cluster *\*-tn-* cannot be found in the Proto-Uralic lexicon, so it is possible that *-nt-* can be explained through metathesis. Rédei and Helimski have postulated an unattested PII form *\*b<sup>h</sup>undas* to account for the Uralic forms, but there is not much substance to this idea. Rédei refers to Pāli *bunda-*, but a Proto-Indo-Iranian development cannot easily be postulated on the basis of a Middle-Indo-Aryan form. The Indo-European cognates that show *-nd-*, Latin *fundus* and Middle Irish *bond*, can probably be explained through metathesis according to EWAia.

The UEW notes that also Mari *pundāš*, *punduš* ‘Baumstumpf’ has been connected here earlier, but because of the different vocalism this word has to be of a different origin. Helimski (2000) is more positive in connecting this other Mari word to *pūndaš*, *pāndaš*, however. He also mentions Mordvin *potmaks*, *potnaks* ‘Boden’ in this connection. The UEW is probably right in arguing that these words cannot be derived from the same PU form. It would be possible to assume that these reflect parallel borrowings, as semantically both the Mari and Mordvin words can be derived from ‘ground, bottom’. The Mari and Mordvin words are not regular cognates to each other, so each of them has to be borrowed parallelly from Indo-Iranian.

An Indo-Iranian etymology

**Mari *pörš*, *pöröš*, *pöršö* ‘Reif’ (< PMa *\*pürš*)**

< ? Pre-Mari *\*pürša*

← PII *\*prušwa-* > OI *pruṣvā́* ‘Reif, Raufrost, Eis’

(Joki 1973: 304; Katz 1985: 85; 2003: 77)

Rédei (2004b: 320) argues that the Mari form shows rhotacism from earlier *\*š*, not metathesis as Katz assumes. In the UEW it is argued that the Mari cluster *rš* continues

---

<sup>36</sup> I owe many of the ideas here to discussions with Niklas Metsäranta.

earlier \*č. Bereczki (2013 s.v., 196–197, s.v. *pörš*) criticizes this idea and notes that the Mari word is of unknown origin, unrelated to the other Uralic forms.

Not an Indo-Iranian etymology

**Ng (der.) *xula*, E Ch *fuřu*, B *fúru*, TN *pul*, FN *pur*, Slk Tym *pəl*, Kam (der.) *pʼulu* ‘bridge’**

< PS \**pul̥* ~ \**pil̥* (SW: 131)

← (Middle) Iranian; cf. MP *puhl* ‘bridge’ < PII/PI \**pr̥tu-*

(Janhunen 1983; Lushnikova 1990: 245; Napol’skikh 2010: 230)

This is one of the few Iranian etymologies that are confined to the Samoyed branch. Because of the *l* the word cannot be a very early loan, so the Middle Iranian dating that is presented in earlier sources is probably correct. From a Proto-Iranian form \**pr̥tu-* one would expect traces of both *r* and *t* in Samoyed. The Proto-Samoyed vocalism (*u* or *ɨ*) fits well with the later Iranian *u*-vocalism, although PU \**u* is retained as PS \**u* in \**i*-stems (Sammallahti 1988: 484), so the vocalism itself does not force us to consider the word a late loan. However, inherited Pre-Samoyed \**l* would become \**j* (cf. PU \**tuli* > PS \**tuj* ‘fire’ [Sammallahti 1988: 540]). According to Napol’skikh the word is a “Middle East Iranian” borrowing, but it remains uncertain what Napol’skikh means with this. The sound change from *-rt-* to *-l-* happened in Middle Persian, which is not East Iranian, so Napol’skikh’s arguments are difficult to understand. As Proto-Samoyed started to diverge during the last centuries before the start of the common era (Janhunen 1998: 457), chronologically a Middle Iranian donor form is suitable.

The word for ‘bridge’ is an obvious cultural borrowing. There was probably no intensive contact between Proto-Samoyed and Iranian, as the borrowings are few and do not necessarily belong all to the same layer.

An Indo-Iranian etymology

**Md E *pušto* ‘zerstampfter Hafer’, *puštā* ‘толокно’; Ma *pušto*, *puštā* ‘dicker Haferbrei, толокно’ (< ? PMa \**puštā*)**

< ? PU

← PII \**pištra-*, > Av *pištra-* ‘Mehl’, *pišant-* ‘zerstampfen’, OI *pišta-* ‘Mehl’, *pináštī* ‘zerstampft, zermalmt’ < PIE *peys* (EWAia; Werba 1997: 206–207 No. 199 s.v. *piš*; LIV<sup>2</sup> s.v. *peis*)

(Joki 1973: 306; Bereczki 1992: 93)

The Mari and Mordvin words are not regular cognates. Mari \**u* cannot be explained from earlier \**i*. It is also uncertain what is the relationship between all the Mari words mentioned here. The Mordvin form does not point to old \**i*, as this would have given \**o* (cf.

\**wiša* > *ožo*). If the Mordvin word is a loan, it should be a rather late loanword then. Semantically the etymology for both words is convincing, but because of the difficulties in the vocalism, the etymology of both Mari and Mordvin words remains uncertain.

? Indo-Iranian etymology

**Mansi North *pāsiγ*, East *pāši* ‘reindeer calf’ (< PMs \**pāššəγ*); Kh North *peši*, East *pečəγ* ‘reindeer calf’ (< PKh \**päčəγ*)**

< POUg \**päččäγ* (Zhivlov 2006: 130)

← PI \**patsuka-* > Avestan *pasu-* ‘livestock’, diminutive *pasuka-*; cf. OI *pásu-*, *pāśuka-* < PII \**paćuka-* (EWAia II s.v. *pásu-*; Gotō 2013: 53, 26, footnote 79)

(Koivulehto 2007: 255 = 2016: 386ff.)

See also \**počaw*

This is a convincing etymology, and a parallel loan from the same Iranian source as the West Uralic/Finno-Permic \**počaw*. Koivulehto (2007) used the reconstruction of Honti (1982), according to which the Khanty and Mansi forms reflect \**e*, but more recently Zhivlov (2006) has reconstructed the vowel as \**ä*, which can probably be a substitute for PI \**a* (or \**ā*, see below). Interestingly this word shows the substitution \**ä* ← PI \**a*, which is also found in most of the other Iranian loanwords in Ob-Ugric. This points to the conclusion that \**ä*-substitution appears in early loanwords, too, in the eastern part of the Uralic language family, not only in later, Alanic-type loans.

The POU affricate points to a Proto-Iranian origin (a similar substitution PI \**ts* → Uralic \**č* as in many Iranian loans in Western Uralic languages, such as \**matsa-* → \**mača*, see the respective entry). Semantically this fits well with some of the other Indo-Iranian etymologies related to cattle breeding and the pastoral economy.

The non-initial syllable of the Ob-Ugric words is open to different interpretations: \*-*γ* could reflect PU \**w*, meaning that this word could show a similar substitution as West Uralic \**počaw*, but on the other hand, \*-*γ* could also be from earlier \*-*k*, meaning that the Ob-Ugrians could have borrowed the word from the derived form \**patsuka-* (Av *pasuka-*, OI *pāśuka-*). A similar \*-*γ* reflex is found also in Mansi \**wšəγ*, which is also borrowed from a word with the same Iranian suffix *-ka*.

An Indo-Iranian etymology

**Mari *piða-*, *piðä-* (< PMa *\*piða-*) ‘to bind’**

< Pre-Mari *\*pänti-*

← PII *\*b<sup>h</sup>and<sup>h</sup>-a-* or PI or later Iranian *\*band-a-* ‘to bind’ > YAv *baṇḍaiieiti* ‘binds’; OI perfect *babandhá* (EWAia II s.v. *BANDH*; LIV<sup>2</sup> s.v. *\*b<sup>h</sup>end<sup>h</sup>-*)

(Aikio 2014: 138, footnote 7)

Aikio’s fresh Indo-European etymology of the Mari word is convincing both semantically and phonologically and it has been supported by Metsäranta (2018: 119). Aikio is right in arguing that Hu *fúz* ‘to tie’ cannot be regularly derived from the same Uralic reconstruction as the Mari word, contrary to the etymology of UEW (s.v. *\*pit<sub>3</sub>-*). Aikio only notes that the Mari verb is an Indo-European loan, but an Iranian source is the most probable here due to the limited distribution of the word. Pre-Mari *\*ä* can be easily explained as a substitute for Iranian *\*a*, but there are no convincing examples of *\*ä* corresponding to Proto-Indo-European *\*e*.

An Indo-Iranian etymology

**Ma *parča*, *pärcä* (< PMa *\*pärcá*) ‘ear of corn’**

← PII or PI *\*paršá-*, YAv *parša-* ‘ear of corn’, OI *paršá-* ‘sheaf, bundle’ (EWAia II s.v.)

(Aikio 2014: 141)

The etymology is convincing, and the vowel substitution *\*ä* is in line with many Indo-Iranian loans found in the centre and the east of the Uralic family. However, the substitution of the *\*-rš-* cluster raises suspicions – why was it substituted by Proto-Mari *\*-rč-* and not by *\*-rš-*? One solution would be to assume that *\*-rš-* would have been impossible at that stage in Proto-Mari/Pre-Mari. A further, yet hypothetical, possibility would be to assume that the Mari word is not a direct borrowing from Indo-Iranian, but was borrowed through some intermediary language, and this could explain the phonetically surprising substitution. According to EWAia, the background of the Indo-Iranian word is uncertain. In any case, the Indo-Iranian origin of the Mari word seems certain.

An Indo-Iranian etymology

SaN *báhtarit* ‘flee, escape’ (psA < *pāteṛe-*), SaS *baataridh* (< PsA\**pātāre-*) ~ *baateridh* SaSk *puât’tred* (irregular), (cognates in all Sa forms except T); Mansi North *pātər-*, East *pöätər-*, West *pṵər-* ‘disappear’, S *pātur-* ‘hide’ (< PMs \**pātər-*)

< PU \**pätäri-*

← PII (or PI) \**patara-*, > OI *pataráh*, *patáruḥ* ‘flying, shooting through the air’, root \**pat(H)* ‘to fly, fall’ < PIE \**pet(H)-* (EWAia II: 71 s.v.; Werba 1997: 301 No. 312 s.v. OI *pat<sup>i-</sup>*)

(Aikio 2013b: 162–63)

Both the Uralic background of the Saami and Mansi verbs and the Indo-Iranian etymology for the word was suggested only recently by Aikio. Also Kümmel (2018a) mentions that the Saami word is borrowed from this Indo-Iranian root. Aikio’s etymology is convincing, although the meanings of the Uralic and Indo-Iranian verbs are divergent. The semantic developments in the Uralic side are possible to explain from an earlier meaning of ‘flying’.

An interesting feature is the substitution of PII \**a* by Uralic \**ä*. This is hardly ever found in the earliest loans with a wide distribution, but occurs frequently in Iranian borrowings into Permic and the Ob-Ugrian languages, some of which have to be quite early. The word has to be a very early borrowing because of its distribution at the two opposite ends of the Uralic language family. Technically the word could have been borrowed separately by Pre-Mansi and Pre-Saami, but this does not remove the peculiarity of the vowel substitution, as the substitution by \**ä* is not found in later West Uralic loans that have cognates in Saami.

Aikio mentions that some of the Saami forms display vowel combinations that were not allowed in Proto-Saami, but because a regular reflex of Uralic \**pätäri-* and a phonotactically suitable form is found in South Saami *baataridh*, the irregular Saami forms can be considered secondary. There are no problems with the Mansi forms, which regularly continue Uralic \**pätäri-*. The Uralic verb stem is “anomalous” according to Aikio, but he explains this anomaly as a result of the word being a loan.

This is one of the few verbs borrowed from Indo-Iranian. In North Saami also forms of a noun \**báhtu* have been attested, as is mentioned by Aikio, but these are probably derived from the verb.

An Indo-Iranian etymology

**Mansi East *roasnə*, North *rāsn*, West *roāasn* ‘Strick’ < Proto-Mansi *\*rāsñā***  
(Zhivlov 2013: 219)

← **Iranian, cf. MP, NP *rasan* ‘Strick’, Chwarezmian *rysn*; OI *raśanā* - ‘Strick, Seil, Halfter’ < PII *\*raćanā* - (EWAia II s.v. *raśanā* -)**

(Paasonen 1909: 120–122; Munkácsi 1920: 165; Collinder FUV: 136; Korenchy 1972: 65, No. 20; Joki 1973: 308; Katz 1985: 340; 2003: 270; Lushnikova 1990: 248–249; Zhivlov 2013: 219)

This is an obvious Indo-Iranian etymology. The Mansi word is borrowed from the same Indo-Iranian word-family as Proto-Uralic *\*rećmā* discussed above (see that entry for more discussion about the background of the Indo-Iranian word).

Korenchy gives the donor form as “Middle Iranian” *\*rasan*/*\*rāsān*. Note that the OI form cited by Korenchy (“*rāśanā* -”) is wrong, as can be seen from the two accented vowels. Probably it is only a typographic error. Korenchy argues that the loan in Mansi is relatively late, and the word was acquired after the change PU *\*ś* > Mansi *s*. The same is also stated by Joki. Katz has the opposite opinion: he argues that the word was borrowed from an Indo-Iranian form with *\*ś* (= PII *\*ć*). EWAia mentions the latter explanation and refers to Katz. The same is stated also in the recent work of Zhivlov (2013). Zhivlov assumes that the Mansi word could be borrowed from the “Andronovo Aryan” language.

This is a crucial question regarding the dating of the etymology. The problem is that Iranian and Mansi have gone through rather similar sibilant developments, and it is very difficult to decide whether the word is an old (PII) or young (Iranian) loanword. The limited distribution in Mansi points to a late borrowing (this is stressed by Korenchy), but as has been repeatedly stated in this work, distribution alone is a poor criterion. The vowel substitution (Mansi *\*ā* from Iranian *\*a*) also offers little evidence, as this appears in Ob-Ugric loanwords of various age. Zhivlov also considers the Khanty word *\*wārəs* ‘horse’s mane’ an early parallel from Indo-Iranian, before the similar sibilant change happened in Khanty.

Although one cannot definitely decide whether the borrowing is old or young, one can argue that there are other loanwords limited to Mansi which have to be late due to phonological reasons: at least *sirāj* ‘sword’, *\*šārkvśv* ‘eagle’ and *\*wīsəγ* ‘calf’ are such words (see the respective entries). Although the distribution itself is not a definitive sign of late borrowing, together with these parallel cases it looks more likely that the word is a relatively late Iranian loan. Note that although the sibilant change does not have to be extremely early, and there are Iranian loans confined to Khanty that predate this change (see PKh *\*aγər*), the Indo-Iranian change from *\*ć* to Iranian *s* (via PI *\*ts*) must have happened quite early, as is shown by Avestan.

Zhivlov’s argument that this word could be borrowed from Andronovo Aryan which would have retained *\*ć/ś* (instead of Iranian *s*) is not very strong, as nothing prevents us from assuming an Iranian borrowing after the change *\*ś* > *s*. The Mansi vowels do not tell us much about the age of the word, as Mansi *\*ā* vocalism can reflect PU *\*ā* (Zhivlov 2006), and it is difficult to determine whether the substitution in this word was Proto-Mansi *\*ā* or some preceding phase, such as short *\*ä*. Lushnikova supports the older view that the word was borrowed after *\*ś* became *\*s* in Mansi.

Interestingly, the Mansi word is probably not borrowed from an Alanic-type language, because modern Ossetic shows irregular *rætæn*.

An Indo-Iranian etymology

**Khanty South** *rõw* 'Unreinigkeit, Abfall, Moder (im Wasser)', East *ráuwyз*, 'Scwhamm, Moder (auf dem Sumpf, von den Wellen angetrieben'

< PKh ? \**rawV*

← PI \**raha-* (< PII \**rasa-*) > YAv *raḥhā-* 'name of a mythological river'; OI *rása-* 'Pflanzensaft, Saft, Flüssigkeit jeglicher Art, Seim, Essenz, Kern, Mark'

(Munkácsi 1901: 533; FUV: 135; Korenchy 1972: 65–66 No. 21; Katz 1985: 122; 2003: 104; Blažek 2005: 172–174)

The Khanty and the Mordvin word *rav*, *ravo* 'Volga' which has often been connected are not regular cognates but they can possibly be parallel borrowings from Iranian (as observed already by Korenchy). The Mordvin word is discussed as a separate entry. The Khanty vowels do not point to a Pre-Khanty \**a-a*-stem, thus weakening the supposed Iranian origin of the word. Semantically the etymology is possible but not obvious. The medial consonantism of the Khanty word cannot be explained from Iranian \**s* or \**h*. A pure chance resemblance is also possible. There is no mention of the Iranian etymology in DEWOS (1268–1269).

The UEW gives a reconstruction \**raḥV*, but it is not at all certain that \**ḥ* should be reconstructed on the basis of either the Mordvin or the Khanty word (*ḥ* is usually retained in part of the Erzya Mordvin dialects).

Korenchy mentions that also the Indo-Iranian verb \**sraw-* 'to flow' (PIE \**srew* > PII *sraw* > OI *sru* '(aus)strömen/fliessen, lecken, (zer)rinnen', Av *sru* [Werba 1997: 266–267 No. 244]) has been considered as the origin of the Khanty word, but also this option is not semantically very convincing.

To sum up, the Indo-Iranian etymology of the Mordvin word is probable, but it is unlikely that the Khanty word has anything to do with either of the presented Indo-Iranian etymologies.

Not an Indo-Iranian etymology

**Md** *rav*, *ravo* 'Volga'

See \*PKh ? \**rawV*

An Indo-Iranian etymology

Mansi East, North, West *rĕk<sup>w</sup>*, West *rek<sup>w</sup> ~ rĕk<sup>w</sup>* ‘steiles Ufer’; Khanty North (Nizjam) *rōwa*, (Kazym) *rōwa*, *rĕwa*, (Obdorsk) *rōwa*, *rĕwa* ‘Uferabhang’

< POUg ?

← PI *\*raga-* > Av *raya-*, OP *ragā* ‘the name of the country in Media (in high plateau)’, Sogdian *r’g* ‘plain’ etc. (AiWb s.v. *raya-*)

(Blažek 1990: 40)

The Khanty and Mansi words are probably not inherited, regular cognates, as the vowel correspondences are obscure, and the word is missing from the material in Zhivlov (2006). Honti (1982: 205) mentions the Khanty and Mansi words in the list of words that do not reflect a regular Proto-Ob-Ugric form. Semantically the etymology is not obvious, and Blažek does not give any detailed arguments in favour of the semantic developments.

AiWb (s.v.) gives no information on the etymological background of these Iranian words. In the light of this, it is not likely that the Ob-Ugric words could have been borrowed from this Iranian form.

Not an Iranian etymology

Fi *repo*, Ka *repoi*, Est *rebane*, *rebu* (< PFI *\*repä-j*) ‘fox’; ? SaN *rieban*, Lu *riepij*; Mo E *riveš* id.; Mari *rəwəž*, *rəwəž* id. < PMA *\*rŭwəž*; ? Ud *ž’iči*, *ž’oči*, Ko *ruć* id. (< PP *\*rući*)

< ? PU *\*repä* (UEW 423–424; Sammallahti 1988: 553)

← Pre-II *\*h<sub>2</sub>rewpēco-* > PII *\*Hrawpācā-* > OI (Vedic) *lōpāsā-* ‘a species of jackal’, OP (attested in Elamite as a personal name) *\*raupāθa-*, MP, NP *rōbāh*, Oss *rūvas/robas* ‘fox’, Khot *rŭwāsa-* ‘jackal’; Av *raopi-* ‘fox’, Khowār *lōw* < PII *\*Hrawp-i-* < ? PIE *\*h<sub>2</sub>lewp-* (EWAia II .v. *lōpāsā-*; Clackson 1994: 96; de Vaan 2000: 289–291; 2009 s.v. *volpēs*; Cheung 2002: 221)

(Munkácsi 1901: 521–522; KEWA III s.v. *lōpāsāh*; Joki 1973: 308–309; Katz 1985: 136–137; 2003: 115–116; UEW No. 859, s.v. *\*repä* [~ *\*repä-ćä*]; Rédei 1986c: 46; Lushnikova 1990: 251–253; SSA III s.v. *repo*; WOT: 1287, 1321; Bereczki 2013 s.v.; Honti 2017: 154–156; Róna-Tas 2017: 69)

This is an old and widely-accepted etymology, but there are several problems in determining the exact donor language and form, the chronology of Indo-Iranian sound changes, the irregular relationship of the several Uralic forms and the reconstruction of the Uralic word. The Indo-Iranian etymology of the Uralic word is mentioned by EWAia, but it has not been discussed in the recent works of Koivulehto or Sammallahti that deal with loanwords.

Semantically the etymology is convincing. In the Indo-Iranian languages the word denotes both ‘fox’ and ‘jackal’, but this is hardly an obstacle, as the meaning ‘fox’ is attested in both language families, and animal names frequently change their meaning due to

various reasons (such as taboo). The fox and jackal are also closely-related canine animals. One might wonder why a word for an animal like the fox which is frequent in the Northern Taiga zone would be borrowed, but this can have something to do with taboo, or perhaps the fox had important cultural meaning (because of the fur trade, for example). Also other animal names have been borrowed from Indo-Iranian, such as Mo *veṛgis*, *vərgas* ‘wolf’ (see the entry).

Because of the vowel *\*e*, the word has to belong to the oldest layer of Indo-Iranian (Pre-Indo-Iranian) borrowings. The simple *\*e* in the first syllable instead of a sequence *\*ew* probably results from phonotactics, as there are no examples of the sequence *\*-wp-* in the Proto-Uralic vocabulary. Otherwise we would expect *\*ew* to have been retained, as this sequence as such was allowed in Proto-Uralic (cf. *\*lewli* ‘spirit’). The archaic vocalism fits well with the relatively wide distribution of the word in Uralic. Concerning the vocalism of the Indo-Iranian word, the word does not have a clear Indo-European background, although various suggestions have been made. KEWA considers the etymological connection with Greek ἄλωπηξ and Armenian *alowēs* ‘fox’ to be likely, and it mentions Latin *volpēs* ‘fox’ and Lithuanian *lāpe* as forms that are also possibly related. The connection to Greek and Armenian is also supported by EWAia. Derksen (2015 s.v.) connects the Lithuanian word and its Latvian cognate *lapsa* to the Greek and Armenian words, as well as to Welsh *llywarn*, reconstructing PIE *\*h<sub>2</sub>lōp-s*. The Indo-Iranian word is not mentioned by Derksen. De Vaan (2008) considers the Greek and Armenian words and Baltic *lāpe* as unrelated to the Indo-Iranian forms, but he derives Latin *volpēs* and Lithuanian *vilpišỹš* ‘wild cat’ from the same IE root as Indo-Iranian *\*Hrawpi-*. De Vaan argues that the Lithuanian word and Middle Persian *gurbag* ‘cat’ continue the original zero-grade form *\*h<sub>2</sub>wlpi*, and that Avestan *urupi-* ‘marten’ reflects a metathesized form *\*Hurpi-*, from which the full-grade *\*Hrawpi-* has been derived.

On the other hand, the vocalism *\*ew* (> *\*aw*) in Indo-Iranian might have been influenced by the unrelated root *lup-* ‘steal, plunder’ through folk etymology, as the Greek and Armenian words point to *\*e/o* and not to a diphthong (Clackson 1994: 96; de Vaan 2000: 289–291). A different view is expressed in EWAia, where it is assumed that the Indo-Iranian form is more archaic and has preserved the original diphthong, and the Greek and Armenian *o* is due to taboo. From the point of view of the Uralic word it does not matter which vocalism is more archaic, but if the *\*ew* instead of *\*o* caused by analogy, the development has to be very early, already taking place in Pre-Indo-Iranian. It is also clear that if the Greek and Armenian words are unrelated, as de Vaan (2008) assumes, this explanation that is based on analogy is not needed.

To sum up, the Indo-European background of the Indo-Iranian word involves various problems, but it is clear that for a pre-form of Vedic *lōpāsá-* a form *\*(h<sub>2</sub>)rewpēco-* can be reconstructed, and the Uralic word with the vocalism *e* can be easily derived from this form. Because Indo-Iranian *r* here almost certainly reflects PIE *\*l*, no other branch of Indo-European can come into question as the donor language.

This is one of the examples of a loan with retained Indo-European *\*e* but Indo-Iranian *\*r* from earlier *\*l*; another example is Sa *geavri* < *\*kekrä*. The Uralic loans thus show that the change *\*l* > *\*r* was very early in Indo-Iranian, as it must have happened before the merger of non-high vowels. It would be impossible to derive the Uralic *\*e* from later Indo-Iranian *\*a*. It has been argued by Mayrhofer (2004; cf. also Kümmel 2018a) that all of the Indo-Aryan words with *l* can be explained as secondary dialectal innovations (Mayrhofer

[2004: 158] lists several cases where Sanskrit shows *l* in place of PIE *\*r*, meaning that the Old Indo-Aryan dialects where *l* is found have not retained the PIE distinction of *\*l* and *\*r* but have innovated and show secondary *l*. The change *\*l* > *\*r* was regular, and later the change from *\*r* to *\*l* occurred in some Indo-Aryan dialects, from which forms with *\*l* have diffused into the Vedic and Classical Sanskrit texts. Also the Tocharian loans from Iranian consistently show *\*r* in place of Indo-Iranian *\*l* (Tremblay 2005a: 424).

Even if the results of this Indo-Iranian sound change would not have been *\*r* but some other resonant sound, the mere fact that the opposition between *\*l* and *\*r* disappeared in Indo-Iranian would be enough to explain the Uralic words with *\*r*. There are no examples of early Indo-Iranian words with retained *\*l* (which would otherwise show clear Indo-Iranian characteristics).

Interestingly, word-initial *\*r*- was probably alien to Uralic phonotactics, as no PU words with *\*r*- can be reconstructed (in the wordlist of Sammallahti 1988 the words beginning with *\*r*- first appear in the Finno-Permic vocabulary), and in Samoyedic the situation has persisted until modern times. Initial *\*r*- was probably introduced through loanwords like this.

Although most of the Uralic words listed above can be derived from a reconstructed stem *\*repä*, the Saami words are probably borrowed from Finnic, and the Hungarian words represent later parallel borrowings from Iranian (see below). Bergsland (1963: 151–153) argues that the word has been borrowed from Finnic into Saami in several different forms: the forms SaN (dial.) *riebân*, SaI *riemñis*, Skolt *rì-èwəŋ*, Kola *rimñe* point to a donor source *\*repäinen*, which is not attested in Finnish or Karelian but is reflected in many Finnic languages such as Veps *rebāne*, Est *rebane*. South Saami *riepie* (analogically from *\*ræbeš*) points to a donor form *\*repäs* (*\*rewäs*).

The Hungarian words *róka* ‘fox’ and *ravasz* ‘cunning; (obs., Old Hungarian) fox’ are often connected to the other Uralic ‘fox’ words. While the two Hungarian words are clearly somehow connected to each other and bear a resemblance to the western Uralic forms, they cannot regularly reflect the same inherited proto-form *\*repä* as the western Uralic languages. Despite the clear irregular relationship of *róka* and *ravasz* both to each other and to the other Uralic forms, Hungarian etymologists have been very stubborn in their attempts to connect these forms etymologically and to derive them from a common source.

Rédei (1986c and UEW) has attempted to explain the back-vowel forms *ravasz* and *róka* through vowel assimilation caused by the following consonant: *\*repä* > *\*rewä* > *\*row* > *\*ró* (+ suffix *-ka*) but his explanation is ad hoc and without clear parallels; Rédei argues that at the time when *w* labialized the vowel, there were not yet palatal labial vowels. The form *ravasz* had *u* in old Hungarian according to the UEW and Róna-Tas, and the *a* can in Rédei’s view be explained in a similar way as the labial vowel in *róka*. However, the form *adruuoz* cited by the UEW is from the earliest possible Hungarian written source, the Tihányi Alapítólevél, and one cannot be sure that we are really dealing with the word ‘fox’ here, and that the grapheme <*u*> really corresponds to a vowel *u*. The development *u* > *a* in the era between Old and Modern Hungarian would also be irregular.<sup>37</sup>

<sup>37</sup> Although traditionally some Hungarian words with *a* or *á*, like *fá* ‘tree’ and *száj* ‘mouth’ have been derived from Uralic forms with *\*u*, almost all of these words involve difficulties in the reconstruction of the vocalism, and it is highly doubtful whether *\*u* > *\*a* is a real sound change in the history of prehistory of Hungarian. In any case this sound change should have happened before the Old Hungarian period, for during or after the Old Hungarian period there were vowel changes that led to the raising of low and mid

Already H. Sköld (1925: 32–33) considered the Hungarian word *ravasz* to be a loan from Alanic (“Ossetic” in H. Sköld’s terminology), and this seems to be the most plausible explanation for the Hungarian word, as the Ossetic forms *rūbás*, *rūwas*, Digor *robas* suit Hungarian more regularly than the phonologically difficult PU form. It is unclear why the idea of a separate, later borrowing has been so difficult to accept for many researchers. H. Sköld states that the word is a ‘Wanderwort’, and that the loans in other Uralic languages can easily be explained as parallel borrowings from Iranian. Also Joki (1973: 309) hints at this possibility, stating enigmatically that “besonders im Ung. ist eine deutliche miran. Einwirkung zu erkennen”. This is a very modern view, and it is a pity that it has not been supported by subsequent researchers of Hungarian etymology. Sammallahti (1988: 553) reconstructs the word for a “Finno-Permic” proto-language, and he does not consider the Hungarian word a cognate of the other forms. In SSA the Hungarian word is added with a question mark.

Recently Róna-Tas (WOT: 1321, 1335; 2017: 69) has suggested that *ravasz* and *róka* are historically different words indeed. For *róka* he assumes the development from *\*repä* as described in the UEW, but *ravasz* is in his view a borrowing from Alanic *\*rowasa*. While the explanation for *ravasz* is certainly correct (the vowels and consonantism of the Hungarian trisyllabic word can be easily explained from the Alanic word), it remains mysterious why Róna-Tas supports such a complicated and irregular history for *róka*. A much better solution would be to consider both *róka* and *ravasz* separate borrowings from different Indo-Iranian forms. In the Iranian languages there exist both *raupasā-* (from which the Alanic form mentioned by WOT) and *raupaka-* (whence Dameli *rōpak*), and the latter form could be the origin of the Hungarian word *róka*. Both forms would thus be relatively late Iranian borrowings. The scarce attestation of the *raupaka-* word is admittedly slightly problematic, but it can be assumed also that Hungarian *ró-* is an early loan from the underived stem *\*raupa-* (from which the Hungarian word could be regularly explained), and the suffix *-ka* is simply a Hungarian (diminutive?) suffix added later.

Honti (2017: 154–156) has criticized WOT’s explanation of separating *róka* and *ravasz* etymologically, but he offers no arguments for this – Honti’s idea is mainly that the traditional explanation by the UEW etc. is better and there is no reason to reject it in spite of the obvious irregularities. Honti’s view is untenable. Honti also states that Sammallahti has “probably forgotten” to mention the Hungarian forms in his etymological wordlist, although it seems obvious that Sammallahti has rejected the Hungarian words from this comparison because of their irregular vocalism.

Regarding the remaining Uralic forms, the relationship between Finnic *repo*, *rebane*, Erzya *riveš*, and Mari *rəwəž* is regular, at least concerning the initial-syllable vocalism, and all the words could be derived from Proto-Uralic *\*repä*. The Mari dialectal form *əwəž* can be explained by Chuvash influence, as Chuvash (like Proto-Uralic) does not allow initial *r-*. However, it is unclear whether the sibilant/affricate found in a number of words is a reflex of a Proto-Uralic *\*-š* or some later suffix. According to Bereczki (2013; 2002: 192–193), the Indo-Iranian *é* has been analogically connected with the Uralic suffix *é* in Pre-Mari. It is further unclear whether this sound reflects the PII *\*-é* found in many of the Indo-Iranian

---

vowels (OHu *hodu* ‘army’ > Hu *had*, OHu *Ystor* [ʔ *uštōr*] > Hu *ostor*), but there are no convincing examples of a change *u* > *a*.

words. The UEW does not take a clear stance on this, reconstructing two alternative proto-forms, and Sammallahti (1988) reconstructs simply \**repä*. Joki does not comment on the issue.

Harmatta (1977: 171) has already considered the words in Permic and Hungarian as parallel loans, rather than reflexes of PU \**repä*. Harmatta assumes that only the Finnic and Saami forms reflect underived \**repä*, the Mordvin and Mari words are reflexes of \**repeše*, Hungarian *róka* reflects a different borrowing \**ropa*, and the Permic words and Hungarian *ravasz* are reflexes of \**ropaća*. Probably the Permic word is a separate borrowing, and so is Hu *ravasz*.

These kinds of explanations are problematic because the derivational morphology of Proto-Uralic is poorly known. Too often in etymological literature various obscure suffixal elements are simply explained as derivational suffixes, even if nothing sure can be said about them. This is an unfortunate situation. In a case like this, where we have different Indo-Iranian forms, such as underived \**rawpa-* (< \**rewpo-*) and different derived forms like \**raupaća*, it seems obvious that a more secure explanation is to connect the obscure \**ć* of the Uralic side to the Indo-Iranian suffixal element. This means that the sibilant in Mordvin and Mari forms and the Permic affricate can all be explained from PU \**ć*, which is a substitute of PII \**ć*.

The Finnic form \**repä*, on the other hand, could be derived from the simple PII \**rewp-* without the “animal-suffix” \**-āca-* (for which see de Vaan 2000), and it is possible that we have two parallel borrowings from Pre-Indo-Iranian into various dialects of Proto-Uralic. However, here it is slightly troubling that a separate Pre-II form is found in Finnic only, pointing to a situation of very early dialectal divergence at the time of the earliest contacts with Indo-Iranian. Another possibility (suggested by Bergsland 1963) would be to derive Finnic *repoi* from earlier \**repäinen* as a back-formation. As the Saami and Scandinavian loans also support the reconstruction of variants PFI \**repäinen* and \**repäs*, then perhaps this is the more original form and *repoi* indeed is a back-formation.

Regarding the Permic words (Udmurt *žiči*, *žóči*, Komi *ruć*), the standard explanation is that they too are part of the same cognate set, and can be derived from PU \**repäci*. However, Niklas Metsäranta (personal communication) has doubted this, as the Permic vowels are not regular reflexes of PU \**e*. This could point to a parallel borrowing from Indo-Iranian, probably from Proto-Indo-Iranian \**Hrawpāća*. The parallel borrowing into Permic cannot be much later, as Permic *ć* probably reflects the same suffix as the \**ć* in other Uralic forms discussed above, and this means that later Iranian forms are out of question.

As a side note, the Scandinavian forms ON *refr* ‘fox’, modern Swedish *räv*. Norwegian *rev* id. have been connected etymologically to the Finnic/Uralic words already by Thomsen (1869: 40) and it is usually assumed that the Scandinavian form is borrowed from Finnic. Bergsland (1963: 153–54) argues in favour of this in more detail, noting that the ON form *refr* could be convincingly derived from Finnic \**repäs* (paradigm with gradation: \**rewäs* : \**repähän*). Although such a form is not attested in Finnic, Bergsland notes that some of the Saami words require a similar donor form, and that this \**repäs* forms corresponds well to the forms in Mordvin and other eastern languages. Keresztes (1997: 146; 1998: 95) mentions the possibility that the Scandinavian word is borrowed from Saami rather than from Finnic, but he does not offer phonological arguments in support of

this. Phonologically, either the Finnic or Saami words could have produced the Scandinavian forms.

Kroonen (2013 s.v. *\*fuhsa-*), on the other hand, states that the Scandinavian forms are borrowed from Iranian. There are, however, no arguments to support the Iranian etymology over the Finnic one, and geographical reasons make the Finnic etymology look much more plausible than a borrowing from geographically distant Iranian. Although there are other Germanic words with tentative Iranian etymologies, such as PG *\*paþa-*, none of them is limited to North Germanic only. Also the context of borrowing the name of a wild animal which is important in the fur trade, an important livelihood in ancient Fennoscandia, would be perfectly plausible via the ancient trade networks.

Indo-Iranian etymology (Hungarian words represent parallel borrowings)

**Ma *rüm*, *rümbâk*, *rämälye* ‘(Abend)dämmerung’; Ko *reṃ* ‘colour’, Ud *žomīt* ‘Dämmerung’ (< PP *\*rōṅV* ‘dusk’)**

< PU ? *\*rVmV* ‘colour’

← ? Middle Iranian *\*rang* ‘colour’, MP *rang* ‘colour, dye’ < PII *\*ranga-*

(Munkácsi 1901: 535; KESK: 243–244; Joki 1973: 309; UEW s.v. *rōm3*; Rédei 1986c: 66–67; Csúcs 2005: 320; Bereczi 2013 s.v. *rümbâk*)

According to the UEW, the Mari and Permic words are cognates and represent a Middle Iranian borrowing into Proto-Finno-Permic. Bereczi (2013) follows the same view and mentions the Iranian origin (but not the donor form) without any further comment.

Rédei (1986c) represents a different view from the UEW, arguing that the Permic word is borrowed from Middle Iranian, and the Mari word is then borrowed from Permic after the sound change *\*ŋ > m* took place in Permic. Csúcs (2005) follows Rédei’s view and does not mention the possibility of Finno-Permic inheritance at all.

All of the explanations are problematic. It is difficult to see how Permic *m* could correspond to Iranian *\*ng*. Also Mari *m* cannot be derived from earlier *\*ŋ*, which is retained regularly in Mari (compare PU *\*wāŋi* > Mari *wiŋə* ‘son-in-law’). Semantically the etymology would be possible if ‘dusk’ is the primary meaning on the Permic side, but because of the phonological problems, the etymology should be rejected.

However, the original etymology by Munkácsi (1901) derived the Permic word from an entirely different Indo-Iranian word, namely from PII *\*rāmá-* ‘dark, black’ (> OI *rāmá-*). This etymology is phonologically clearly better, but similar semantic issues remain as with the other etymologies discussed above.

? Indo-Iranian etymology

SaN *reašmi* ‘rope at the end of the fishing net’ (< PSa *\*reaćmē*; distribution Lu–I); Mo E *rišmé* ‘Kette’, M *rišmä* ‘Seil’

< ? P(W)U *\*rećmä*

← Pre-II *\*Hrećmi*-> PII *\*Hraćmi*-, root *\*Hrać*- > OI *rašmi*- ‘Schnur, Strick, Seil, Zügel, Peitsche, Strahl’, *rašmán*- ‘Zügel’, root *raš* ‘binden’ (EWAia s.v. .v. *rašmi*-; Werba 1997: 227 No. 160 s.v. *raš*; Lubostky 1988: 39, 112)

(Munkácsi 1901: 180; FUV: 136; Joki 1973: 308; KEWA III: 46 s.v. *rašmih*; Katz 1985: 340; 2003: 270; UEW: 744; Lushnikova 1990: 259; Sammallahti 1998: 126; Koivulehto 2001b: 250)

See also *\*rāšnā*

This is a well-known old etymology (already discussed by Munkácsi) that is in principle convincing, but various views about the exact donor language and the relationship of the Saami and Mordvin words have been put forth. The word has to be an early borrowing, as the vowels clearly reflect earlier *\*e*. However, it is unclear whether this is actually an Indo-Iranian borrowing, as Koivulehto has argued that the word could be borrowed from Proto-Indo-European (or Northwest Indo-European), where *\*ć* could substitute for *\*k*<sup>1</sup>. The problem is that this substitution is impossible to prove, and a more economical solution is to opt for a Pre-Indo-Iranian loan, as this is a well-founded loanword layer in Uralic. Furthermore, the Indo-Iranian word does not have an undisputed Indo-Iranian etymology, making the idea of a PIE borrowing less likely. Sammallahti (1998: 126) lists the word among the Pre-Indo-Iranian loanwords that have reflexes in Saami, but he does not comment on the etymology in more detail, and the word is interestingly absent from Sammallahti’s 2001 listing of early Indo-European loans in Saami. Also Dolgopolsky (1989: 21–22) mentions the Uralic etymology but does not comment on the possible IE origin of the Indo-Iranian word.

Concerning the Indo-European origin of this Indo-Iranian word, the word does not have a secure Indo-European etymology. Some sources (Schneider 2010: 212–215; Werba s.v.) reconstruct a possible Proto-Indo-European root pre-form *\*lak*<sup>1</sup>-. However, no such root is found in LIV<sup>2</sup> or NIL. According to Schneider (2010: 213; footnote 249, 215) the noun *rašmán*- could reflect earlier PIE *\*lak<sup>1</sup>-mon*. The Uralic evidence does not fit well with the idea that the Indo-Iranian word reflects PIE *\*a* vocalism, and here the Uralic word can be used as evidence for the old *\*e*-grade of the Indo-Iranian word. The decision to reconstruct *\*a* seems not to be based on any Indo-Europeanist arguments. Furthermore, if the word indeed is inherited from Proto-Indo-European, *\*e* would be more expected than the very rare *\*a*.

An older view reflected in (IEW: 863) connects the Indic word with Germanic words such as ON *rakki* m. ‘parrel’ on the basis of the Indic word and Germanic words such as ON *rakki* m. ‘parrel’. IEW (863) reconstructs an IE root *\*rek*<sup>1</sup>-/*\*reg*<sup>1</sup>- as the PIE form. It is unclear why the alternation of a voiceless and voiced velar is assumed. The *k* in the Germanic words is probably due to Verner’s law. Kroonen (2013: s.v. *\*rakkan*-) accepts the comparison and reconstructs PIE *\*Hrok<sup>1</sup>-on-* as the preform of the Germanic word; OI

*rašmi-* reflects PIE *\*Hre/ok-mi-* and OI *rašānā* reflects PIE *\*Hrek-en-eh<sub>2</sub>-* according to Kroonen.

Lubotsky (2001b: 314) mentions that the formation *\*rać-m-* is confined to Indo-Iranian. The forms without *-m-* (OI *rašānā-*, MP *rasan* ‘rope’) are explained by Lubotsky as possible results of dissimilation (< *\*raćmna-*?). This would mean that the Uralic word could only be the result of borrowing from Indo-Iranian, albeit at very early date.

Grünthal (2012: 328–329) has argued that the Mordvin word E *rišmé*, M *rišmä* is rather a borrowing from a Baltic source akin to Lithuanian *rišimas* ‘binding’ (derived from the verb *rišti* ‘to bind’; see ALEW: 868–869 s.v. *rišti*, Derksen 2015: 383 s.v. *rišti* and LIV<sup>2</sup>: 669 s.v. *\*wreyk-* for competing views on the Indo-European etymology of the Baltic word family). The word obviously cannot be a proper Baltic borrowing because of the palatal sibilant, so it could rather be an early borrowing from Proto-Balto-Slavic; Grünthal assumed that *\*ś* is an archaism (in Proto-Baltic *\*ś* would have become *\*š*) so he was on the same track even if he somewhat confusingly mentions the Proto-Baltic form *\*rišima-* as the donor form.

Junttila (2018: 83–84) has recently discussed the Mordvin word and modified Grünthal’s Baltic etymology, assuming that the word can belong to the layer of ‘old Latgalian’ loans in Mordvin. In Junttila’s view the Proto-Mordvin palatal sibilant *\*ś* could go back to earlier *\*s*, because according to Junttila there are no non-palatalized *\*-sm-* clusters in Mordvin. This *\*s* could then be a substitution for Baltic (Latvian/Latgalian-type) *s*, from Proto-Baltic *\*š*.

From the Uralic point of view, there is no compelling reason to assume that the Mordvin word is a parallel borrowing, as it could be derived from exactly the same proto-form as the Saami form, as has been done in earlier sources (such as the UEW): for the Mordvin vocalism, cf. *\*mertä* > Mo E *mird’e*, M *mird’ä* (see the entry *\*mertä*). Finnic *rihma*, on the other hand, is a well-known borrowing from Baltic *\*rišima*, and Fi *räismä* ‘Seil vom Baumwurzeln’ is borrowed from the Saami word, as noted by the UEW.

Semantically either a Baltic or an Indo-Iranian source is equally convincing. If the Saami and Mordvin word are indeed cognates, it is impossible to derive their common proto-form from Baltic, as the Saami word cannot reflect an earlier *\*i-ä* stem, whereas the Mordvin forms could reflect either *\*rićmä* or *\*rećmä*.

Koivulehto notes that the word is probably not borrowed from a form leading to OI *rašmi-* because of the stem vowel *\*-ä* in Uralic. This is not a very strong argument, however, as the substitution of stem vowels is poorly known, and *\*a/ä-* stems are frequent in Indo-Iranian loanwords, irrespective of the stem vowel of the original. Furthermore, Koivulehto does not mention the Vedic *n-*stem *rašmán-* (< PII *\*raćman-*) In any case, both *rašmán-* and *rašmi-* appear in Indo-Aryan, and both are attested already in Vedic.

Gamkrelidze & Ivanov (1984: 921), referring to Abaev 1958–1989 (II: 382–383), argue that the Uralic *\*e* vocalism could reflect the later fronted *æ* attested in Ossetic, and there is no need to argue that the word is very early borrowing showing the retained IE *\*e*. This is an improbable idea, as Iranian *æ* would have been more likely substituted by Uralic *\*ä*, as the vowel *ä* is found in many borrowings as the substitution of Iranian *\*a* or its later reflexes. Gamkrelidze & Ivanov’s idea clearly stems from their reluctance to accept borrowings that are earlier than Proto-Iranian.

An Indo-Iranian etymology

## Mari *rweze* ‘boy’

See the entry *\*orpV* for discussion and references.

? Indo-Iranian etymology

**Fi *saada*, *saa-* ‘to get, obtain’; SaKld *soagke-* ‘to catch; to get’ (< PSa *\*sāke-*; reflexes only in Kld and T); Md E *sa-* ‘to come’, *saj-* ‘to take’, M *sa-* ‘to come’, *šav-* ‘to take’; Ko *su-*, Ud *suti-* (der.) < PP *\*su-* (UEW s.v. *\*saye-*; Sammallahti 1988: 553; SSA III s.v. *saada*; Zhivlov 2014: 131–132; Aikio 2015b: 53)**

< PU *\*sayi-*

← PII *\*saH-* > OI *san<sup>i</sup>/sā* ‘to win, reach, get, obtain’, Av *han* < PII *\*sanH* < PIE *\*senh<sub>2</sub>* (EWAia II s.v. *SĀ*; Werba 1997: 322–333 No. 360 s.v. *san<sup>i</sup>/sā*; LIV<sup>2</sup> s.v. *\*senh<sub>2</sub>*)

(Parpola 2010: 313)

or

← PII/PIA *\*saǰh-* > OI *sah-* ‘to vanquish, overcome, conquer, be victorious, have power, gain (battles etc.)’, Av *\*haz* attested only in the noun *hazah-* ‘Gewalt, Macht; Überlegenheit’ (= Vedic *sáhas-*); < PIE *\*seǵh-* ‘überwältigen, in den Griff bekommen’ (EWAia s.v. *SAH*; Werba 1997: 252–253 No. 214 s.v. *sah*; Stüber 2002: 145–147; LIV<sup>2</sup> s.v. *\*seǵh-*)

(Koivulehto 1999a: 227 = 2016: 229)

The Indo-Iranian etymology for the Uralic word was first suggested by Koivulehto. Koivulehto’s etymology has not been commented on since, except by Parpola, who proposes a competing Indo-Iranian etymology. Of these two etymologies, Parpola’s suggestion is more convincing.

In Parpola’s etymology, the Uralic velar spirant *\*ɣ* substitutes for the PII laryngeal. While there are few examples of laryngeals being substituted in the Indo-Iranian loans in Uralic, it would actually be strange not to find reflexes of them, since it is universally accepted that at least one laryngeal (reflecting PIE *\*h<sub>1</sub>*, *\*h<sub>2</sub>* and *\*h<sub>3</sub>*) should be reconstructed for PII, and it is even possible that there were two different laryngeal phonemes in Proto-Indo-Iranian (Kümmel 2018b: 169–170). However, here one would gladly see more parallel examples. It is also possible that although a reflex of the laryngeal was retained, it had become so “weak” that it was not substituted in Uralic: this kind of idea has been supported by Kallio (2008: 267) with regard to the Balto-Slavic loans of Finnic, as these loans likewise show no reflex of the Balto-Slavic “laryngeal”. Here the Indo-Iranian laryngeal would go back to the laryngeal *\*h<sub>2</sub>*, which according to Kümmel (2018b) remained distinct from PIE *\*h<sub>3</sub>* in Proto-Indo-Iranian. There is thus no obstacle to assume that here the Indo-European laryngeal is reflected in Uralic, other than the general absence of laryngeal reflexes in loanwords.

Also semantically Parpola's etymology is more convincing, as this Indo-Iranian root denotes 'obtaining', whereas Koivulehto's suggestion 'to win, conquer' is semantically more far-fetched. There is thus reason to reject Koivulehto's etymology and to support Parpola's new suggestion.

Koivulehto's Proto-Indo-Aryan etymology would mean that the Indo-Aryan deocclusion would have taken place already at the time of borrowing, and the Pre-Finnic \* $\gamma$  would substitute for the IA \* $h$  (see also the entry \**meyi*). The Indo-Aryan evidence from Mitanni clearly shows that this cannot be the case, but \* $\zeta^h$  was still retained in Proto-Indo-Aryan and even later.

The distribution of the word within Uralic also requires comment, as different cognate sets can be found in different sources. The UEW and SSA connect possible cognates from Saami, Mordvin, Permic and Samoyedic languages to the Finnic word *saada*. The potential Samoyed cognate PS \**tâjwâ-* is rejected by Aikio (2002: 29–30), as mentioned by Parpola.

Hyllested (2009: 11) has considered the Uralic word and Indo-European \**segh-* (> OI *sah-*) as representing an inheritance from a hypothetical Indo-Uralic proto-language. Hyllested lists a number of examples which show the correspondence Uralic \* $j$  ~ Indo-European \**gh* and assumes that the predecessor of these sounds was Indo-Uralic \**gj* (a palatal stop?), which would have developed into \* $j$  in Uralic in word-initial position, and to \* $\gamma$  in medial position, while in Indo-European this sound developed into \**gh*. A change from \**gj* to aspirated \**gh* would be phonetically unlikely and typologically rare, and Hyllested's idea of an inherited Indo-Uralic word is much less convincing than Parpola's etymology.

An Indo-Iranian etymology (← \**saH-*)

**Fi *sammua* 'to be extinguished' (also in Ka, Lu, Ve, Vo) < PFi \**sammu-* (SSA s.v. *sammua*)**

← \**ćamH-* > OI \**sami* - 'to be extinguished, die out', OI \**śamyati* < PIE \**kémh<sub>2</sub>* (EWAia II: 610–611 *sami*; Werba 1997: 319–320 No. 353 s.v. *sami*)

(Parpola 2010: 313)

This is in principle a convincing etymology. However, there is a competing Germanic etymology recently suggested by Koivulehto<sup>38</sup> (2016: 458–59): PG \**stammian-* > ON *stemma* 'to stop, staunch'. Koivulehto notes that the Indo-Iranian etymology is also flawless (both semantically and phonologically), but the limited distribution is the only thing (a statistic argument) that can be used as a support of the Germanic etymology over the Indo-Iranian one.

If the word is borrowed from Indo-Iranian, it can be either an early Proto-Indo-Iranian loan or a later borrowing from some Iranian language with \* $s$  (< \* $ć$ ), because the Finnic sibilant can reflect either \* $ś$  or  $s$ . This is not commented on by Parpola. The Finnic geminate *-mm-* is suspicious, and can be better explained from the Germanic original.

<sup>38</sup> Koivulehto's etymology is published in his memorial volume (Koivulehto 2016), in a chapter that has been written by Petri Kallio.

The more widespread Uralic stem *\*ćoma-* is borrowed from the same source according to Aikio (2014: 88–89) (see the entry).

? Indo-Iranian etymology

**Fi *samm*s ‘pillar’, *sampo* ‘the mill of abundance’, Vo *samm*az, Est *samm*s, arch. *samb*s < PFi *\*samp*s ‘pillar’ (SSA s.v. *samm*s; EES s.v. *samm*s)**

< Pre-Fi *\*samp*s

← PII *\*stamb<sup>h</sup>Ha-s*

← PII *\*stamb<sup>h</sup>Ha-s* ‘pillar’ > OI *stambha-* ‘Pfosten, Pfeilen, Säule’, from the root *stamb<sup>h</sup>H-* ‘befestigen, stützen, festmachen’, OP Pr. *stabava-*, Av Ni. *stambana-* < PII *\*stamb<sup>h</sup>H-* < PIE *\*stemb<sup>h</sup>H-*;<sup>39</sup> cognates in Baltic: Lith *stámbas* and Greek *στέμβω, άστεμφής* ‘inébranlable, immobile’, Tocharian B *śama*, A *śäm* ‘stand’ (EWAia II: 753–754; LIV<sup>2</sup> s.v. *stemb<sup>h</sup>H-*; Werba 1997: 383–384 No. 476 s.v. *stamb<sup>h</sup>*; Chaintraine 2009: 122, 1015, 1018; Derksen 2015 s.v. *stambas*)

(Kalima 1933: 128–; Uotila 1973: 7; 1977: 95; Koivulehto 1999a: 230 = 2016: 232; SSA III s.v. *samm*s; Parpola 2005: 47–48; 2012: 161; 2015: 66)

This Indo-Iranian etymology is probably correct: semantically the Indo-Iranian original is very close to the Finnic word, and there are no phonological problems either: the sound substitution *\*s-* ← *\*st-* has convincing parallels in Germanic and Baltic loans in Finnic (Junttila 2015: 171; LÄGLOS I: XX) However, the numerous derivations and words whose connection to *samm*s is etymologically unclear make the etymology more complicated.

Also a Baltic etymology from the reflex of the same Indo-European word has been proposed by Nuutinen (1987: 55–56), who assumes that Proto-Baltic *\*stamba-* (> Lith *stamba* ‘stem of a plant’) is the source of the Finnic word. From the point of view of phonology, both are equally likely. Nuutinen’s idea has been supported by Koivulehto (1992: 176). The Baltic cognate’s meaning suits the Finnic word less well, but Koivulehto has argued that the word could have been acquired from a stage of Baltic where the more archaic meaning of ‘pillar’ would have still been retained. If the original meaning of the Indo-European verbal root was ‘befestigen, stützen, festmachen (to support, to bind)’, which is the meaning of the Indic verb, it is possible that the Baltic word also originally had a meaning akin to ‘pillar’, and that ‘stem’ is indeed a later development, as Nuutinen suggested. However, LIV<sup>2</sup> gives a more abstract meaning ‘sich stützen, sich stemmen (Derksen: fix, support)’ for the Indo-European verbal root, and in the light of this semantic reconstruction it seems unnecessary to argue that a meaning ‘pillar’ would have to be primary for the Baltic noun *\*stambas*.

---

<sup>39</sup> LIV<sup>2</sup> (588–589, s.v. *\*steb<sup>h</sup>-*) mentions that the root *\*steb<sup>h</sup>-* ‘erstarren’ is not easily distinguished from *\*stemb<sup>h</sup>H-*, and Baltic verbal forms like Lith *stebù* ‘in Erstaunen geraten’ (< nasal-infix present *\*st-ŋ-b<sup>h</sup>-*), *stabaù* ‘anhaltten, hemmen, hindern’, *stebiù* ‘betrachten, beobachten’ reflect the root *\*steb<sup>h</sup>-* rather than *\*stemb<sup>h</sup>H-*.

Later Koivulehto (1999a: 230) mentioned *sammās* among the Indo-Iranian etymologies, but he does not discuss the etymology in further detail. Parpola (2005: 47–48) has supported the Indo-Iranian etymology. Both etymologies are formally faultless, but there is no reason to reject the Indo-Iranian etymology in favour of the Baltic one. There are many Indo-Iranian borrowings connected to mythology found in the Western Uralic languages (see *\*juma*, *\*tajwas*), and it can be assumed that *\*sampās* belongs to the same layer of borrowings.

At the time when the Baltic option was suggested, there were not yet many Indo-Iranian borrowings known from Finnic only, and this was seen as support for the Baltic etymology of the word. But since then, especially thanks to the works of Koivulehto, it has been clear that there are a number of borrowings found in Finnic only. It is thus no obstacle to consider this an Indo-Iranian loanword.

The syllable *-as* is attested in a number of borrowings, which are usually found in only a couple of (westernly) branches. It is possible that *\*as*-stems were not yet possible in Proto-Uralic, although the apocope in most eastern branches makes it more difficult to determine this. The same ending is also attested in some Balto-Slavic and many Baltic borrowings (see *\*porćas*). It is possible that the loans with *\*-as* are later, perhaps Iranian loans, borrowed before the *\*s > h* change took place in Iranian. The root *\*stemb<sup>h</sup>H-* has reflexes in Iranian branch (Av *staβra-* ‘strong, solid’, MP *stabr* ‘strong, big’), but a noun corresponding to OI *stamb<sup>h</sup>a-* is not attested in any Iranian language.

The Finnic word has also been derived from other Indo-Iranian words, which I find much more unconvincing: *skāmb<sup>h</sup>a-* ‘Stütze, Strebepfeiler’, (Katz 1985: 307; 2003: 244; Kuzmina 2007: 56). This explanation is also mentioned by EWAia (II s.v. *SKAMBH*). The substitution *\*s ← \*sk* is not very likely, as there are no parallel examples to this, and the traditional etymology is clearly better.

Uotila (1973) has derived the Finnic word from Proto-Indo-Iranian *\*śamba-* ‘Stange, Keule’. This would phonetically be an equally fitting etymology, but the word does not have a solid Indo-Iranian etymology. EWAia considers it possible that the noun is derived from Indo-Iranian root *\*śam-*. The formation *\*śamba-* is attested only in Indo-Aryan (EWAia II: 612–613, s.v. *śamba-*, *śamyā-*). There is no compelling reason to reject the earlier etymology, but Uotila’s idea remains possible too.

Various other Finnic words have been connected etymologically with *\*sampās* in earlier research. The words for ‘frog’ (Fi *sammakko*, Veps *zamba*), and their possible but irregular Saami cognates, such as North Saami *cuobbo* (< ? Pre-Sa *\*čampaj*), have been connected here by Setälä (1902–03: 146–147) and this possibility is also mentioned by the SSA, but it is probable that these two words have nothing to do with Proto-Finnic *sampās*. See Holopainen (2018b: 142–146) for a discussion of the Saami words. The idea that the *sampo*, the cornucopia-mill in Finnic mythology, is derived from *sammās*, has a long history, and this view is probably correct (see SSA s.v. *sammās* with references). Also the fish name *sampi* (: *sampe-*) ‘acipenser sturio’ has been etymologically connected to *sammās* (Uotila 1973: 18–19). Uotila lists parallel examples from other European languages, where the name of this fish denotes a mythologically important creature. Some terms in Greek, Turkic and other languages bear a resemblance to the Finnic word, and Uotila assumes that this can perhaps be a Wanderwort.

Saami *soabbi* (< Pre-Sa *\*sompā* or *\*sampi*) is, according to Sammallahti (2001), a parallel loan from Proto-Indo-European *\*stomb<sup>h</sup>Ho-* or a reflex thereof in some archaic

Indo-European language. An Indo-Iranian etymology is also possible for the Saami word, but there is no compelling reason to assume that this word is borrowed specifically from Indo-Iranian.

An Indo-Iranian etymology

**Md E *šumor'd'e*** ‘to worry, grieve’; **Mari *šuma-*, *šâma-*** ‘to become tired, languish’ (< PMa *\*šûma-*); **Ko *šumav-*, Ud *šuma-*** ‘be hungry’ (< PP *\*šuma-*); **Hu *szomorú*** ‘sad’, ***szomjas*** ‘thirsty’ (UEW s.v. *šom<sub>3</sub>*; Aikio 2014b: 88–89; 2015b: 61)

< PU *\*ćoma-*

← *\*ćamH-* > **OI *sám<sup>i</sup>*** - ‘to be extinguished, die out’ < PIE *\*k<sup>ʷ</sup>emh<sub>2</sub>* (EWAia II: 610–611 *sám<sup>i</sup>*; Werba 1997: 319–320 No. 353 s.v. *sám<sup>i</sup>*)

(Aikio 2014b: 88–89)

Aikio’s recent etymology is convincing based on both phonological and semantic criteria. The UEW gives two different entries, *\*šom<sub>3</sub>* ‘hunger, thirst’ (with the Mari and Udmurt reflexes and Hungarian *szomjas*), and *\*šom<sub>3-r<sub>3</sub></sub>* ‘trouble, sorrow’ (> the Mordvin word and Hungarian *szomorú*), but Aikio argues that both semantically and phonologically these two entries can be combined. It remains unclear whether Mordvin *šumor'd'e* and Hungarian *szomorú* include an identical derivational suffix or not, but this has no relevance for the Indo-Iranian etymology.

This word has to be an *o*-loan: the Mordvin and the Udmurt word could point to either an *\*o-a* or *\*a-i*-stem, but the Hungarian vowel *o* cannot be regularly derived from earlier *\*a* (here one would expect Hungarian *a* or *á*), and Mari *u* also more frequently reflects Uralic *\*o* than *\*a*, even though the conditions for the reflexes of *\*o* and *\*a* are not very well known. Here the *o* substitution could have been influenced by the labial consonant *m*.

As the Uralic word stem has verbal reflexes in most branches, it is probable that the Indo-Iranian verb was borrowed as a verb into Proto-Uralic. The UEW assumes that *\*šomV-rV* was a *nomen verbum* in the proto-language. There is no evidence for an underived noun *\*šoma* in the daughter languages, so it is more probable that the word was simply a verbal stem, and the Hungarian adjectives are later derivatives.

Fi *sammua* is possibly borrowed from the same Indo-Iranian word (see the entry).

An Indo-Iranian etymology

**Fi *suoja* ‘shelter; warmth’, SaN *suodji* ‘shadow, shelter’; ? Mari *šai-* ‘back’;  
? Ko, Ud *saj* ‘shelter’ < PP \**saj***

< ? PU \**sajV*

← Iranian PI \**stsāyāH-* < PII \**šcaH-yaH-*, OI *chā-yā-* ‘Schatten; Widerschein, Abbild’, YAv *a-sāiia-* ‘the one who has no shadow’, MP *s’dk’ /sāyag/ id.* (EWAia II s.v. *chāyā-*; Lubotsky 2001a: 35)

(Joki 1973: 319–320; Rédei 1986c: 57–58; Koivulehto 2001a: 363–364 = 2016: 262–263)

The Iranian etymology for the Uralic word is an old idea, but it has been discussed in more detail by Koivulehto (2001a), who argues that the word is probably a late, post-Proto-Iranian borrowing because it already has \*s (resulting from PII \*šć, as argued by Lubotsky 2001a). The correspondence Uralic \*s ~ Iranian s (< PII \*šć) is plausible. Finnic s could also reflect earlier \*š, which could be the substitute for Proto-Indo-Iranian \*šć. Also the Saami s could be explained from this, but not Komi s. It seems, however, that the Komi and Mari words are parallel loans, as they cannot be derived from a Uralic form \**saja*. The Finnic and Saami words are regular cognates.

The etymology of the Finno-Saamic word involves some problems, however. It is clear that Finnic *oo* (> Finnish *uo*) in \*i-stems is a result of Proto-Finnic lengthening of PU \**a* (Aikio 2012a), but there is no such phenomenon attested in \*a-stems. Pystynen (2018: 67–68) has recently argued that the word is a relatively late borrowing into Finnic, acquired during the change from \**a* > \**aa* (which later yielded *oo* through the development known as “Lehtinen’s Law” [Pystynen 2018: 60]) had already taken place, and long \**oo* was thus already a possible substitution. This explanation is possible, although there are not many other late loans that could provide parallel evidence for such late loans into Proto-Finnic. It also remains unclear why \**oo* was the sound substitution here (instead of \**a* or short \**o*).

Koivulehto has also derived the verb \**šoji-* > Fi. *hoi-taa* ‘to take care of’ from the same Indo-Iranian word (as an earlier borrowing from a Proto-Iranian form \**stsaHyā-*; see the entry). Koivulehto remarks that as the Saami word has the meaning ‘shadow’, it can be assumed that the same meaning was found in Finnic too, originally.

? Indo-Iranian etymology

**Fi *suola* ‘salt’; ? Md *sal*; Ma *šind’ž’al*, *sandzal*, *šñd’ž’al*; Ko *sol*, Ud *šilal id.*  
(< PP \**sol*)**

< ? \**salV*

← ? PII or PI \**sal-* (? \**sar-*), attested in *salilá-* ‘an epithet of the ocean’ (< ? \*‘salty’) < PIE \**sal-* (EWAia II: 712–713 s.v. *salilá-*; NIL: s.v. \**sal-*, 586–590; Kölligan 2018: 2254)

(FUV: 137; KESK: 258–259; Joki 1973: 316; Katz 1985: 171; 2003: 141; Napol’skikh 2015; Pystynen 2018: 67–70)

The words for ‘salt’ in various Uralic languages are clearly borrowed from Indo-European languages, and there is an old idea that the words in the Finno-Permic languages are borrowed from Indo-Iranian. However, this is not a very good idea, as the Indo-European ‘salt’ is poorly attested in Indo-Iranian. It is unclear whether Old Indic *salilá-* (which appears in the Rígvēda [10,109,1]) really comes from the meaning ‘salty’, as the word denotes ‘sea’. It has however been argued by Thieme (1953) that the word is originally an adjective with the meaning ‘salty’. EWAia (II: 712–713) considers this possible, and also the possibility of borrowing into Ob-Ugric (!) is mentioned by EWAia, with references to Joki (1973: 316) and the Habilitationsschrift by Katz. The same Indo-European etymology for *salilá-* is also found in NIL, where also the Vedic word *sarirá-* N. ‘Meeresflut’ is mentioned, which could reflect earlier *\*sal-iRo-*. Katz (1985: 171; 2003: 141) considers the ‘salt’ words to be Indo-Iranian borrowings, but he assumes that the differences in the vowels of the Uralic forms are due to Uralic ablaut. It is certain that the forms in the different languages cannot be derived regularly from a Proto-Uralic form, and this clearly points to parallel borrowings from Indo-European.

Burrow (1976: 62) has criticized the etymology because in his view, the OI word does not mean ‘salty’, and he does not consider the Indic word as a cognate to the ‘salt’ words of other Indo-European languages. Burrow assumes that the Uralic word could have been borrowed from Tocharian (cf. A *sāle*, B *sālyi*). Also KESK assumes that the loans are not from Indo-Iranian, but that the Finnic word is borrowed from Baltic whereas the Permic words represent an “Old Indo-European loan”.

Kölligan (2018: 2254) supports the connection of OI *sar-ít-* ‘river’ here, although the semantics of this connection is rather shaky: it is assumed that the word continues PIE *\*sal-hi-t-* ‘going to the salt’. Even if the connection of Vedic *salilá-* with the Indo-European word for ‘salt’ can be upheld, it does not mean that this is automatically the best source for the Uralic words. Among the Proto-Indo-Iranian or Proto-Iranian loan etymologies that are convincing, PIE *\*l* is always reflected as Uralic *\*r*, and this word would be a counterexample to that rule. But it is much more convincing to derive the Uralic words for ‘salt’ from some Indo-European language with retained *\*l*.

Liukkonen (1999: 133–4) considers Finnic *\*soola* to be a Baltic loanword (he mentions that this has been suggested already by Būga [1908], albeit based on false arguments). Referring to earlier research Liukkonen argues that Finnic *suola* and the more easterly Uralic forms are parallel borrowings from another Indo-European language. The *a*-stem with long *\*ō* is a sign of a more recent loan according to Liukkonen. The Finnic *a*-stem is unexpected from the Baltic original according to Liukkonen, who supposes that the word is borrowed from a collective *\*sālā*. Pystynen (2018: 67–70) has suggested that the Baltic word was borrowed into Finnic after the change *\*a > \*aa* (from which later Proto-Finnic *\*oo*) had taken already place.

Also the Mordvin word can be borrowed from Baltic, as *sal* could reflect earlier *\*sīli* or *\*sala*. The Mari word *šinžal* is an opaque compound according to Bereczki (2013). It can be analysed as *\*san-cal*, and Bereczki argues that *san* is from “Iranian or Caucasian”, Ossetic *cæn* or Avar *can* ‘salt’, whereas the latter part is from the Uralic ‘salt’ word. As it is very doubtful whether the word in question can be considered Uralic (or Finno-Permic), it is possible that also the Mari word is a parallel loan from somewhere. The idea that *\*san* reflects Ossetic *cæn* is possible but not unproblematic, as one would expect a Mari affricate or *\*š* in this case.

In any case it is clear that none of the forms that have traditionally been derived from Finno-Permic *\*sala* or the like are borrowed from Indo-Iranian. The idea that these words are from Tocharian is not very convincing neither. There are no known examples of Tocharian loans in Permic, Mari or Mordvin, and as these languages could have been in contact with the predecessors of Tocharians only in very early times before their migration to Central Asia, the words could at best be Pre-Tocharian loans. It looks much more likely to assume that all these forms are in some way borrowed from Baltic, either as direct borrowings or through some intermediary. As ‘salt’ is a very typical cultural borrowing, it is probable that the word has spread as some kind of a Wanderwort whose exact route is difficult to track.

Not an Indo-Iranian etymology

**Fi *suoli* ‘intestine’ (cognates in all Finnic languages); SaN *čoalli* < PSa *\*ćōlē* (cognates in all Saami languages except SaA); Md E *śulo*, M *śulǎ*; Ma *šolo*, *šolâ* (< PMa *\*šolǝ*); Ud *sul*, Ko *suv* (< PP *\*sul*); Khanty North, East *sol*, South *sut* ‘intestine’ (< PKh *\*saA*)**

< PU *\*śali* (UEW s.v. *śola*; Sammallahti 1988: 545; Aikio 2015b: 53)

← PII *\*ćālā* > OI *śā́lā* ‘hut, room’ (EWAia II: 631)

(Koivulehto 1999a: 217, 218 = 2016: 219, 220)

This etymology was first presented by Koivulehto (1999c) and it is mentioned as a possible etymology by SSA (s.v. *suoli*). Koivulehto considers this word either an Indo-Iranian or earlier Indo-European borrowing (in Koivulehto’s view, Uralic *\*ś* could also reflect PIE *\*k’*). It was traditionally assumed that the Uralic word was an *\*o-a*-stem, but the study of Aikio (2015b) on the reflexes of *\*a-i* and *\*o-a*-stems in Finnic, Saami and Mordvin has convincingly shown that this word has to be reconstructed as an Uralic *\*a-i*-stem. The new reconstruction makes the Indo-Iranian etymology somewhat more unlikely, as there seems to be no reason for the choice of an *\*i*-stem here. The *\*i*-stems are in any case very few among the early borrowings from Indo-Iranian. The initial syllable vowel is not a problem, as either Uralic *\*o* or *\*a* could substitute for PII *\*a*. In the case of Uralic *\*a* it would be very unlikely that the word was borrowed from an Indo-European form with *\*o*. The etymologies which point to retained PIE *\*k’* are also for the most part unconvincing (Holopainen: forthcoming).

The etymology is also not semantically straightforward, as there is no meaning ‘intestine’ attested for the Indo-Iranian word. The etymology has been criticized by Rédei (2002) because of this. Koivulehto mentions that the meaning is attested in Greek cognates of the Indo-Iranian word, but Beekes (2010: 735, 739) considers the Greek words *κολεον* ‘sheath’ and *κόλος* ‘large intestine’ possible substrate words and doubts their Indo-European etymology (also Frisk 1954–1972, Chantraine 2009 and Van Windekens [1986: 126 s.v. *κόλον*] state that the etymology of the Greek word(s) is unclear). Also de Vaan (2008: 150) considers Latin *culleus* ‘leather sack’ (which is also mentioned by Koivulehto as a possible cognate) and the aforementioned Greek words possible parallel substrate

borrowings (de Vaan does not mention the Indic word). EWAia (II: 631) and de Vaan (2008: 104) mention Latin *cella* as a possible cognate of the Indic word *śā lā*.

Interestingly, if the etymology was corrected, it would be the only example where both Uralic and Indo-Iranian show *\*l* in the place of PIE *\*l*. In almost all cases the loanwords in Uralic show *\*r*, lending support to the view of Mayrhofer (2004: 158) that the *l* forms in Indic are secondary.

Considering all this, one has to reconsider whether it is still worth supporting this etymology. The stem vowel alone is not sufficient reason to reject this, as similar problems are involved with the etymology of *\*śarwi* (see below), but combined with the unclear Indo-European background of the Indo-Iranian and Greek words, it means that this etymology does count as credible evidence for a substitution PII *\*-a* → PU *\*-i*.

Although the Uralic word does not have an attested cognate in Samoyedic, according to Aikio (2014: 75) the Proto-Yukaghir word *\*solije* ‘intestine, gut’ (> Old Yuk <*šolje*>) is borrowed from the Samoyedic cognate of *\*śali* (which has since disappeared from the attested Samoyedic languages). Koivulehto reconstructs the hypothetical Proto-Samoyedic form as *\*śālā*. Although the probable earlier existence of the word in Samoyedic does not give extra evidence for or against the Indo-Iranian etymology, it is interesting to note that if the etymology was correct, it would be one of the few examples of Indo-Iranian loans attested both in the Samoyedic branch and in the other branches that are traditionally grouped as the “Finno-Ugric” macro-branch.

Not an Indo-Iranian etymology

**Hu *ár* ‘flow; flood’; Mansi East *tūr*, North *tūr*, West *tūr* (< PMS *\*tūrə*) ‘lake’; Khanty East *lar*, North *lar*, South *tor* (< PKh *\*āār*) ‘lake’**

< **POUg** *\*θēra* (Zhivlov 2006: 163) < **PU/Pre-Ug** ? *\*sara* ~ *šara* (? *sora*) (UEW: 843–844 s.v. ; Zhivlov 2014: 127)

← **PII** *\*śáras-* > **OI** *sáras-* ‘See, Teich’, **YAv** place-name (‘mit Gewässern versehen’) *haraxvaitī-* = **Ved** *sarasvatī-* (< *\*saras-wat-iH-*); < **PIE** *\*sél-os-* (EWAia II: 709; Werba 1997: 256–257 No. 221; Stüber 2000)

(Munkácsi 1901: 68; Harmatta 1977: 171; Katz 1985: 119–120; 2003: 102; Koivulehto 1999a: 215 = 2016: 217; J. Häkkinen 2009: 22)

This is a good etymology (one of the rare etymologies for Ugric words suggested by Koivulehto). There is no compelling reason to search for a foreign origin for the Ugric word for a basic concept such as ‘lake’ or ‘flood’, but it resembles the Indo-Iranian word so much that a coincidence would be strange. It is known that almost anything can be borrowed in intensive contact situations (Koivulehto 2001b: 293; K. Häkkinen 1987: 244), and there are parallels of such words being borrowed: WOLD (s.v. *the lake*) lists several cases of the word for ‘lake’ being borrowed, for example English *lake* from French *lac*, Vietnamese *hồ* from Chinese *hú*.

Although the etymology itself is convincing, the vowel substitution is unclear because of the troubling and complicated development of the Ugric vowels. The Ugric cognates here

seem to be regular and probably go back to an earlier *\*a-a*-stem, as is reconstructed by Zhivlov (2014). Honti (1982: 139, No. 158) reconstructs the Proto-Ob-Ugric word as *\*θūr₃*, but this does not tell us much; the reconstruction seems incorrect, as Proto-Khanty has *\*ā* (*a* in Honti's reconstruction), and Proto-Mansi *ū*, which is frequent in Uralic *\*a-a*-stems. Zhivlov (2006) reconstructs the Proto-Ob-Ugric forms as *\*θeṛ₃a*, and the Uralic/Pre-Ugric pre-form is reconstructed as *\*sara* (or *\*šara*) by Zhivlov (2014: 127). According to Zhivlov (2014) this is one of the words for which *\*a<sub>l</sub>* should be reconstructed for the non-initial syllable, as in his view the divergent Khanty reflexes of Uralic *\*a*-stems justify the reconstruction of two different stem vowels in the place of traditional *\*-a*. This reconstruction has not acquired much of a following, but here it suffices to say that in any case the relationship of the Ugric words is regular. In Indo-Iranian the word is an *s*-stem, but this has been borrowed into Uralic as an *a*-stem, like most of the other *s*-stem loans.

However, the Ob-Ugric and Hungarian words could equally well be derived from earlier *\*sora* (or *\*šora*), as Mansi *\*ū* and Khanty *\*ā* are regular reflexes of this stem type as well (see the discussion under the entry *\*mēja*).

Katz (2003: 102, footnote 81) argues that Hungarian *ár* probably does not belong here, as the 'flowing water' does not fit semantically very well, and a mediaeval attestation (the Bécsi Kódex) of the verb *arrad* (= modern Hungarian *árad* 'to flood') with a geminate *-rr-* points to a pre-form with *\*-rk-*. A. Molnár (2003) has argued that Old Hungarian geminate resonants are secondary and do not correspond regularly to old clusters of a resonant and *k*. In any case, a single attestation of a derivation in an Old Hungarian source is not enough to refute the connection of the Hungarian word with the Ob-Ugric ones.

If the vowel reconstruction is correct, it is interesting that it is *\*a* that is the Uralic vowel here, as in some other loans into Ugric, the Indo-Iranian *\*a* is substituted by *\*ā*, and in many early loans, Indo-Iranian *\*a* is reflected by Uralic *\*o* in front of *\*r*.

J. Häkkinen (2009: 22) has attempted to connect Permian Ud *šur*, Ko *šor* 'lake' (< PP *\*šor*) and Finnic *sarajas* 'river of the dead' to the Ugric word, deriving them from a common Proto-Uralic form, but this is impossible because of irregularities in the consonantism and vocalism, and because the Finnic word is a hapax. The Permian words reflect PU *\*šerV* according to the UEW (s.v.). A connection of the Ugric and Permian word was already suggested by Harmatta (1977: 171).

An Indo-Iranian etymology

### **Fi *sarajas* 'mythological river in the land of the dead' (only in old folk poetry, not found in other Finnic languages)**

**? ← Iranian *\*zraya-* (= Proto-Iranian *\*dzraya-*) > Av *zraiih-*, OP *drayah-* 'sea' (EWAia I: 606–607, s.v. *jráyas-*)**

(Setälä 1912: 189; Jacobsohn 1922: 122–123; FUV: 141; Joki 1973: 151; Katz 1985: 119–120; SSA s.v. *sarajas*; Blažek 2005: 175; J. Häkkinen 2009: 22)

Setälä's etymology is both phonetically and semantically unproblematic, but since this word is attested only in Finnish folk poetry and in no other Finnic or Uralic language, it is dubious whether this can really be borrowed from an Iranian language.

Komi *sarid'ž* 'sea', Udmurt *zarid'ž* 'sea; a warm (southern) region where birds migrate for winter' are true borrowings from this Iranian word (this was established already by Munkácsi 1851: 382; see the entry in Appendix I). J. Häkkinen (2009) has tentatively attempted to derive the Finnish word from the same Uralic source as the Ugric *\*sara* (see the entry), but this seems very unlikely. Instead, the Ugric word is borrowed from a different Indo-Iranian word (PII *\*sarás-* from PIE *\*selós-*).

Not an Indo-Iranian etymology

### **Ko *sar* 'sea'**

< **PP** *\*sar* (Csúcs 2005: 374)

← **PII/PI** *\*saras-* > **OI** *sáras-* 'lake'

(Joki 1973: 310; Rédei 1986c: 76; Lushnikova 1990: 294; Csúcs 2005: 374)

See also *\*sara*

This is one of the more troubling etymologies, as Permic *s* can point to an old borrowing, whereas the *a* vocalism speaks in favour of a late loan. In the Andronovo Aryan theory of Helimski and Zhivlov, forms that retain *\*s* are borrowed from this Andronovo language. Among the three loans which point to a late loan with *s*, this is the most troubling, as *\*sur* 'beer' can be a Wanderwort, and the sibilant in *\*šumiš* can perhaps be explained otherwise (see the respective entries).

However, both Joki (1973) and Rédei (1986c) also mention the possibility that Komi *sar* is just a shortened form of *sarid'ž* 'see'. The word *sarid'ž* has an irregular cognate in Udmurt, where the word is *zarid'ž*. The Komi and Udmurt words were borrowed from Iranian *\*zraya-* (compare Av *zrayō*), probably after Proto-Permic times, as the Komi and Udmurt sibilants point to different substitutions: Komi shows *s*, as *z* was possibly not yet allowed when the word was borrowed, and Udmurt shows *z* which corresponds better to Iranian *z*. The *-id'ž* suffix of the Permic words does not correspond to any Iranian suffix, so it can even be assumed that the word was borrowed as *\*sar*, *\*zar*, and the suffix was added later.

If the Komi and Udmurt words really are parallel borrowings, this has implications for the dating of later loans into Permic. There are a number of loans where Permic shows *z* as the substitute of Iranian *z*. Among these *zon* 'son' is confined to Komi, and *zar'ni* is reflected by both languages. It can be assumed that in loans that were acquired into Proto-Permic, *z* did not yet exist as a phoneme, and the words that show *z* have been borrowed after the Proto-Permic language had split up. *zar'ni* in Udmurt and Komi could then reflect two parallel borrowings.

? Indo-Iranian etymology

**Fi *sarvi* (cognates in all Finnic languages); SaN *čoarvi* (cognates in all Sa languages); Md *E šuro*; Ma *šur* (< PMa \**šur*); Ko, Ud *šur* (< PP \**šur*); Hu *szarv*; ? Ms North *sōrp* ‘male elk’, East *šōrp* (< PMs \**šārəp*), id.**

< PU \**šarwi* (UEW s.v. \**šorwa*; Sammallahti 1988: 549; Zhivlov 2014: 130; Aikio 2015b: 53)

← PII \**črwā-* > Av *sruuā-* ‘nail, horn’ (Nussbaum 1986: 3–4, 16, 100, footnote 136)

(Jacobsohn 1922: 124–125; FUV: 136; Korenchy 1972: 69–70; KESK: 274; Joki 1973: 311; Katz 1985: 135; 2003: 113; Gamkrelidze & Ivanov 1984: 925; Rédei 1986c: 47; Lushnikova 1990: 260–26; SSA III s.v. *sarvi*; Koivulehto 2001b)

The Indo-Iranian etymology for this Uralic word is transparent, and this is one of the best-known examples of an Indo-Iranian loanword in Uralic (see Joki 1973; Rédei 1986c). As in the case of \**šali* (see the respective entry), it was traditionally assumed that the stem vocalism in Uralic was \**o–a*, but Aikio (2015b) has convincingly argued that these go back to an \**a–i*-stem in Proto-Uralic. Due to the obvious resemblance and semantic similarity of the Indo-Iranian and Uralic words, this etymology cannot be rejected, unlike \**šali* which might be a false etymology.

The Uralic words are regular cognates, but the Mansi word \**šārəp* is not mentioned by Aikio. A similar word is also found in Khanty South *t’ārpə*, *šārpə*, North *šārpi* (< PKh \**čərpä*). Honti (1999: 129) and the UEW consider the Ob-Ugric words to be cognates with one another and reflexes of the same Uralic word as the other Uralic forms listed above. The East Mansi form *sōrp* has an irregular sibilant, but as the other Mansi forms reflect *s/š*, we can assume that the KO development is secondary. The Mansi vocalism can be regularly derived from Uralic \**a–i* (compare PU \**kali* ‘to die’ > PMs \**kāl*, \**kaji* ‘hair; stalk’ > PMs \**kāj*, but the consonant cluster \**-rp-* raises some questions. Usually \**w* does not develop to \**p* in Mansi, but there are no examples of Uralic \**-rw-* words that have reflexes in Mansi. The kinship-term \**arwV* is one possible example, but this is not a very good etymology (see the respective entry).

For Khanty we do have at least some examples of Uralic \**-rw-*, and the form \**čərpä* (this would be the Proto-Khanty reconstruction following the principles of Zhivlov 2006; Honti reconstructs the form as \**čěrpä*) is irregular in this respect. Also the fronted vocalism of Khanty is irregular (it cannot reflect earlier \**a*), the non-initial syllable vowel is unexpected, and so is the affricate \**č*. This means that the Mansi and Khanty words are not regular cognates (even Honti 1982: 206 lists them among the words that cannot be reconstructed for Proto-Ob-Ugric), and it is possible that they reflect two separate parallel borrowings from (Indo-)Iranian. It remains unclear why Khanty has \**č*: it is uncertain whether PU \**č* can be regularly reflected by Khanty \**č*, as many of the cases traditionally reconstructed with \**č* yield Khanty *s*, but some counterexamples remain (Honti 1999).

Katz (1986: 134; 2003: 113) argues that the Ob-Ugric words reflect a loan from PII \**čarabhá-* ‘a kind of animal (< \*horned animal’) (> OI *šarabhá-*, EWAia II s.v.). PII \**čarabhá-* is probably derived from the same root \**čar-* ‘head, horn’ (< PIE \**kér-*) as \**črwā-*. The same has been suggested by Blažek (2005: 180). Katz’s etymology is also plausible, but because the formation \**čarabhá-* has no traces in Iranian, we cannot be certain of its Proto-Indo-Iranian age. Also the Khanty vocalism is difficult to explain from a

form with *\*a*. Deriving the Ob-Ugric forms from *\*ćarab<sup>h</sup>á-* remains a possibility, but this etymology is not better than assuming that the Khanty and Mansi words reflect early parallel loans from *\*ćr<sub>wa</sub>-*.

Riese (2001: 52) assumes that Mansi *śōrpi* is a reflex of the PU (= FU) word for ‘horn’ word but includes the Mansi suffix *\*-pī-/pī-*, a reflex of the similar Ob-Ugric suffix which according to Riese has both denominal and deverbal functions. Korenchy notes that Mansi *-p* and Khanty *-BV* are derivational suffixes for Nomina possessoris, the semantic development being ‘Männchen von Elentier’ ← ‘Gehörntes’. In inherited words there are only two examples of the use of this suffix (according to Riese), the other being South Mansi *tērpi* ‘Köder’, which might be inherited from Proto-Ob-Ugric. This makes this explanation sound dubious; also for the deverbal use, Riese lists only one example, PU (?) *\*ama-* ‘schöpfen’ > North Mansi *ūmpi* ‘Schöpfkelle’. The use and origin of the Mansi suffix needs more research, but at the moment it cannot be ruled out that the Mansi word really reflects a derivative. For some reason Riese does not reconstruct the word into Proto-Mansi and mentions only the North Mansi form. The Khanty word is not comment in Riese’s book in general, but the absence of Proto-Ob-Ugric reconstruction seems to imply that Riese does not consider the Mansi word as a cognate of Khanty *\*ćerpä*. The Ob-Ugric suffix is considered an “extension” of POug *\*p* by Riese; it is perhaps possible that they are the same suffix, and this seem to be the idea of Zhivlov who does not reconstruct *\*-pi* for the Proto-Mansi word. The suffix *-p* is attested in many Ob-Ugric words of Uralic origin.

As Kallio (2012a: 163–165) has rehabilitated the idea that the Uralic *\*i*-stems are actually *\*ə*-stems, it would be phonetically plausible to assume that the unstressed Indo-Iranian *\*-a* could have been substituted by Uralic *\*i* [ə]. Here the Indo-Iranian accent cannot be reconstructed on the basis of the Vedic stress as this exact formation is absent from Indo-Aryan, and the Avestan accent is less well known, although it can be reconstructed in some cases based on the different reflexes of stressed and unstressed *\*r* (Fortson 2010: 207–208). Although the form *\*ćr<sub>wa</sub>-* lacks parallels in Indo-Aryan (where a different formation *śr̥ṅga-* ‘horn’ is found instead), the formation itself is a regular reflex of a Proto-Indo-European derivational process, and there is no reason not to consider this word a Proto-Indo-Iranian loan (an Iranian source is anyway ruled out by the palatal sibilant of Proto-Uralic). Joki (1973) argued that the borrowing would be from an early form of Iranian, but this is criticized by Dolgopolsky (1989: 22) because of the palatal sibilant.

However, there might be another solution for the unexpected *\*i*-stem in this loanword. The word ‘horn’ often appears in the dual in (Young) Avestan texts, which is logical, considering that such animals tend to have a pair of horns. According to de Vaan, there are many analogous dual forms for the feminine noun *sruua-* in Young Avestan (for a detailed survey see de Vaan 2003: 405–407), but an early Avestan form for the dual accusative *\*sruuie* would be regular, and this points to Proto-Indo-Iranian form *\*ćr<sub>wy</sub>ai*. It is not difficult to imagine that a sequence like *-yai* could be substituted by Uralic *\*i*. The nominative dual form *sruuaē-* appears in a sandhi form (followed by the enclitic *-čā*), regularly reflecting PII *\*ćr<sub>wai</sub>-* (de Vaan 2003: 406). Either of these forms is a suitable original for Uralic *\*śarwi*. An Indo-Iranian dual form of ‘horn’ could have been easily reanalysed as a singular in Proto-Uralic during the time of borrowing, and in fact we do not know in what number the speakers of Proto-Uralic used the words for body parts (compare

the modern Hungarian way of using singular forms of words like *kéz* ‘hand’, *láb* ‘foot’ when referring to a pair of limbs).

An Indo-Iranian etymology (Ob-Ugic words are parallel loans)

**Md E *sazor*, *sazoro* ‘sister’, M *sazâr*, *sazârâ*; ? Mari *šuzar*, *šūžar*, W *šâžar* ‘jüngere Schwester’ (TschWB 755) (< PMa \*šūžar); Udmurt *suzer* id., Komi *sozor* ‘Fadenbruch (im Gewebe)’ (< PP \*sozer)**

< PU \**sasarV*

← PII \**swasar-* > OI *svásar-*, YAv *χ<sup>v</sup>anhar-* ‘sister’; < PIE \**s<sup>w</sup>esor-* ‘sister’ (EWAia II s.v. *svásar-*)

(Munkácsi 1901: 68, 79–80; Paasonen 1908–1910a: 69; Jacobsohn 1922: 181–182; FUV: 136; Joki 1973: 312; Harmatta 1977: 171; Katz 1985: 216; 2003: 175–176; UEW s.v. *sasare*; Rédei 1986c: 58–59; Helimski 2000: 192; SSA III s.v. *sisar*)

This established, old etymology is in principle credible in both its form and semantics, although the relatedness of the various forms in the Uralic languages entails some uncertainties. While Helimski (2000) still lists the Finnic word \**sisar* (> Fi, Est *sisar*, Karelian *sizâr*, Ludic and Veps *sizar* id.) as a cognate of this Indo-Iranian loanword, it is now a universally accepted view that the Finnic word is borrowed from Baltic \**sesor-* (> Lith *sesuõ* ‘sister’) (SSA III s.v. *sisar*). This is evident from the first-syllable vowel correspondences, as Finnic \**i* could hardly reflect PII \**a* but there are examples of Finnic \**i* as a substitute of PB \**e*. Helimski also did not take into account the fact that the Finnic word actually represents two parallel borrowings from Baltic (cf. Lith *sesuõ*), as the vowel correspondences within Finnic are not regular: in addition to the Finnic words presented above, the word appears in Estonian (*sõsar*), South Estonian *sysar* and Livonian *sezâr*, *süzar* which have to reflect earlier \**sesari* ~ \**sešari* and show that the word was borrowed separately into (at least) two Finnic dialects (Kallio 2018a: 255, footnote 6).

Joki (1973) presents the data for this etymology rather ambiguously. He supports the idea that the Mordvin word (he reconstructs the Proto-Mordvin form as *sasar-* or \**s<sup>u</sup>asar-*; the latter form is impossible because of Mordvin phonotactics) is borrowed from this Indo-Iranian form (Joki classifies it as “Aryan” or “early Proto-Iranian (frühuriran.)”). Joki supposes that both the Finnic words (the irregularity of which he does not comment on) and the Mari words can be derived from \**sesor* or \**sesar*. Joki mentions that these, as well as the Permic words (Joki reconstructs the Proto-Permic form as \**soz3r*), can be borrowed equally well from Baltic as from Indo-Iranian. But the reality is, as discussed above, much more simple: the Finnic words represent parallel borrowings from Baltic, and the rest of the words can be derived from the same Indo-Iranian origin. Here it is possible to argue that the word is an early borrowing into the proto-language level, and that in Finnic the word has been replaced by a later borrowing from Baltic. Deriving the rest of the Uralic words from Baltic is much more complicated, and an Indo-Iranian origin is preferable: Mordvin clearly reflects an early \**a*–*a*-stem, and so does Permic, making the Indo-Iranian

origin probable and the Baltic one impossible. Also no other Indo-European source could come into question because of the Uralic *\*a-a*-stem.

The etymology of the Mari word *šüz̄zar* (< PMA *\*sūžzar*) is given in TschWB with a question mark (“FP?/FW?”), implying the editors’ uncertainty about the relatedness of the word to forms in the other Uralic languages. Although no reasons for this are given, it does indeed look like the Mari word might be an independent borrowing from the same (Indo-)Iranian source as the Permic and Mordvin words. The problem here is that the development of the Uralic vowels in Mari is so poorly understood that it is difficult to determine what the regular reflex of PU *\*a* is in Mari (here in *\*a-a*-stems). However, the vowel correspondences of the Mari dialects in this particular word are very peculiar for an inherited word with *\*a* in the initial syllable (parallels are difficult/impossible to find), and assuming a parallel borrowing would explain this. Also Katz considers the Mari word as a separate loan, but because of his ablaut theory, he assumes that the Mari vocalism can be derived from PII *\*sasar-*. Harmatta (1977: 171) does not mention the Mari word at all, and considers only the Permic and Mordvin words reflexes of Proto-Finno-Ugric *\*sasar* (in Harmatta’s reconstruction and terminology). Niklas Metsäranta has stated in personal communication that the dialects with *ü* show a secondary development, and that the word can be reconstructed for Proto-Mari with *\*š̄*.

The idea that the words for ‘sister’ represent parallel borrowings from Indo-Iranian is also expressed by Katz, but his conclusions are different and hard to uphold, as he argues that Mordvin *sazor* corresponds regularly to Estonian *sõzar* and Votic *sezàr*, and that the Permic and Mari words are either different borrowings from Indo-Iranian or reflexes of different Proto-Uralic ablaut grades.

The Indo-Iranian word could be either a PII or PI borrowing, as *\*s* was possibly still retained in Proto-Iranian (Mayrhofer 1989: 7). There are not many borrowings shared by Mordvin, Mari and Permic, and if we suppose that the word was borrowed into a common proto-language of these branches, we have to argue that the reflex of this word was lost in Finnic, Saami (and possibly Ugric, depending on what kind of stance on taxonomy we take), or that the word was borrowed into neighbouring Proto-Uralic dialects that eventually developed to Mari, Mordvin and Permic.

As this Indo-Iranian borrowing does not have a reflex in Finnic or Saami it is difficult to say anything about the possible original vowel in the third syllable. In some sources this is reconstructed as *\*sasarV*, but there are really no arguments for reconstructing this word-final vowel, and Helimski (2000: 192) supposes that this word, together with a number of Indo-Iranian loans ending in *\*-as*, could be an early consonantal stem in Uralic; Helimski considers the emergence of consonantal stems a Finno-Permic innovation.

Aikio (2018) reconstructs consonantal stems already for Proto-Uralic. This word could probably have been borrowed as a consonantal stem, as the Indo-European original gives no reason to postulate a syllable-final vowel. The word for ‘sister’ does not appear in the list of back-vocalic words in Aikio (2015b), however.

Culturally this is an interesting borrowing. Many other kinship terms in Uralic languages are borrowed from the Indo-European neighbours (such as Finnic *äiti*, probably due to intensive cultural influence (as Milanova, Holopainen & Bradley [forthcoming] state, kinship terms are basic vocabulary and cultural vocabulary at the same time). However, interestingly among the early Indo-Iranian borrowings there are not many

kinship terms (see the entry *\*icü* for a dubious Indo-Iranian etymology), while in later Iranian loans into Hungarian some possible examples can be found.

Rédei (1986c) reconstructs the meaning of the proto-word as ‘younger sister’, as this is the meaning of the word in both Udmurt and Mordvin. However, the Indo-Iranian original had the more general meaning ‘sister’, and it can be argued that this word was borrowed into Proto-Uralic in a more general meaning of ‘sister’, and the later meanings referring to ‘younger sister’ developed later due to Turkic influence: both in Udmurt and Mordvin the words for ‘elder sister’ (Ud *apa(y)*, Md *aka*) are later loans from Turkic (Mészáros 2001: 172; Milanova, Holopainen & Bradley [forthcoming]).

An Indo-Iranian etymology (Mordvin, Permic)

**Fi *sata-* ‘to rain’, *sade* ‘rain’ (cognates in all Finnic languages); Ng *soru-*, Slk (Taz) *sēri-* ‘to rain’ (< PS *\*sārā-*)**

< PU *\*śaḍa-* ‘to rain’ (Sammallahti 1988: 540; Aikio 2015b: 56)

← PII *\*ćad-* ‘to fall’, OI *śad-* ‘abfallen, ausfallen’, perfect *śaśāda-* < PIE *\*k’ad-* ‘to fall’ (EWAia II s.v. *ŚAD*; LIV<sup>2</sup> 318 s.v. *\*k’ad-*)

(Koivulehto 1991: 50)

Koivulehto’s idea of an Indo-European origin of the Uralic word is convincing. However, instead of assuming that the word is borrowed from Proto-Indo-European, like Koivulehto did, with PU *\*ś* (*\*ć*) as the substitute of PIE *\*k’*, it can be assumed that the word is a later borrowing from Proto-Indo-Iranian, showing the substitution *\*ś* ← *\*ć*. In the lack of convincing Proto-Indo-European loanwords in Uralic (see Simon [in press]), the Indo-Iranian etymology is preferable. Semantically the Indo-Iranian etymology is equally plausible. The substitution of PIE *\*d* by PU *\*d* has few convincing parallels, but this is an equal hindrance to both the Proto-Indo-European and Proto-Indo-Iranian etymology.

The Finnic and Samoyed words are regular cognates, so there is no reason to doubt the Proto-Uralic status of the etymology. This is one of the few examples that show an early Indo-Iranian loanword having a distribution stretching to Samoyed.

An Indo-Iranian etymology

**Mari *šör*, *šəžer* ‘milk’ < ? PMA *\*šür***

← PII *\*kšiHrá-* ‘milk’, OI *kṣīrā-*, Oss *æxsir* (EWAia I 433 s.v. *kṣīrā-*)

(Katz 1985: 239; 2003: 193–194; Hyllested 2014: 20)

The “frühurarisch” etymology of Katz for this Mari word is not very convincing, as Katz assumed that *š* is a reflex of the Indo-Iranian laryngeal here, deriving the word from a Pre-Indo-Iranian form *\*kših<sub>2</sub>ró-*. The etymology has been criticized by Hyllested, who notes that Mari *šəžer* is secondary to the basic word *šör* (both are attested in Mari dialects in the

same meaning), and the second sibilant in the disyllabic form cannot reflect the Indo-Iranian laryngeal. One can agree with Hyllested, even though it remains obscure what the relationship of Mari *šör* and *šəžer* actually is. The latter looks like some kind of an opaque compound. Theoretically one could assume that Mari *šör* would reflect the Indo-Iranian original, with \*š as the substitution for \*kš-. The Proto-Mari vowel has to be \*ü here, if we follow the reconstruction of Aikio (2014: 143). This vowel usually reflects Uralic/Pre-Mari \*ä or \*i (Aikio 2014c: 155), neither of which is a known substitute of PII \*i, and this Mari development usually appears only before \*k.

The Mari word is certainly a loanword from somewhere, as it does not have a Uralic etymology but it is not at all obvious that it is borrowed from Indo-Iranian.

Not an Indo-Iranian etymology

**Fi *syteä* ‘to hit’, *syttyä* ‘to set on fire’ (cognates in all Finnic languages) (SSA s.v. *syttyä*)**

< Pre-Finnic ? \**seči*-

← Pre-Iranian \**tsewč-e*- > PI \**tsawč*-, cf. OAv *saocaiiat* ‘inflames’ (EWAia II: 655–656, s.v. *ŠOC*; Werba 243–244 No. 196; Cheung 2007: 338–340, \**sauč*)

(Koivulehto 1999a: 223–224 = 2016: 225–226)

The same Indo-Iranian root appears in Fi *huhta* < \**šukta* ← PI *tsuxta* (see the entry). If both etymologies are correct, *syttyä* has to be a later loan because here Finnic has *s*-, not *h*- (< \*š). Alternatively, the word could be a very early Indo-Iranian loan (from Pre-II \**čewč*-), if Finnic *s* reflected earlier \*š here. However, this etymology is much more problematic than the one for \**šukta*. Koivulehto assumes that this is a Pre-Iranian loanword, where the Iranian affricate \**ts* is reflected by Finnic \**s*, but this substitution is probably not correct, as it lacks convincing parallels and is not phonetically expected. The whole concept of Pre-Iranian loans with retained PIE \**e* and Iranian affricates \**ts* and \**dz* is based on rather scanty evidence (see also \**sentü*-, \**teksä*).

Koivulehto assumes that Finnic \**ü* has developed from an earlier sequence *-ew-*. The sequence \**ew* is usually retained as some kind of diphthong in Finnish as shown by such examples as PU \**lewli* (> Fi *löyly*) ‘spirit, steam’ (Sammallahti 1988: 545) and PFI \**kewhä* (> Fi *köyhä*, Est *kehv*) ‘poor’ (SSA s.v. *köyhä*). However, Aikio (ms.) argues that the *ü* vocalism can be explained through regressive assimilation caused by the \**ü* in the second syllable, so Koivulehto’s explanation of the vowel development is probably correct.

The Finnic vowel \**ü* could be explained also otherwise, as it could reflect PI \**u*. There are examples of Finnic *ü* corresponding to *u* in the donor languages among various Indo-European loans (see \**mükkä*). The Finnic word could reflect a zero-grade form (\**suč*-) in Iranian. In Old Indic there are nouns which show the zero grade, such as *śuc*- ‘flame’ and *śuci*- ‘gleaming’. If the Finnic word was derived from such a form there would also be no need to suggest a “Pre-Iranian” origin, as the Finnic *s* could simply reflect later (Avestan-type) Iranian *s* and not PI \**ts* (although here again Finnic *s* could also result from earlier PII \**č*). While there are few examples of Iranian *s* (< \**ts*) in loanwords that are found in

only Finnic and other West Uralic languages, loanwords in Ugric, as well as later Ossetic evidence, show that the Iranian languages of the steppe had \*s (< \*ts) and \*z (< \*dz) like the majority of the Iranian languages.

From the Finnic point of view, it looks like *syttyä* is derived from the verb *syteä* ‘to hit’. This makes the Iranian etymology look more unlikely, as the Iranian verb clearly denotes burning, and semantically it would be dubious to derive the verb ‘to hit’ from this. The semantic hindrance is also noted by Aikio (ms.).

? Indo-Iranian etymology

**Md E *sed*, M *säd* ‘bridge, floor’; Ko *sod* ‘ladder, stairs’, *sojd*, *sojt* ‘bridge, footbridge’ (< PP ? \**sojt*)**

< PU \**sejti* ~ \**seti*

← Pre-II \**seh<sub>2</sub>itu-* > OI *sétu-* ‘binding; bond, fetter; dam, bridge’, Av *haētū-* ‘dam’, from the IE root \**seh<sub>2</sub>i-* ~ \**sh<sub>2</sub>ey-* ‘to bind’; cf. also Lith *saitas* ‘tie, leash, cord’, Proto-Slavic \**sěťŭ* ‘net’ < PBSl \**soi-t-o-*; (EWAia II: 745, s.v. *sétu-*; LIV<sup>2</sup>: 544 s.v. \**sh<sub>2</sub>ei-*; Werba 1997: 253–254, No. 216; Derksen 2015: 386 s.v. *saitas*)

(FUV: 136; Joki 1973: 313–314, No. 139; Katz 1985: 318; 2003: 254; UEW s.v. \**set<sub>3</sub>* (\**sejt<sub>3</sub>*); Lushnikova 1990: 257; Koivulehto 1999c; 2001b: 250 = 2016: 292)

This is a well-known and convincing etymology. Semantically the Uralic and Indo-Iranian meanings fit well. The Indo-Iranian root is derived from the root meaning ‘to bind’, but it is the meaning ‘bridge’ that has been borrowed into Uralic, and the more divergent meanings in Komi and Mordvin can be easily derived from this.

Although the etymology is limited only to Mordvin and Permic (and within Permic only to Komi), it has to be an early (Pre-Indo-Iranian) borrowing, as the Uralic vowel clearly points to retained Indo-European \**e*. Joki considers the word an Iranian borrowing from a language with retained \**s*, but this makes little sense because of the archaic vocalism (Indo-European \**e* could not have been retained until Iranian times). Koivulehto reconstructs the Uralic form as \**se(j)te*, not taking a stance on whether the Mordvin and Komi words reflect a diphthong/sequence \**ej* or not, and the same is done already by Joki, although his reconstruction of the stem vowels is the imprecise *ɜ* (it can by now be considered certain that the word was an \**i*-stem). Interestingly, Indo-European loans in original *u*-stems often became \**-i*-stems in Uralic, possibly because both are high vowels.

The Indo-Iranian word certainly had a diphthong, and as diphthongs of the type \**Vj* can be reconstructed for a number of Uralic words, there would have been no obstacle to substitute \**ey* by the identical Uralic \**ej*. In Mordvin, the \**j* of the diphthongs regularly disappears, as is shown also by the Indo-Iranian loan *vanoms* from \**wanaiti* (see the respective entry; see also the entry for \**pejmä* ‘milk’). In Komi one can see both forms with and without the retained \**j*. In Permic the Uralic *Vj* sequences are usually retained, cf. \**kajwa-* ‘to dig, spade’ > PP \**koj* > Ko *koj*, Ud *kuj*, \**majaka* ‘beaver’ > PP \**mɔjɨ* > Ko *moj*, Ud *myj* (Sammallahti 1988: 552; UEW s.v. *maja*). One can suppose that in front of \**t* the

development would have been different, but the problem is that even in the same dialects forms with and without *j* are found (such as Udora dialects *sod* ‘ladder’ and *sojd* ‘bridge’).

It is interesting that there are a couple of borrowings showing a “central Uralic” distribution (*\*mertä* ‘human’ is another such a case). As there is no such taxonomic entity, and at least no intermediary proto-language for Mordvin and Permic can be reconstructed after PU/PFP, then it is probable that these words have simply been retained in these branches but were borrowed in earlier times (into PU or an early daughter language thereof), or that the words in Permic and Mordvin are extremely early parallel borrowings. It is also easy to explain that reflexes of this word have been lost in some languages, such as Finnic, which borrowed its word for ‘bridge’ (Fi *silta*) from Baltic *\*tiltas* (the Finnic word has been further borrowed into Saami as PSa *\*šeltē*). In Samoyed, the word for ‘bridge’ (PS *\*pulā ~ \*pilā*) is a later borrowing from a different Iranian word (cf. MP *puhl* ‘bridge’ < PI *\*prtū-*, Janhunen 1983, Napol’skikh 2010: 230; see the entry), and Hungarian *híd* ‘bridge’ is a separate, obviously much later parallel loan from Alanic, cf. Ossetic *xīd, xed* (H. Sköld 1922: 23; Joki 1973: 263, No. 42.). It is, of course, impossible to prove the previous existence of a word in all these branches. It is theoretically possible that neighbouring dialects that later emerged as Mordvin and Permic borrowed a word from Indo-Iranian simultaneously, but the very archaic vocalism of this particular borrowing makes this kind of scenario unlikely.

According to Koivulehto (1991), Fi *sitoa* ‘to bind’ and Saami *čadnat* id. are also borrowed from the same Indo-European verbal root, possibly from Indo-Iranian. This is a convincing etymology, although some other Indo-European form could also come into question.

Katz (1985; 2003) argues that the Komi word *sojd* is borrowed from PII *\*samdhi-* ‘bond’. The traditional Indo-Iranian loan etymology is semantically more convincing.

An Indo-Iranian etymology

**Fi *syntyä* ‘to be born’ (cognates in all Finnic languages); ? Ko *sod-*, *sud-* ‘to increase’**

< ? *\*sen-tü-* ‘to be born’ (SSA s.v. *syntyä*; UEW s.v. *sente-*)

← “Pre-Iranian” *\*dzenH-* > PI *\*dzanH-* ‘to be born, to grow’, YAv *zan-*, 3PL *zīzanānti* < PII *\*žanH-* < PIE *\*g’enh<sub>1</sub>-* ‘to be born’ (Cheung 2007: 464 – 465, s.v. *\*zanH-*; Werba 1997: 288–289 No. 286 *jan<sup>i</sup>*; NIL: 139–154, s.v. *\*gēnh<sub>1</sub>*; LIV<sup>2</sup>: 163, *\*gēnh<sub>1</sub>*)

(Koivulehto 1999a: 222–223; 2001b: 254–255; 2005: 330 = 2016: 224–225, 296–297, 332; Häkkinen 2004 s.v. *syntyä*; EES s.v. *sündima*)

Koivulehto’s etymology is semantically plausible, but both the vowel and consonant substitutions require more detailed commentary. Koivulehto assumes a sound change *\*e* > *\*ü* (= Fi *y*) in Finnic. There are some examples of PU/Pre-FI *\*e* developing into Finnic *\*ü* (such as *\*jewä* > *\*jüwä* ‘grain’, also an Indo-Iranian or other Indo-European loanword, see the entry), but this change is usually caused by phonological factors which are absent from this word. The example *\*jewä* > *\*jüwä*, *\*jüwä* has been explained as a regular development

by Pystynen (2015), as *\*je* becomes *\*ji* in Pre-Finnic, and *\*i* is labialized before *\*w*. Pystynen further notes that *\*e* is usually not labialized even before *w*. However, it seems that a sound change *\*e > \*ü* in Finnic can nevertheless be assumed, as Aikio has recently argued in favour of such a change (see the entry *\*seči-*).

The Komi word cannot be a regular cognate of the Finnic word (Komi *o* regularly reflects PU *\*e* in *\*ä*-stems, but it cannot have developed from earlier *\*ü*). One possibility to explain the irregular relationship would be to consider the Komi word a loan from Finnic, but this is improbable because the borrowing would have to be extremely old, from the time before the Proto-Permic denasalization (Niklas Metsäranta: personal communication). The Saami word *šaddat* ‘to grow, to be born’ is a well-known loan from Finnic (Sammallahti 1998: 264).

The idea that the Finnic or the Komi word is borrowed from Iranian *\*žanH-* involves a substitution *\*s ← \*dz*, for which there are not many examples. Finnic *s* could also reflect Pre-Finnic *\*š*, which could also point to an earlier Indo-Iranian donor form. The word might then be from Pre-Indo-Iranian *\*ženH-o-*.

? Indo-Iranian etymology

## Md E *sire*, M *sirä*, *širä* ‘old, aged’

< West Uralic/Pre-Md *\*serä*

← “Pre-I” *\*dzerH-o-* (= PI *\*dzarH-a-*), root *\*dzerH-/\*dzarH-* [Cheung 2007 *\*zarH-*], > Oss *zæronð* ‘old’; < PII *\*žara-* > OI *jaran-* ‘alt, gebrechlich’, a participle/verbal adjective with amphidynamic inflexion; cf. also OI *juránt-/jurát-* < PIE (transp.) *\*gr̥h₂-ént-/\*gr̥h₂-nt-* < PIE root *\*ger-*, cf. Greek *γῆρω* ‘to become old’, *γῆρων* ‘old’, Arm *cer id.* (EWAia s.v. *jaran-*; Bielmeier 1977: 156; LIV² s.v. *\*ger*; Werba 1997: 289–290 No. 288; Gotō 2013: 46; Steer 2015: 165ff.)

(Joki 1973: 314, No. 141; Katz 1985: 143, footnote 147; Koivulehto 1999c, 2001b; 2016)

The etymology is convincing: the semantic correspondence is one-to-one, and the Mordvin word bears a close resemblance to the Indo-European word family denoting ‘old’. Although there is at least one word for ‘old’ that can be reconstructed for Proto-Uralic, *\*ammi-*, many words denoting ‘old’ have been borrowed from Uralic into Indo-European (see *\*sojka*). This is one of the few Iranian borrowings reflected only by Mordvin, and a rare example of a clearly Iranian loanword (words such as *\*pakas* are harder to date). Although the substitution PI *\*dz > U \*s* is a bit peculiar (one could argue that the affricate *\*č* would be more logical), it is clear that the PII palatal affricate cannot come into question. Koivulehto has argued that the word is a “Pre-Iranian” borrowing: the consonantism clearly points to Iranian, but the Pre-Mordvin vowel *\*e* has to be a substitute of the retained PIE *\*e*. Koivulehto notes that although later Iranian languages show a front-vowel development from PI *\*a*, in the later loans the prevailing substitution in Uralic is *\*ä*, not *\*e*, and the Mordvin words clearly reflect an *\*e-ä* stem.

However, it is also mentioned by Koivulehto that the Iranian word appears also in the zero grade (with syllabic *\*r*). As the “Pre-Iranian” borrowing is an anachronistic idea (it is

difficult to suppose that the merger of non-high vowels happened after the Iranian development  $*č > *ts$ ,  $*č > *dz$ ) and is supported by very few parallel examples which often involve other dubious developments, a borrowing from a zero-grade form  $*dzr̥Ha-$ , which is attested *inter alia* in Indic *jurant-*, looks more plausible. The substitution of  $*r̥$  by the sequence  $*er$  is also found in a number of borrowings, such as  $*mertä > Mo\ mir'd'e$ , Ko *mort*. Most of these borrowings are difficult to date, but if the  $*er$  loans belong to the same layer, this Iranian loan  $*serä$  proves that they are relatively late loanwords from Iranian.

Here one has to note also that if the word does not have to be a very old borrowing from “Pre-Iranian”, the initial sibilant can also be re-evaluated. Although there is no compelling reason to reject Koivulehto’s idea that  $*s$  is a substitute for  $*dz$ , it could equally well substitute later “Old Iranian”  $*z$ ; this was already the view of Joki; the steppe languages which were in contact with Uralic eventually developed this sound, as shown by the Ossetic example, and as we do not know the exact time of borrowing, it is possible that the Mordvin word was borrowed from an “Old Iranian”  $*zara-$ . This reduces the number of good examples of Proto-Iranian developments in the loanword material. It seems that the substitutions of  $*dz$  and  $*z$  are very difficult to distinguish. The words for ‘gold’ in Mordvin reflect the same problem, and probably belong to the same layer of borrowings (see ***\*sernä*** below).

The  $*e-ä$  vocalism is found also in another Iranian loan into Mordvin, namely  $*sernä$  ‘gold’ ( $> si'ñe$ ). In both forms the  $er$  seems to substitute for Iranian  $*r̥ \sim *r̥H$  or its later reflex  $*ər$ . Katz (1985: 143, footnote 147) is doubtful of the etymology because of problems with the word-initial consonant, the vocalism and word-formation. It is difficult to understand what Katz means with the problems with word-formation (“Bildung”). For the vocalism his cautiousness is understandable, as the  $*e$  vocalism does not often appear in loans where Iranian has  $*a$  or  $*r̥$ , but as there are at least two cases in Mordvin that show this vocalism, we have no reason to doubt the explanation. The initial  $s-$  as the substitution of  $*z-$  is expected, as mentioned above.

Although Koivulehto does not mention it, the Hungarian adjective *öreg* ‘old’ has been considered a cognate of the Mordvin word (UEW s.v.  $*serä$ ). Technically the Hungarian word could reflect a PU form  $*serä-$  with a derivational suffix: the development of  $*e$  to  $*ö$  is late in Hungarian, but usually there are dialectal reflexes of the retained  $*e$ . The Hungarian suffix is obscure – according to the UEW it is “ein seltenes Nominalsuffix”. It is also uncertain whether Hungarian and Mordvin could reflect the same Iranian loan, as most of the Iranian borrowings in all probability stem from the period after the split of Proto-Uralic. Hungarian could have borrowed the word separately, but not from a very late Iranian language, as the development  $*s > *ø$  had occurred before the Alanic-type loans entered into Hungarian. However, the Hungarian word also has a Turkic etymology: the UEW mentions that it can be borrowed from Turkic  $*irik$  akin to Old Chuvash  $*irik$ , Osm., Kyrgyz *iri* ‘big, thick’, Chagatay *irik*, *ärik* ‘old, big’. It is thus more probable that the Mordvin and Hungarian words are unrelated and only the Mordvin word is borrowed from Iranian.

## An Indo-Iranian etymology

**Fi *sitoa* ‘to bind’; Saami *čadnat* id.**

< PU *\*sita-* ‘to bind’

← PII *\*sHay-* or PIE *\*seh<sub>2</sub>i-* ~ *\*sh<sub>2</sub>ey-*

(Koivulehto 1991: 51, footnote 23)

The etymology is semantically and phonologically plausible, and there is no reason to doubt the Indo-European origin of the Uralic word. It is uncertain whether *-ta-* is a Uralic suffix, as it cannot be explained from the Indo-European verbal forms. Nevertheless, it is not certain that the word was borrowed from Indo-Iranian, as the root is widespread in Indo-European and another source is also possible. Koivulehto has also considered this to be Proto-Indo-European loan. There are no grounds for deciding this based on semantic or formal criteria, so the Indo-Iranian origin of the Uralic word remains open.

? Indo-Iranian etymology

**Fi *sini* : *sine-*, *sininen* (der.) ‘blue’, Est *sini*, *sinine* id. (+ cognates in Karelian, Veps, Ludic, Votic and Livonian); Mo E *šeń*, *šǎń*, M *šeńem* id.**

< WU *\*sini*

← ? PI *\*axšaina-* ‘nicht hell’ > Av *axšaēna-*, OP *ahšaina-*, Afg *šīn* ‘blue; green’, NP *χasīn* ‘bluish’, Oss *æχsin* ‘dark grey’

(Mikkola 1894: 162; Joki 1973: 314; UEW s.v. *\*sine*, No. 1555; SSA s.v. *sininen*)

This is an old etymology, but it has not been commented on in detail in recent etymological works. SSA mentions the Iranian etymology with a question mark but without further comment. This is one of the few colour names that can be reconstructed at any Uralic proto-language level. The Iranian etymology is probably not correct, which should be obvious to anyone who takes a closer look at this etymology.

The Iranian word is a transparent formation with the negative prefix *\*a-*: the word originally meant ‘non-white’, and thus it came to mean a dark colour, blue. If the word had been borrowed into (West-)Uralic from Proto-Iranian or Old Iranian, or from “Alanic”, there would have been no reason to drop the initial *a-*. This kind of development would go against all we know about Uralic substitution patterns. Thanks to Greek attestations of Scythian names (listed also by Joki), we know that the languages of the East European steppe still retained the initial syllable *\*ax-*, and this is obvious also from the modern Ossetic form. The forms which have simplified the initial syllable, such as Afghan *šīn*, are late, and the Uralic word can hardly be borrowed from such a language due to chronological and geographical reasons. A prefixless Proto-Iranian or Old Iranian form *\*χšaina-* is out of the question, as semantically it would make no sense to borrow a word for ‘blue’ from a word meaning ‘white’. It is also important that change from *š* to *s* would be required for the origin to be suitable, as borrowing from a form *šīn* or the like would result in Uralic *\*\*šini*.

The relationship of the Uralic words to Russian *синий* ‘blue’ is unclear. The UEW refutes the etymological connection to that Russian word, although it states that both the Finnic and Mordvin words might have been “influenced” by the Russian word. Vasmer (s.v. *синий*) mentions the Finnic and Mordvin words in connection with the etymology of the Russian word but points out that Iranian etymology has been suggested by Mikkola. The Russian word is etymologically unconnected to the Iranian word, but it might belong to another Indo-European word family; according to Vasmer, the Russian word is either related to *syát* ‘glänzen’ or to *sivuj* ‘grau’, which is cognate with the root attested in Baltic *\*šīwas* ‘grey’ and Indo-Iranian *\*čyāwá-* ‘dark brown’ (EWAia II s.v. *šyāvá-*), but the Uralic word cannot be borrowed from these languages, as the nasal formation is attested in Slavic only. It is possible that the word is a very early Slavic loan, but this idea is complicated by the fact that there are no convincing parallels of Slavic borrowings this early (into West Uralic). There are convincing examples of earlier Balto-Slavic loans, but it is unclear whether a pre-form for *синий* can be reconstructed for this stage in the view of the Baltic data. However, Derksen (2015) does not mention the Slavic words in connection with the Baltic ones (s.v. *šemas*, s.v. *šyvas*), so the Slavic word might be etymologically unrelated. Černyh (1993, s.v. *синий*) simply states that the Russian word is derived from the root *\*si-*, and that both Slavic *\*sijati* and *\*sivŭ* are derived from the same root. The *Etimologični slovník ukrajnskoj movi* (s.v. *sini*) also states the relation to Baltic *šyvas* and other words for ‘grey’.

Not an Indo-Iranian etymology

TN *sēr*” ‘white’, FN *xjèr*, E *si*, Nga *syr*, Slk (der.) *seṛi*, Kam (der.) *siri*, Mat (der.) *сыры* id.

< PS *\*ser* ~ *\*seṛ* ‘white’ (SW: 138)

← “frühurarisches” *\*kʷitróm* ? (= PII *\*čwitra-* ‘white’, root *\*čwait-* ‘to be white’) > OI *švitrá-* ‘white’ (EWAia II: s.v. *švitrá-*; Werba 1997: 248 No. 206 s.v. *švit-*)

(Katz 2003: 86)

This etymology is not very convincing. First of all, there are absolutely no reasons to suppose that this word to be borrowed from the “early Proto-Indo-Iranian” form postulated by Katz, as the form itself is anachronistic, being practically identical to Proto-Indo-European. If this word was an Indo-Iranian borrowing, it could be simply borrowed from a Proto-Indo-Iranian word with *\*č* substituted by PU *\*ć*. Katz also argues that the Samoyed word was borrowed into Proto-Uralic, but as no cognates are found in any other branch, this is very unlikely. Katz gives the PU form as *\*ćīrə*, following his unconventional reconstruction of the proto-language.

If this etymology was correct, the borrowing should be very old, Pre-Samoyedic, because of the development of the affricate *\*ć*, and therefore this would probably belong to the same layer of loanwords as PS *\*wātā* (see the entry).

The Samoyed word is possibly related etymologically to PS *\*sirā* ‘snow’ (> TN *sira*) and *seṛ* ‘ice’ (> TN *sēr*) (SW: 138, 140), as is suggested by Aikio (2014: 75), but the more

precise relationship between these words would require further research. According to Aikio, the Proto-Samoyedic word is also borrowed into Proto-Yukaghir as \**sēr* (> Yukaghir (Kolyma) *sērīl* ‘snow on trees’, (Tundra) *sierul* ‘hail’).

Not an Indo-Iranian etymology

**Hu *arany* ‘gold’; Ms South *tarəń*, West *tarəń*, ‘Kupfer’ (< PMs *tarəń*); Kh East (Vakh dialect) *lorńz* ‘Kupfer (Samowar)’, (Vasjungan dialect) *jorńi* ‘Messing, Kupfer’ (< PKh ? \**lorńV* or \**larńV*)**

< PUg/PU \**sīr(a)ńa* ‘gold, metal’ (Sammallahti 1988: 550; UEW s.v. \* *šarańa*, *sarańa*)

← I \**zr̥Hanya-* (?\**zərańya-*) or \**zarańya-* or

← PI \**dzr̥Hnya-* (?\**dzərańya-*) > Av *zarańya-* ‘gold’, Manichean Sogdian *zyrn*, Khotanese *ysirra-*, Oss *zærīn/zærīnæ* id.; cf. OI *hīrańya-* ‘gold’; < PIE \**ǵʰh₃-en-* (EWAia II s.v. *hīrańya-*)

(FUV: 129; Korenchy 1972: 77, No. 35; Joki 1973: 250; Katz 1985: 323–324; 2003: 258; Ligeti 1986: 134; UEW s.v. *šarańa*, *sarańa*; Lushnikova 1990: 292; Helinski 1991: 222; Abondolo 1996: 40; J. Häkkinen 2009: 23–24; Parpola 2012: 161; Blažek 2017: 270)

This old etymology is also mentioned by EWAia. As noted by Korenchy (1972), Ligeti (1986), the UEW, as well as Abondolo (1996), the words in Mari, Mordvin and Permic with similar meaning are probably parallel borrowings from various Iranian languages at different times (however, see below). The word was acquired before the Proto-Ugric sound changes, so it has to be a very old borrowing. According to a traditional view the word is borrowed from an (Old) Iranian form \**zarańya-*. Lushnikova (1990: 292, No. 202) and Helinski (1991: 222) consider the words as specifically “Scytho-Sarmatian” loans.

An earlier borrowing from Proto-Iranian could perhaps come into question too: in the Western Uralic loans from Iranian the substitution \**s* ← PI \**dz* has been suggested to prevail in word-initial position (according to Koivulehto 1999c, 2001f). However, examples are few and in the case of this Ugric word it is impossible to prove that the \**s* here reflects Iranian \**dz* and not \**s*. It is most probable that this word is an Old Iranian-type loan as has usually been assumed. A Proto-Indo-Iranian loan is out of question because of its palatal affricate (this would have given Uralic \**čirańa*).

Korenchy (1972) reconstructs the word as \**sarańa*, but in modern reconstruction the Ugric vowels point to Uralic \**ī*. Contrary to some other borrowings, here the relations of Hungarian, Mansi and Khanty regularly point to a unitary form, and this word can theoretically be reconstructed for the Ugric proto-language. Abondolo (1996: 40) notes that the Hungarian word with *a–a* vocalism “neither contradicts nor confirms a Ugric unity at the time of borrowing”. Although this seems a bit enigmatic, Abondolo seems to imply that a Hungarian \**a–a*-stem can point to either a Uralic/Ugric \**a–a* or \**ī–a*-stem. The Proto-Ugric meaning is reconstructed as ‘yellow/red metal’ by Abondolo.

This word’s relationship to the possible Mari and Mordvin cognates has raised much discussion. They are nowadays usually considered parallel borrowings from the same

Iranian source, contrary to what is often claimed in the earlier sources, such as FUV. However, it has been again claimed by J. Häkkinen (2009) recently, that the words in Mari, Mordvin and Ugric could all reflect Proto-Uralic *\*sṛ(a)ńa*. Häkkinen considers this an important notion, as he assumes that the word is an Old Iranian (= post-Proto-Iranian) borrowing into Proto-Uralic. This would have important implications for the dating of the Iranian-Uralic contacts.

The Mari word *šörtńa* (TschWb: 717–718) could reflect an earlier *\*ṛ* so it could theoretically be derived from the same reconstructed form as the Ugric words. Mari *ö* in front of *\*r* often reflects *\*ṛ*. In both old *\*i-a* and *\*i-i*-stems Mari shows the stem vowels E -*o* ~ W -*ə* or E -*o* ~ W -*ə*, so the stem vowel is no obstacle either (the -*i* in some western dialectal forms has to be secondary).

The Mordvin word *sirńe*, on the other hand, cannot reflect earlier *\*ṛ*. Although Häkkinen is right in assuming that it is difficult to explain the Mordvin word as a later loanword from the same Iranian source, it cannot be regularly connected to the other words that are mentioned here.

Häkkinen compares the development of *sirńe* with some other Uralic words. The reflexes of Uralic *\*wäškä* (> Mo E *viškä*, *uške*) do not serve as a good parallel because this word cannot be regularly reconstructed for Proto-Uralic according to Aikio (2015b), and the same can be said of the reflexes of *\*wašara* (> Mo E *višir*), which is a loanword that can have been borrowed independently into Mordvin, Finnic and Saami. The last example that Häkkinen mentions is the relationship between Mordvin *mirďe* and Finnic *marras*, but these reflect two different borrowings from Indo-Iranian (see the entries).

Interestingly, the word *sirä* ‘old’ in Mordvin shows similar vocalism as the word for gold: this points clearly to Pre-Mordvin *\*e*. Both the words for ‘gold’ and ‘old’ can be explained from forms which had *\*r* in Iranian, and Pre-Mordvin *\*er* could have been the substitute of this in both cases.

Rédei (1986c: 82) has derived the Mari and Mordvin words from a “Middle Iranian source”. Although Rédei rightly considers them parallel borrowings and not regular cognates, it is extremely unlikely that the words were borrowed from a “Middle Iranian” language but probably from some much earlier form. Rédei reconstructs the Mari word as *\*serńä* but this is certainly incorrect.

Udmurt *zarńi* and Komi *zarńi* are obviously much later borrowings from Iranian because of both the vocalism (Permic *a* corresponding to Iranian *a*) and the sibilant *z* which is not found in early loans.

Within Ugric, the word is attested only in part of the Khanty and Mansi languages, but this has to be due to later disappearance of the word from certain languages, as the attested forms with regular correspondences and the distribution of the reflexes in all the three Ugric language groups clearly show that the word was present in Proto-Khanty and Proto-Mansi.

An Indo-Iranian etymology

**Md E *sírne*, M *sírñä* ‘gold’; Ma *šürtñö***

**< \*? PU *serñä* ‘gold’**

← PI \**dzr̥Hanya-* or later I \**zaranya-*, cf. Manichean Sogdian *zyrn*, Khotanese *ysirra-*, Oss *zærīn/zærīnæ* id.; cf. OI *hírunya-* ‘gold’; < PIE \**g<sup>h</sup>lh<sub>3</sub>-en-* (EWAia II s.v. *hírunya-*)

(FUV: 129; Korenchy 1972; Joki 1973: 250; Helimski 2000: 193; J. Häkkinen 2009: 23–24)

See also \**sír(a)ña*

It is probable that the Ugric words for ‘gold’ are parallel loans and not true cognates of the Volgaic words dealt with here (see above for the Ugric words). Even though Mari *šürtñö* and the Mordvin forms have often been connected, even they cannot be derived from a common source, which is not surprising, as there do not seem to be any Indo-Iranian loans shared by only these two branches, which is also consistent with the other ideas concerning the relationship of Mari and Mordvin within the Uralic family (= the absence of Proto-Volgaic). The Mari word could formally reflect a similar pre-form as Pre-Ugric \**sír(a)ña*. J. Häkkinen argues that the Ugric, Mari and Mordvin words all reflect a Proto-Uralic word, and states that on the basis of this there were contacts between Proto-Iranian and Proto-Uralic, but the Mordvin word cannot be derived from such a pre-form.

The Mordvin vowel *i* can in no case reflect PU/Pre-Mordvin \**a* or \**i*. Although J. Häkkinen tries to argue that there are cases where inherited Uralic cognates display a correspondence Fi *a* ~ Mo \**i* (which practically means “PU \**a* > Mo *i*”, although this is not explicitly stated by Häkkinen), the examples that Häkkinen presents to support this are clearly false, i.e. \**wačara* ‘hammer’ is irregularly reflected also in Saami, and \**wäškä* ~ \**wäski* is a notoriously irregular item and potential Wanderwort which cannot be used to support any regular sound law (see Aikio 2015 for the most recent and comprehensive analysis of \**wäškä* as a Wanderwort).

The Mordvin \**i* could reflect earlier \**e*, and here one could think that the word was borrowed into Pre-Mordvin in the form \**serñä*, where \**er* reflects Iranian \**r̥H* or \**ər*. This would mean a similar substitution as in the word \**serä* ‘old’, which is attested only in Mordvin (see the entry).

For some reason, Häkkinen mentions the Permic words (Ud, Ko *zarñi* < PP \**zarñi*) in connection with the Mari, Mordvin and Ugric words, and considers the Permic sibilant *z*-inexplicably irregular (“selittämättömän epäsäännöllinen”). It is obvious that the Permic word was borrowed from some later Iranian form, cf. Sogdian *zyrn*, Ossetic (*syy-*)*zærīn*, (*syy-*)*zærīna*.

According to Rédei (1986c: 82) the Mari and Mordvin words are Middle Iranian loanwords (this view is echoed also by Bereczki 1992: 94). There is no reason to consider that these clearly early loanwords were borrowed from Middle Iranian forms.

An Indo-Iranian etymology

**SaN *soamis* ‘some’ (also in S, U, L and I < PSa *\*soamēs*)**

**< Pre-Sa *\*soma* or *\*sami***

← PII *\*samH-* > OI *samá-* ‘irgendein’ < PIE *\*sm̥h₂o-* (Beekes 1988: 99; EWAia II: 703, s.v. *samá-*)

(Sammallahti 1999: 80–81; 2001: 402, 408; Holopainen 2018b: 165–166)

As stated by Holopainen (2018b), Sammallahti’s recent etymology is convincing, and he is right in arguing that a Proto-Scandinavian origin for the Saami word is impossible because of the first-syllable vowel. Although the word has cognates in most branches of Indo-European, the Indo-Iranian original seems to be phonologically the most suitable original for the Saami word. Theoretically a Proto-Indo-European or Pre-Indo-Iranian source would also be possible, as the Pre-Saami vowel + nasal could also substitute for the syllabic resonant *\*m̥*.

Following the new vowel reconstruction by Aikio (2015b), the Saami word could reflect either an old *\*a-i* or *\*o-a*-stem.

? Indo-Iranian etymology

**Mari *šonyo*, *šongə* ‘old’, Bolshoy Kilmez *sonyo* (< PMa *\*songə*); Hu *agg id.***

**< *\*sonka* or *\*san̥ka* or *\*s̥y̥ka* ‘old’ (UEW s.v. *sonkə* [*son̥kə*]; Sammallahti 1988: 548)**

← PII *\*sana-ka-* > OI (Vedic) *saná-* ‘old’, *sanaká-* ‘former, ancient (alt, ehemalig, früher)’; an IE word, compare Latin *senex*, *senis*, Welsh *hen id.*, Greek *ἔνος* ‘old, from previous year’ < PIE *\*sén-o-*, root *\*sen-* (EWAia II s.v. *sána-*; NIL s.v. *\*sen-*; Kölligan 2018: 2257)

(Koivulehto 2001b: 254 = 2016: 296)

Koivulehto’s etymology is convincing: the semantics correspond perfectly, and the substitution of *\*s* by *\*s̥* is expected, as well as the vowel substitution PU *\*o* ← PII *\*a* (here it has to be noted that the PIE vowel was *\*e*, see NIL s.v. *\*sen-*, and thus there is no possibility to consider this a borrowing from Proto-Indo-European or some other branch of the language family). The Uralic *\*a*-stem is also expected from the PII *\*a*-stem. The Mari and Hungarian words can be regularly derived from a common Proto-Uralic form and there is no need to assume parallel borrowings (the parallel borrowings would have to be very early, as Hungarian shows *\*s* > *ø*). The Hungarian word has at times been also connected to the word family *ó* ‘old, ancient’ (UEW s.v. *sonkə*; EWUng s.v. *agg*), but these two words cannot be connected via regular means of word formation within Hungarian. The word *ó* is probably from another Proto-Uralic stem, *\*ammi-* ‘old’ (Aikio 2015b: 51), although the development of geminate *\*mm* > Hu *ø* lacks parallels (the regular development of intervocalic *\*m* is Hu *v*, and here we have to assume that the geminate has developed similarly; *v* is still visible in derivatives like *avul-* ‘to become obsolete’). There is

no alternative loan etymology for the Mari word *\*soŋka* or *\*saŋka*. Bereczki 2013 (s.v. *šongo*) does not mention the Indo-Iranian etymology.

While some might claim that there is no reason to suppose that such a central concept would be borrowed, Koivulehto notes that there are many other adjectives meaning ‘old’ in the Uralic languages which are borrowed from Indo-European, such as Mordvin *\*serä* (from Iranian, see the entry above), and Finnic *vanha* from Germanic *\*wanha-*.

The exact dating of the borrowing is difficult. If it is true that the Iranian *\*s > h* sound change is a late, Post-Proto-Iranian development, the original *\*sana-ka-* could be either Proto-Indo-Iranian or Proto-Iranian. In a footnote Koivulehto has tentatively suggested that Finnic *honka* might also be a reflex of this same Uralic word, but then PU *\*š* would have to be reconstructed (as is well-known, the reflexes of Uralic *\*s* and *\*š* merge in Hungarian and most dialects of Mari, but see below). Koivulehto considers a substitution *\*š ← \*s* possible, because PII *\*š* did not appear in word-initial position, but I find this unlikely, as the two sibilants were different phonemes in the donor language, and thus the situation is more complicated than with the Proto-Germanic sibilant *\*s*, which is usually substituted by *\*š* in early borrowings (LÄGLOS I: XVI). There are no convincing parallels of such a substitution among the Proto-Indo-Iranian loanwords. Moreover, the Bolshoy Kilmez (Malmyzh) dialect of Mari reveals that the initial consonant of this Uralic word has to be reconstructed as *\*s-*, not *\*š-*, as the two proto-phonemes have not merged in this dialect of Mari.<sup>40</sup> The dialectal form with *s-* is also cited by the UEW.

If the Finnic word does not belong here, the Uralic word can be reconstructed also as *\*saŋka* (Mari *o* and Hungarian *a* are ambivalent, as both could reflect either PU *\*a* or *\*o*), and even *\*siŋka* would be possible, as *\*i* appears as the substitute of PII *\*a* in numerous examples. Sammallahti (1988) reconstructs the word as *\*sonka*. In Hungarian *\*i* merged with *\*a* in *\*a*-stems, but in Mari the development was more complicated and is poorly known. Examples (from Aikio2015b) of *\*i > \*o* include *\*d’imi > MariE lombō, W lombâ, \*ippi > MariW owâ, \*iikti > MariNw toktâ-*, as well as *\*miksa > mokš* and *\*piŋka > poŋgo* ‘mushroom’. Only the last two examples are *\*a*-stems, and here probably the labial consonants *m* and *p* have influenced the development of *\*i* (see the entry *\*piŋka* for more comments on the phonology of this word). Zhivlov (2014: 125) has argued that in both *\*i*- and *\*a*-stems *\*i* became *\*o* before *\*k* and *\*ŋ*, but this assumption is clearly based on limited evidence, and as we can see from the examples listed above, the development was more complicated than this. Also Zhivlov admits that there are unexpected reflexes of *\*i* in Mari. At the moment it is obvious that in most cases *\*i-a*-stems are thus not reflected by Mari *\*o*, and *šoŋgo* rather reflects *\*saŋka* or *\*soŋka*, as *šoŋgo* could completely regularly reflect either of the forms.

The Indo-Iranian etymology has a convincing Indo-European etymology. The root *\*sen-* ‘old’ is reflected in numerous branches of the family. However, there are details in the formal background of the noun *\*sanaka-* that require comment. The Indo-Iranian word and its cognates, Germanic *\*seniga-* and especially Latin *senex*, are possible examples of the so-called laryngeal hardening, meaning that the IE laryngeal could have “hardened” into a velar stop before *\*s* already in Proto-Indo-European. This idea of laryngeal hardening was first suggested by Martinet (1955), and it has met with various views ever since (see Lindeman 1987: 94–98 and Schrijver 1991: 148–154 for sceptical discussions of

---

<sup>40</sup> I am grateful to Niklas Metsäranta for pointing this out to me.

this disputed phenomenon and Olsen 2009 for a recent defence of the laryngeal hardening). The Latin cognate *senex* might thus result from *\*seneh<sub>2</sub>-s*, and if the laryngeal can be reconstructed for the PIE word that is reflected by both Latin *senex* and Vedic *sanaká-*, it was probably still present in the Indo-Iranian word when it was borrowed into Uralic. Uralic *-k* could thus be a substitute for a laryngeal.

However, the origin of the Latin *-k* suffix in this word is a disputed question, and it is not at all certain that the Indo-Iranian and Latin words reflect the same suffix. The Latin *-k* has been explained as an old *\*-k* suffix (Lindeman 1987: 96; Cowgill 1965: 176; cf. Pinault 2001: 102–104), but this view has also been criticized (Watkins 1965: 182, 186). It is interesting that also the Germanic cognate *\*senīga-* ‘senior’ (> Goth *sineigs*) contains a similar-looking suffix: Kroonen (2013: 433, s.v. *\*senīga-*) states – a bit obscurely – that the suffix *-īga* is “not identical to, but nevertheless close to that of Lat. *senex*”). Olsen (2009: 207–208, 217) has defended the idea that the *k* in the Latin suffix could reflect a laryngeal. Borrowing from a form *\*sana-ka* and syncopation could lead to *\*soŋka*, as Koivulehto himself noted, and so could a borrowing from *\*sanaH-*. In Martinet’s original idea the Germanic, Indo-Iranian and Latin suffix all reflected masculine stems ending in *\*-h<sub>2</sub>* (cf. Schrijver 1991: 149).

The Latin suffix is reconstructed as *\*Vk* by NIL, and this source notes that this may be identical with the Indic suffix. De Vaan (2008, s.v. *senex*) only mentions that the Latin nominative has a suffix *\*-ek*, but he offers no further comments on the origin of this suffix. As it is thus highly uncertain whether the *-k-* of the Indo-Iranian results from the laryngeal hardening or not, the Uralic word cannot be used as external evidence for laryngeal hardening or the reconstruction of laryngeal for this root.

An Indo-Iranian etymology

### **Hu *úr* (: *ura-*) ‘lord, gentleman’**

< ? Pre-HU *\*suri* ~ *\*sura*

← PII *\*Hsu-Hrī-* > OI (Vedic) *sūrī-* ‘Opferherr, Herr, Schirmherr’ (EWAia II s.v. *sūrī-*)

(Katz 2003: 169–170; Holopainen 2016; Aikio 2018)

This Hungarian word has a complicated Indo-Iranian research history. While most of the modern-day etymological dictionaries (TESz; EWUng s.v. *úr*; WOT: 969–974) of Hungarian offer cautious support for the Turkic etymology for this word (← Proto-Turkic *\*ur-* ‘megtermékenyít, nemz’ (> Turkish *urī* ‘young man’), various Indo-Iranian sources have been put forth for this word in earlier research. Aikio has recently reconsidered the old possibility (originally suggested by Munkácsi 1901: 614–615) to derive the Hungarian word from PU *\*asora*, another Indo-Iranian loan. This remains a possibility, although the vowel development is unclear. The Hungarian word has also been connected to Avestan *sura-* ‘stark, gewaltig; gebietend Herr über...’ ~ OI *śura-* (Harmatta 1955). Moór (1965) rejects Harmatta’s etymology and supports the old etymology of Munkácsi, arguing that it is chronologically impossible for the Iranian *\*s* (< PII *\*é*) to be reflected by Hungarian *ø*.

This argument is obsolete, as there are several other loans borrowed into the Ugric languages from the Old Iranian stage with the sibilants \*s, \*z, such as Hu *arany* ‘gold’ (see the entry \**sīrña*). This means that Harmatta’s etymology remains a possibility too.

The old idea that the Hungarian word is cognate to Finnic *uros* ‘male; (arch.) man’ (Nirvi 1952: 29–31), and Saami L *årēs* ‘male, reindeer bull’, N *varris* (SSA s.v. *uros*; UEW s.v.) can be rejected, as the Finnic and Saami words have a convincing Germanic etymology, from PG \**ūruz* (> Old Norse *úrr* ‘Auerochse’, Old English *úr*, Middle Low German *ur-osse* id.; LÄGLOS s.v. *uros*). Both the Finnic and Saami words probably reflect earlier \**uras*, with irregular *va-* in some Saami varieties, and *-as* > *-os* in Finnic. Honti (2017: 68–72) is still supportive of the old Uralic etymology, assuming that the Hungarian word can represent a merger of a Turkic loan and the inherited Uralic word. As all the Uralic words have convincing loan etymologies, there is no reason to support the Uralic etymology, however.

Also other words for ‘man’ were borrowed from Indo-Iranian into Uralic, such as \**wirä* (> Ko *ver*), possibly Ma *mari* and Ugric \**mVńíci* (see the entries). Katz (2003: 17) attempted to derive the Finnic word *uros* and its Saami cognate from the same Indo-Iranian source as Ko *ver*, reconstructing the Finno-Saamic form as \**uīrəs*, but this explanation is phonologically unconvincing, as is argued by Holopainen (2016: 60). Holopainen (2018b: 170–172) assumed that South Saami and Lule Saami *nïere* ‘man; person’ could possibly be borrowed from Indo-Iranian, but it is also possible that the word was borrowed from some other branch. In Germanic the word is not retained in the meaning of ‘man’, although cognates of this IE root are possibly attested in Germanic names such as ON *Njorðr* ‘name of a god’ (Zimmer 1999: 131–132). There is thus no compelling obstacle to assume that the word can be an early Germanic loan.

An Indo-Iranian etymology

### Ud, Ko *sur* ‘beer’

< PP *sur*

← PII \**sūrā-* > OI *sūrā-* ‘beer’, Av *hurā-* ‘fermented mare’s milk’ (EWAia II s.v. *sūrā-*)

(KESK: 266; Joki 1973: 317; Rédei 1986c; Lushnikova 1990: 262; Helimski 1997b: 121; Csúcs 2005: 376; Zhivlov 2013: 219–220; Parpola 2015: 66)

This is a problematic etymology. Although the similarity of the Permic and Indo-Iranian words is transparent, the Permic vowels do not point to an old word, whereas the *s* means that the borrowing cannot have taken place recently. The situation is complicated by the fact that the Indo-Iranian word was borrowed early into Turkic languages as well, and from there it has been borrowed by various Uralic languages, such as Hungarian (*ser*, *sör*). It cannot be ruled out that the Permic word is not a borrowing from Turkic too.

According to Helimski and Zhivlov, the Permic word was acquired from the Andronovo Aryan language, an unattested branch of Indo-Iranian where \**s* would have been retained longer than in Iranian languages. There are some additional examples of a preserved \**s*, but these etymologies are more or less problematic and at least most of them can be explained otherwise (see *śumīs*). Note that the “Middle Iranian” forms (Csúcs: \**sur*) with \**s-* that are listed by Rédei (1986c) and Csúcs (2005: 376) are impossible, as Indo-

European \*s obviously did not survive in any Middle Iranian language. Parpola (2015: 66) mentions this word as a Proto-Indo-Iranian loan. From the Indo-Iranian point of view this would be possible, but because of the Permic vocalism it is unlikely that the loan could be so early.

The word looks like a Wanderwort, and due to its large distribution in various languages in the Volga-Kama area and beyond, it seems risky to draw historical conclusions based on it. The ultimate origin of the word is certainly in Indo-Iranian, but the exact route to Permic could be via Turkic.

? Indo-Iranian etymology

**Hungarian *hét* (: *hete-*); Mansi South *sāt*, East *sə̄t̄*, West *sāt*, North *sāt* (< PM \**sātə*); Khanty East *läwət*, North *apət*, South *tápət* (< PKh \**apət*) id.** (Zhivlov 2006: 143)

< PUG ? \**säptä*

← PII \**sapta-*/ PI \**safta-* > YAv *hapta-* ‘seven’, OAv *haptaθa-* ‘seventh’, OI *saptá-* ‘seven’ < PIE \**septm̥-* (EWAia II s.v. *saptá-*)

(FUV: 130; Korenchy 1972: 70, No. 27; Joki 1973; Ligeti 1986: 134–135; UEW s.v. *θäptz*; Helimski 1997b: 121)

This much-discussed word is most probably a parallel loan in all of the Ugric languages, as it is impossible to reconstruct a common proto-form to the Ugric words because of the irregular reflexes of the sibilant. Only Khanty reflects Ugric \**säptä* regularly. The sibilant \*s should disappear in Hungarian completely; it is often assumed that this has indeed happened, and that the *h-* in *hét* is the result of analogy on the word *hat* ‘six’. The attempt of Honti (1997: 35–37) to explain the Mansi sibilant with a flip-flop change where PUG/PM \**θ* returns (!) to \*s in front-vocalic words which have *j* word-medially is not convincing due to lack of examples and general implausibility (there is no trace of *j* in this Mansi word), but rather it leads us to assume that the Ugric sound change \*s > \**θ* is a late, parallel development (see the discussion on page 30 in the **Introduction**). Interestingly, Sammallahti (1988) mentions that this word shows irregular vowel correspondences. This seems not to be the case, in fact, as all the Ugric forms could be derived regularly from an earlier \**ä–ä* stem.

It would not be impossible to consider the Hungarian word a later loanword from an Iranian language. In fact, borrowing of a later form such as *hæft* would regularly lead to Hungarian *hét*. This possibility has been mentioned in FUV, which however also considers it possible that Hungarian *h* is a result of analogy.

The Avestan form *hapta* is irregular, as *f* would be the expected Iranian outcome of \**p* in this consonant cluster. The later Iranian languages point to regular *f*. It is impossible to see what was the outcome in the form from which the Ugric words were borrowed, as \**f* would have been substituted by \**p* in any case.

The Iranian etymology itself is convincing, but there is no reason to consider the word specifically an *Iranian* borrowing, as it can equally well have been borrowed from Proto-

Indo-Iranian as well. The dating mainly depends on the disputed age of the Iranian change *\*s > \*h*.

Napol'skikh has attempted to derive this word (as well as the Samoyedic word for 'seven') from Tocharian, but the flaws of his argumentation are shown in detail by Widmer (2002), who lists several phonological arguments against Napol'skikh's assumption and considers the Indo-Iranian etymology better. Interestingly, Widmer sees the Ugric vowel (cited as *\*ä* in his work) as a reflex of Indo-Iranian *\*ä*, an intermediary stage of the PII development *\*e > \*a*. It is indeed the case that even in the early loanwords acquired into the Ugric languages before the sibilant-changes, the substitution of the vowel *\*a* is often *\*ä*. However, it is unlikely that this is an intermediary stage between *\*e* and *\*a*; it is more probable that *\*ä* is just a substitution of *\*a*. This substitution is found in many loans in Ugric/Ob-Ugric, some of which have to be quite late. This shows that *\*ä* was probably the prevailing substitution for a long period, starting with very early loanwords, which means that at least at the time of the Proto-Iranian loanwords, the substitution patterns in the east (Ugric) and the west (West Uralic?) of the Uralic family were different (in the west *\*a*, *\*o* and *\*i* are found, but there are no reliable cases of *\*ä* among the early loanwords). This is only natural in the light of the large geographical area where the contacts occurred.

An Indo-Iranian etymology (parallel borrowings)

### Md E *šeja*, M *šava* 'goat'

< ? Pre-Md < *\*šaya* (Joki 1973: 317–318)

← ? PII *\*sčāga-* > OI *chāga-* M 'billy-goat', *chāgā-* F 'she-goat', Oss *sæy*, *sæyæ* id. (no further cognates within or outside Indo-Iranian) < ? PIE/Pre-II *\*skēgo-* (EWAia I: 558–59, s.v. *chāga-*; Lubotsky 2001a: 35; 2001b: 312)

(FUV: 136; Joki 1973: 317–18; Katz 1985: 257–258; Lushnikova 1990: No. 44; Helimski 1991: 221; Rédei 1986c; Zaicz 1998: 212)

This is an old and well-known etymology. As Zaicz (1998: 212) has pointed out, the borrowing has to be old, despite its limited distribution to Mordvin (other such cases are *\*pakas* and possibly *\*tarvas*). The Mordvin palatal sibilant points to an early borrowing, as a later loan from Iranian would have resulted in a form like *\*sava*. If the word was acquired already into PU or some other proto-language, it is easy to explain its disappearance from other languages, such as Finnic, where the word for 'goat' is borrowed from Baltic (PFi *\*wooši* from Baltic *\*ōži*).

The relationship of the Erzya and Moksha words seems irregular, but the front vowel in the first syllable of the Erzya word has been explained as a result of the palatalizing influence of word-initial *ś*. It can then be assumed that the Proto-Mordvin vowel was *\*a*, and it is still retained in the Moksha word. The word-internal consonant (E *j*, M *v*) results from hiatus caused by the lost *\*y*.

Helimski(1991) has noted that the Mordvin word is borrowed rather from Iranian because of Md *ś* pointing to early *\*ś* and not *\*č*. It is not clear what Helimski's arguments are, but a form with palatal *\*ś* would not typically be an Iranian loan according to modern

views of Proto-Iranian reconstruction. Also Zaicz (op. cit.) mentions that the word “looks like an Old Iranian word”, but it remains enigmatic what he actually means, as Zaicz does not mention any criteria for the definition of an Old Iranian loanword.

According to EWAia the background of the Indo-Iranian word is unclear<sup>41</sup> (no cognates in other branches of Indo-European have been presented), but as the word has cognates in both Indo-Aryan and Iranian, it can be reconstructed for Proto-Indo-Iranian; EWAia notes a bit mysteriously that the Ossetic words have to be taken into account, but it does not take a clear stance on the exact relationship between the Indo-Aryan and Ossetic words. It is surprising that this word is not attested elsewhere in Iranian, but it is of course possible that the word has been retained only in Ossetic. Lubotsky (2001b: 312) considers *\*sčāga* a loanword (a substrate word) in Indo-Iranian. EWAia refers to both Joki and Katz, who argue that the direction of borrowing is nevertheless from Indo-Iranian to Uralic and not vice versa (Burrow 1955: 26 had assumed that the Indo-Iranian word could be borrowed from Uralic). It would be very difficult to explain the Indo-Iranian cluster *\*sč* (which has to be reconstructed on the basis of the Ossetic and Indic forms) from Pre-Mordvin *\*č* or *\*š*, and also the confinement of the Uralic word to Mordvin speaks against the idea of a reverse borrowing.

Lubotsky (2001b: 309) further argues that the Uralic word has to be borrowed from the Iranian branch due to chronological reasons (cf. *\*worači*, *\*werkkä*), but as noted above, this does not look probable because of the palatal sibilant of Mordvin – here one can refer to Lubotsky’s own (2001a) work, where *\*sč* is reconstructed for the Proto-Indo-Iranian word, and the Iranian reflect of this cluster was probably *\*sts*. Lubotsky (2001b: 9) notes that Oss. *sæγ*, *sæγæ* ‘goat’ points to a short vowel in the root, although he also states a bit ambiguously that “theoretically, *\*skēgo-* is possible”.

This is one of the etymologies where the Indo-Iranian donor form begins with PII *\*sč* (from PIE *\*sk*). *suoja* and *hoitaa* are other such cases, but they have to be loans from later Iranian forms because of the divergent sibilant substitutions. There are different views on the precise background of the PII *\*sč* (resulting in the correspondence Iranian *s* ~ Indo-Aryan *(c)ch*; see Lubotsky 2001a and Kobayashi 2004 for various views on these developments), but from the point of view of loans like Mordvin it is irrelevant whether the background of this sequence was simply PIE *\*sk* in palatal environments (as argued by Lubotsky 2001a) or PIE *\*skʹ* (as the more traditional view goes). The Mordvin word’s palatal sibilant fits well with Lubotsky’s reconstruction of Proto-Indo-Iranian *\*sč*.

To sum up, it is clear that the Mordvin word has to be a relatively early borrowing from Proto-Indo-Iranian.

## An Indo-Iranian etymology

---

<sup>41</sup> According to Katz (2003: 207–8), West-Germanic (?) *\*skēpa-* N ‘sheep’ might be related to the Indo-Iranian word through a dissimilation of the word-internal consonant, reflecting a vrd̥dhi-derivation *\*skʹēgo-* from PIE *\*skʹēgo-*. Katz’s idea of assimilation is an ad hoc solution, and it cannot be supported. Furthermore, Kroonen (2013: 443, s.v. *\*skēpa-*) states that the Germanic word is probably derived from the verb *\*skaban-* ‘to shave, shear’. Furthermore, Kroonen argues (based on OHG *sčāpāri*, *sciiāppāri* and MHG *schāpære*, *schæper* M ‘sheepskin’), that the Germanic word goes back to a PIE form with *\*bʰ-*. There are thus no reasons to support Katz’s idea of the etymological connection of the Germanic and Indo-Iranian words.

**PU \*śaŋka ~ \*ćaŋka- ‘to sting’**

See \*śaŋkV for references and discussion

? Indo-Iranian etymology

**\*śaŋka ~ ~\*ćaŋka ‘handle’**

See \*śaŋkV for references and discussion

An Indo-Iranian etymology

**Fi sata (present in all Finnic languages); SaN čuohti (present in all Saami languages; < PSa \*čuoṭē); Md E śado, M śada; Mari E šüðö W šüðə ‘100’; Ko so, Ud sú (< PP \*šo); Hu száz (: száza-); Mansi South šā<sup>et</sup>, East šāt, West sēt, sāt ‘hundert; viel’ (< PMs \*šitV); Khanty South sot, East sat, North. sat ‘100’ (< PKh \*sāt)**

< PU \*śita ‘hundred’ (UEW s.v. śata; Sammallahti 1988: 549; Zhivlov 2014: 119, 132; Aikio 2015b: 60)

← PII \*ćatá-m > OI śatá- ‘100’, OP θata-, Av. sata- (satəm) id. < PIE \*k̑m̑tóm ‘100’ (EWAia II s.v. śatá-)

(Munkácsi 1901: 559–562; Jacobsohn 1922: 38, 124; FUV: 136; Korenchy 1972: 71 No. 28; Joki 1973: 311–312 No. 135.; Harmatta 1977: 171, 175; UEW s.v. śata; Katz 1985: 367; 2003: 291; SSA s.v. sata; Koivulehto 1999a: 216; 1999b: 329, 331, 343; 2001b: 248; 2016: 218, 270, 290; Kallio 2006: 12, footnote 9; J. Häkkinen 2009: 21; Bereczki 2013 s.v.)

This is one of the best-known Indo-Iranian loanwords in Uralic, and it has usually been assumed that the word was borrowed already into Proto-Uralic (~ Proto-Finno-Ugric). The etymology is convincing both semantically and phonologically. However, even this very transparent etymology involves numerous details that require comment.

One issue is whether the word was borrowed into a unitary proto-language, or whether the forms (or a portion of the forms) in various Uralic languages represent parallel borrowings from Indo-Iranian. The word is reconstructed as \*śata in earlier sources such as the UEW and even by Sammallahti (1988), but it is obvious that many of the words cannot be explained from an old \*a-a-stem. The Mansi and Mari clearly point to \*i, and in many recent works (such as J. Häkkinen 2009; Zhivlov 2014; Aikio 2015b) the word is reconstructed as śita. Almost all of the Uralic words can be regularly derived from this reconstructed form: Khanty \*sāt could reflect either \*śata or \*śita, and as the \*a-a and \*i-a-stems merged in Saami and Mordvin, as well as in Hungarian, and \*i and \*a in almost all instances in Finnic (Aikio 2015b; Zhivlov 2014: 115–117), the forms in these languages can also regularly reflect \*śita.

However, the Permic forms are more problematic, as *\*i*-*a*-stems regularly produce *u* in both Udmurt and Komi. According to Metsäranta (personal communication) the Permic correspondence Udmurt *u* ~ Komi *o* which is not typical of either Uralic *\*a*-*a* or *\*i*-*a* combinations but is typical for loans (the same vowel correspondence is interestingly found in the Permic reflexes of *\*asora*: Aikio [2015b: 37–38] argues that in that case the Permic vowels reflect the PU *\*a*-*o* combination, see the entry *\*asora*). Zhivlov (2014: 124–125) argues, however, that the Permic development is regular before *\*t*. If this is correct, it is possible that also the Permic words can be regularly derived from *\*šita*. but as the vowel-combination appears in other early loanwords too, it remains a possibility that the Permic words reflect an early parallel borrowing.

Bereczki (2013 s.v. *šüdö*) has attempted to derive the Mari vocalism from PU *\*a* through a chain of changes: *\*ā* > *\*o* > *\*u* > *\*ü*, but this is not a convincing explanation, as PU *\*a* is never reflected as Mari *ü*, and reconstructing *i* is more economical.

Because of the irregularity in the vowel correspondences, Kallio (2006) has suggested that the word could have been borrowed separately at least into Ugric and Finno-Permic. Kallio’s idea of parallel borrowings is certainly plausible, but it seems unlikely that the word was borrowed separately into western and eastern languages, but possibly the picture is more complicated because of the Permic words. As the forms in all of the other languages could regularly be derived from *\*šita*, we would only need to assume that Pre-Permic borrowed the word separately. It is unlikely that Permic was the first one to branch off from Proto-Uralic, so this explanation is problematic. One possibility would be to assume that the word was borrowed separately into several dialects of Proto-Uralic, and the regularity of the non-Permic languages is only apparent. Some languages would have substituted the PII vowel *\*a* by Uralic *\*a*, some by *\*i* (this remains uncertain in most cases because of the vowel mergers discussed above).

As a side note it can be mentioned that Zhivlov (2014: 119, 132) reconstructs the word as *\*čita<sub>l</sub>*, with *\*a<sub>l</sub>* in his system where the Mari stem vowels and Khanty vowel combinations point to two different vowels in Uralic *\*a*-stems (see op. cit. 117–120). Interestingly, Zhivlov notes that in Permic *\*i* could develop into Udmurt *u*, Komi *o* before *\*t*, which would make the Permic words a regular cognate of the other Uralic forms.

Even if the words in Uralic are parallel borrowings, as seems probable, the loans have to be very early because of Uralic *\*š* that has to be the substitute of PII *\*ć* and not a later Iranian reflex (*\*ts* > *\*s*). The claims by Korenchy (1972) that the Uralic words could be borrowed from “PI *\*šata*” are now outdated (they reflect the view that Iranian had *\*š* from PIE *\*k<sup>ʰ</sup>*). This view was also supported by Joki who mentions Iranian as one possible original for the Uralic words (Joki considers the donor form “ar. (bzw. frühuriran.?)”).

PU *\*š* is the usual substitution for PII *\*ć* (for example PII *\*ćṛwā-* > PU *\*šarwi*), but at times this idea has been questioned and the word for ‘hundred’ and other words that show *\*š* have been considered to be borrowings from Proto-Indo-Aryan where *\*ć* had already become *\*š*. Koivulehto (1999a: 219–220) mentioned such a possibility, and the Indo-Aryan origin of these words has been supported by J. Häkkinen (2009: 24). The idea that *\*š* would reflect Indo-Aryan *\*š* would of course make sense phonologically, but already Koivulehto (1999c) mentioned cautiously that the PU *\*š* might be due to lack of *\*ć*, at least in word-initial position. Zhivlov (2014: 114, footnote 3) has more clearly expressed the view that PU *\*š* and *\*ć* are actually the same phoneme (see the Introduction), and if this is indeed the case, this would explain the substitution.

To return to the vocalism of the Uralic word, \**ɨ* is found as the substitution of PII (and PI) \**a* in several convincing loanwords. J. Häkkinen (2009: 22–24) argues that the different substitutions of \**a* and *ɨ* are due to differences in the substitutions of accented and unaccented \**a*. While Häkkinen’s idea is promising, he based it only on a limited number of etymologies. It also leaves the problem of Uralic \**o* from PII \**a* unexplained. It seems that actually a very large number of the loans can be explained as having original \**ɨ* vocalism.

An Indo-Iranian etymology (possibly parallel borrowings after the split-up of Proto-Uralic)

**Ko *śurs*, Ud *śurs* ‘1000’ (< PP \**śurs*); Ms North *sāter*, West *šoater*, South *šatera* id.**

< PU ? \**śasra* ~ \**šisra*

← PII \**ǵhásra*- ‘thousand’; \**sa-ǵhásra*- “one-thousand” > OI *sahásra*- N ‘thousand’, (y)Av *hazaŋra*- id.; < PIE < \**sm̥ǵhéslo*- “ein habend”, root \**ǵhés*- ‘[mit der hohlen Hand] fassen’ (EWAia II: 719–20, s.v. *sahásra*-; Rix 1991: 226–228; Winter 1992: 20–21)

(Munkácsi 1901: 242; Jacobsohn 1922: 105–110; FUV: 137; Korenchy 1972: 71 No. 29; KESK: 175; Joki 1973: 318–19; Gamkrelidze & Ivanov 1984: 921; Katz 1985: 367; 2003: 291–292; UEW s.v. *śasra*; Rédei 1986c: 46–47; Lushnikova 1990: 293; Koivulehto 1999a: 216; Parpola 2015: 66)

This is a well-known and convincing etymology, first presented by Munkácsi. Interestingly words for both ‘hundred’ and ‘thousand’ were borrowed into Uralic from Indo-Iranian; \**śata* ‘100’ is better represented, but \**śosra* has to be an equally old borrowing, as the Permic and Mansi words clearly go back to an early stage and reflect an inherited word. (As has been noted in earlier research, a later borrowing from Permic to Mansi is ruled out by the regular vowel correspondences as well as by the development of \**ś* to \**s* in Mansi.) The word is clearly borrowed from Proto-Indo-Iranian (and not from Iranian) because of the palatal sibilant/affricate in Uralic (note that Korenchy reconstructs “uriran. \**ǵhasra*”). Although it is rightly debated whether the various Iranian developments of sibilants/affricates had taken place in Proto-Iranian already, here Korenchy makes a simple mistaken by reconstructing an aspirated (!) palatal sibilant into Proto-Iranian). No other Indo-European branch could come into question as a donor language; Latin *mille* (< \**smih<sub>2</sub>ǵh<sub>2</sub>esl-ih<sub>2</sub>*)<sup>42</sup> and Greek χίλιοι ‘thousand’ (< Proto-Greek \**χ<sub>2</sub>εσλ-*, cf. Ionic χεῖλιοι, Emmerick 1992a: 176) ultimately include the same elements as the Indo-Iranian word, but no exact formal parallels to \**sa-ǵhásra*- are found in other branches (see the entry on EWAia), and in any case, the change \**l* > \**r* and the affricate reflex of PIE \**ǵh* point clearly to Indo-Iranian (\**ǵhéslo*- is not attested in any other satem language).

From the point of view of semantic and morphology, it is both interesting and troubling that only the element \**ǵhasra*- was borrowed into Uralic, as it is exactly the compound

<sup>42</sup> But note that Emmerick (1992b: 315) is more critical of this etymology of the Latin word.

\**sa-ǰ<sup>h</sup>asra-* that denotes ‘thousand’ in Indo-Iranian. EWAia, referring to the Habilitationsschrift of Katz even notes that the Indo-Iranian etymology for the Uralic word for ‘1000’ is wrong, exactly because of the odd borrowing of the sole \**ǰ<sup>h</sup>asra-*; EWAia also refers to an article by Emmerick (1992b) on Iranian numerals, where it is noted that in some later Iranian languages the first syllable of \**sa-ǰ<sup>h</sup>asra-* became lost because of secondary developments (cf. Sogdian *z’r*, Pashto *zəŕ*), and a simple form \**ǰ<sup>h</sup>asra-* cannot be reconstructed for Proto-Indo-Iranian in the meaning of ‘1000’. It would be too big a coincidence for the Uralic and Indo-Iranian words to resemble each other so much, so despite the understandable cautiousness of EWAia, the Indo-Iranian origin of the Uralic word has to be considered secure.

Katz (2003: 296–97) attempts to explain the Uralic word in the following way: the word was borrowed as \**sašara* from PII \**sa-ǰ<sup>h</sup>asra-*, and later the Uralic word underwent metathesis to \**śasara*, from which the Mansi word can be derived from through syncope and the Permic word through syncope and metathesis; Katz explains this through influence from the word \**śata* ‘100’. This is an interesting explanation, but it is more economical to support Jacobsohn’s original idea (although Katz calls this the “worst possibility”) and simply suppose that the Uralic word reflects PII \**ǰ<sup>h</sup>asra-* and not \**saǰ<sup>h</sup>asra-*.

Regarding the sound substitutions, it has to be stated that the exact vowel substitution is difficult to determine, as it is uncertain whether the Mansi and Permic word actually reflect a regular Uralic proto-form. Koivulehto (1999a: 219) mentions that the word could reflect either \**o* and \**i*. More traditionally the Uralic word was reconstructed as \**śarsa* or \**śasra* (Gamkrelidze & Ivanov 1984: 921; UEW s.v. \**śasra*; Korenchy 1972: 71–72; Joki 1973: 318–319), but the reconstruction of this word as an \**a*-*a*-stem can hardly be correct. The Mansi word looks like an old \**o*-*a*-stem. The Permic vowel correspondences *Ko u ~ Ud u* rather point to \**a*-*a* or \**i*-*a*.

The \**ś*/*ć* in the word-initial position clearly points to Proto-Indo-Iranian donor, as mentioned above, so the words will have to reflect a very early borrowing.

Also the UEW mentions the possibility of separate borrowings into Proto-Ugric and Proto-Permic (probably rather meaning Pre-Permic here), but this is because of the anomalous correspondence between the medial consonant clusters. Here probably a metathesis of the medial consonant cluster \**-sr-* > *-rs-* can be seen in the Permic words, but Mansi represents the earlier state of affairs.

A well-established fact is that Hungarian *ezer* ‘thousand’ is a much later (Iranian) loan from a reflex of this Indo-Iranian word; cf. Khotanese *ysāra-*, Middle Persian *hazār*.

An Indo-Iranian etymology

**Hungarian *szeg*, *szög* ‘nail’; Mansi North *saŋk<sup>w</sup>*, East (KL) *säŋk<sup>w</sup>* ‘Keil’, West *säŋk*, South *süŋ* (< PMs \**säŋ.kū* Zhivlov 2006: 11)**

< ? PUG \**seŋkV* > Pre-Ug \**śeŋkV*

← PII \**ćanku-* > Old Indic *śaŋkú-* ‘a peg, nail, spike’ < PIE \**kónkú-* (EWAIA II: 604, s.v. *śaŋkú-*; Garnier 2004–2005: 4–6)

(Munkácsi 1901: 565–566; Korenchy 1972: 67 No. 23; UEW; Abondolo 1996: 35; Katz 1985: 331; 2003: 263; Lushnikova 1990: 254; Helimski 1997b: 121)

This etymology is an example of a very early borrowing into Ugric, as both the Ugric and the Indo-Iranian side point to palatal \*ć in the donor form. The Mansi and Hungarian cognates are regular, which point to an inherited word (from Proto-Ugric?) and thus to an early borrowing. If the etymology is correct, it has to be a Pre- or Proto-Indo-Iranian loan, not Iranian. Korenchy considers the loan “Iranian”, but this is based on her views on Proto-Iranian reconstruction which assumes that the Iranian reflex of PIE \*k<sup>h</sup> would be \*ś. Because it now is obvious that the Iranian reflex was something else, probably \*ts, the palatal sibilant in the Uralic form can only reflect the palatal affricate of Proto-Indo-Iranian but not its later Iranian reflexes.

The background of the Indo-Aryan word is considered uncertain by EWAia (II: 604), but possible cognates in Lithuanian (*at-šankė* ‘Widerhaken’) and Old Norse (*hár* ‘Ruderpflock’) are mentioned. It is probable that the word continues a PIE form \*k<sup>h</sup>onkú- (Garnier 2004–2005: 4–6). Within Indo-Iranian, *śakala-* ‘Span etc.’ (602) might stand in a derivational relationship with *śankú-* (to this Mayrhofer connects, with reservations, also Lith *šakalys* ‘Holzspan’).

As the vowel goes back to IE \*o, it is unlikely that the Ugric word is a Pre-Indo-Iranian borrowing. Korenchy reconstructs the Ugric vowel as \*e, explaining this with influence from palatal \*ś – the “Iranian” original had \*a in Korenchy’s view. Korenchy also notes that one would expect \*ä instead of \*e to be the “palatalized” development of \*a. The Ugric vowel here seems to be rather \*ä. The Proto-Mansi form is \*sänk<sub>u</sub> according to Zhivlov (2006: 11) and Helimski (1997: 121). In the earlier reconstruction of the Proto-Mansi word, the first-syllable vowel was usually reconstructed as \*ü, and it is has been mainly the Indo-Iranian evidence which has made the reconstruction of an illabial high/central-vowel more probable – from a Mansi-internal point of view, both Proto-Mansi \*ü and \*ä would be possible. Mansi \*ä is found as a regular reflex of PU \*ä at least in many \*ä-stems, such as \*käjä ‘moth’ > PMs \*käj, \*äktä- ‘to cut’ > PMs \*jäkt-.

It seems that the Mansi form could reflect either Proto-Uralic/Ugric \*ä or \*e. Thus the word could theoretically be a very early “e loan”, or it could be a later loan where PII \*a is substituted by Ugric \*ä. There are many later loans in Ob-Ugric and Permic where this kind of substitution prevails, but there are not many possible examples of this substitution in earlier loans that are shared with Hungarian and the Ob-Ugric languages. The word for ‘seven’ is one such example, but because of its very complicated consonantism (pointing perhaps to independent borrowings by the various Ugric languages), it does not serve as a credible parallel. Widmer (2002) considers the \*ä in the word ‘seven’ to be a reflex of an early form of PII \*ä (see the entry for ‘seven’), and this word could be a similar example.

It seems that the Indo-Iranian vowel \*a here does not reflect PIE \*e but \*o, which forces us to consider this a loan from a “proper” Indo-Iranian form with \*ä. As mentioned above, an Iranian loan is out of question because of the palatal sibilant/affricate, so interestingly this has to be a rather old loan from Proto-Indo-Iranian (pointing to interesting differences in the substitution of PII \*a in the eastern and western Uralic languages) or a later loan from an unattested language where PII \*ć would have been retained longer. The last scenario could lend support to Helimski’s “Andronovo Aryan” idea, but is impossible to

prove (or to refute), and thus we are left to support the first scenario. This means that this etymology is a very early borrowing with the substitution \*ä ← \*a.

However, Pystynen (personal communication) has noted that the Hungarian word should rather reflect earlier \**siŋkV*, because Hungarian *ë* usually reflects PU \**i* (compare etymologies such as PU \**miñä* ‘daughter-in-law’ > Hu *mëny*, \**miñV* ‘heaven’ > Hu *mënnny*). This makes the loan-etymology more uncertain. However, counterexamples to this rule can be found: PU \**kelki-* ‘to be necessary’ is reflected by Hu *këll-*, but the cognates in at least Permic (Udmurt *kul-*, Komi *kov-*, *kol-*) show that the word cannot go back to PU \**i-*; also PU \**meni-* is reflected by Hu *mën-* (Finnic \**mene-* proves that the PU word had to have \**e-*), and a similar case is \**sewi-* ‘to eat’ > Hu *ë-*, *ëv-*. The precise reflexes of PU \**e*, \**i* and \**ä* would require more research, but it does seem correct that \**ä* is never reflected by Hungarian \**ë*. The Mansi word could also be derived from \**siŋkV*, as PU \**i* is usually reflected by Mansi \**ä* (for example PU \**ipsi* ‘smell’ > PMs \**ät* (~ ə). Mansi could reflect also \**e*, cf. \**miñä* ‘daughter-in-law’ > PMs \**mäñ* (~ ə); Mansi \**ü* often reflects PU \**ü*, cf. \**künči* ‘fingernail’ > PMs \**küş* (~ ə) id. (Mansi examples cited from Zhivlov 2006).

A parallel borrowing in the West Uralic languages from the same Indo-Iranian source or from a related Indo-European form might be \**ćaŋka* (Fi *sanka*, Mordvin E *šango*, M *šangă*, *ćangă* ‘handle’), as is suggested by Koivulehto (1983 [the PIE form is reconstructed as \**kanku-* by Koivulehto]). Holopainen (2018a) has argued that most of the examples of PIE \**k* being reflected by \**ć* in Uralic loans can be explained from later satemized forms, and thus this West Uralic word could also be considered as an Indo-Iranian loan. Also an alternative Germanic etymology has been suggested, PG \**stangō-* (LÄGLOS s.v. *sanka*). This does not explain the Mordvin word, though (*ś-* could hardly result from Germanic \**st-*), so the Indo-Iranian etymology is more convincing.

Another parallel loan could be the PU verbal stem \**ćaŋka-* (homonymous with the West Uralic noun \**šaŋka-* ~ \**ćaŋka*) which is reflected by SaN *čuoggut* ‘to sting, stick’ and Slk (Taz) *seŋki-* (< PSam \**sāŋkâ-*; Aikio 2006a: 24), if we assume that the verb ‘to sting’ can be derived from a noun with the meaning ‘nail’. Älgu (s.v. *čuoggut*) mentions that an Indo-European etymology for this word has been suggested by Sammallahti (1998: 126), but no such explanation is found in Sammallahti’s book and it remains uncertain whether Sammallahti was thinking of the same Indo-European word. Älgu gives no donor form or a more specific classification of the Indo-European donor. If \**ćaŋka-* is an Indo-European loanword, an Indo-Iranian source is the most probable, but because the Uralic word is a verb, the etymology remains uncertain.

? Indo-Iranian etymology

Ma *šār*, *sār* ‘Gemütsart, Charakter, Art und Weise, Gewohnheit, Beschaffenheit, Bau’; Ko *šer* ‘Gewohnheit, Sitte, Mode, Weise’ < PP \**šer*; Mansi North, West, East *sir* ‘Art, Weise’; Khanty East *sür*, *sěr*, North *sir* ‘Art, Weise’; Hu *szer* ‘Teil des Dorfes; Reihe, Ordnung; közösség; Gemeinschaft; Geschlecht; Art, Sorte, Gattung; Werkzeug, Instrument’; ?TN *šer* ‘Sache’, Enets *sieʔ* id., Nganasan *sier* id.

< ? PU \**šerV* ~ \**šertV* (UEW s.v. \**šer*<sub>3</sub> No. 957; Mikola 2004: 30–32; Honti 2017: 55)

← Pre-II \**sčerdʰo-* > PII \**sčardʰa-* > OI *sárdha-* ‘strength; army’, YAv *sarəða-* ‘Art, Gattung’ (EWAia II s.v. *ŠARDH*; Lubotsky 2001b: 50–53)

(see the entry on SaN *čearda*)

(Munkácsi 1901: 571; Jacobsohn 1922: 128; FUV: 137–138; KESK: 250–251; Harmatta 1977: 172; Katz 1985: 372; Katz 2003: 292–293)

The possible Indo-Iranian origin of this Uralic word is an old but problematic idea, and it is also problematic (or even unlikely) that the Uralic forms listed under this reconstruction are even cognates with each other. The etymology is not found in the sources written by modern Finnish loanword researchers but is supported by Katz (2003). The UEW mentions that Indo-Iranian etymology has been suggested, but it is not probable because supposedly there are no Indo-European loanwords in the Uralic proto-language (“Die Übernahme ist jedoch aus chronologischen Gründen unwahrscheinlich, da bisher noch keine ieur. Lehnwörter aus uralischer Zeit authentisch nachgewiesen wurden”). This argumentation is of course circular. Katz (2003) has actually derived this word from the stage he calls “frühurarisches” but which coincides with the widely accepted view of Proto-Indo-European: in Katz’s view the word was borrowed from a stage that had still retained \**k*, meaning that there is no sense in considering that stage “Indo-Iranian”. This assumption of Katz is not based on any real evidence (there is no way to show that Uralic \**š* here would correspond to PIE \**k* and not PII \**č*). Katz also could not take into account the later reconstructions of the Indo-European word: Lubotsky (2001a) has convincingly argued that the Indo-Iranian forms go back to a form with an *s*-mobile, and the Proto-Indo-Iranian word should be reconstructed as \**sčardʰa-*.

Even though the Uralic \**š* or \**č* could well be explained from a later Indo-Iranian form, the etymology is not very convincing. First of all, the forms of most Uralic languages cannot be explained from the Indo-Iranian form with the cluster *-rdh-*, as this would probably have been substituted by \**-rt-*, which has left no traces on the Uralic side here. The palatal sibilant in Khanty and Nenets is totally irregular. The Nenets word probably does not belong here at all, as SW (67–68) reconstructs Proto-Samoyed \**ker* which cannot be derived from an earlier form with \**š*-. Mikola (2004: 27–34) has attempted to connect the Samoyed and other Uralic forms by assuming a Proto-Uralic affricate \**č* which would give Samoyed \**k*, but this idea is based on very few examples and is not convincing.

The Hungarian word *szēr* is probably borrowed from West Old-Turkic \**šer* (cf. Ottoman Turkish *yer* ‘Erde, Boden; Ort’; WOT: 795–796). The authors of WOT also suspect that there might be two (nearly) homonymous words in Hungarian: *szer* ‘Art’ derived from this Uralic word and *szēr* ‘a group of houses’ borrowed from Turkic, but this

is refuted by Honti (2017: 55), who argues that there is only one word *szér* in Hungarian, and this is inherited from Proto-Uralic. Neither the authors of WOT nor Honti take into account the severe problems connected with the reconstruction of the Uralic word. Despite Honti's reluctance, Róna-Tas (2017: 54) further argues that Hungarian *szēr* with the meaning 'a group of houses, a territory with a group of houses' is borrowed from Turkic. Possibly some of the other Uralic words can also be derived from the same Turkic word, but this will have to be left to future research.

Not an Indo-Iranian etymology

**Md E *kšta*, *šta*, M *šta* 'beeswax'** (< PMd \**šššta* Keresztes 1986: 161–162); **Ma *ššte*, *ššte*** < PMA \**šišta*; **Ko *šis*, Ud *šus* id.** < PP \**šiš(t)*

< PU \**šišta* 'beeswax'

← PIA/PII \**čišta-*, > OI *madhu-šišta-* 'beeswax'; *šišta-* 'übergeblieben; Rest', verbal adjective from the root *šes-* 'to leave behind' < ? PII \**čayš-* (EWAia II: 655, s.v. *šes*)

(Katz 1985: 280–281; 2003: 224; Carpelan & Parpola 2001: 122–126; Parpola 2012: 161; 2015: 63; 2017: 261)

This etymology has been suggested independently by both Katz and Carpelan & Parpola. The etymology is both formally and semantically credible, and offers a good explanation for a seemingly PU word limited to Mordvin, Mari and Permic. It is interesting to speculate whether this word was absorbed into the central dialects of PU, or whether it was acquired independently by the three branches. Formally there is no difficulty in deriving the words from a common PU or "Finno-Permic" form. One could suppose that the reflexes of the word in other languages, such as Finnic, were ousted by more recent loanwords, i.e. the Finnic word for wax, *vaha*, is a well-known Baltic borrowing. There is a handful of loans which have a "Central Uralic" distribution, and it is often very difficult to determine exactly whether they are very early parallel loans or words that have been lost in other branches. See the entries on \**mertä*, \**sejti*. The words in Mordvin, Mari and Permic could also be parallel loans, but there is no need to assume this because they can reflect PU \**šišta* regularly.

According to Kallio (p.c.), Mordvin *k* in *kšta* (< Pre-Md *šikšta*) results from secondary epenthesis. A similar development is also found in Mordvin E *kšna*, M *šna* 'Riemen' (< \**šišna*), a probable Baltic loan<sup>43</sup> (Keresztes 1986: 68–69). Keresztes (1986: 161–162) also notes that the *k* in Erzya *kšta* is "unetymologisch". It seems that Keresztes assumes that the *k* cannot be reconstructed even for the Proto-Mordvin form, which he reconstructs

<sup>43</sup> The suggested Baltic origin \**šišna* (> Lithuanian *šikšna* 'Riemen'; also Fi *higna* is borrowed from this Baltic source) is probable, but the Baltic word does not have a proper Indo-European etymology (the suggested cognates are doubted by KEWA s.v. *šikya-* and Beekes 2010: 704–705, s.v. κίσσοϛ). According to Kallio (p.c.), here the Baltic word shows similar *k*-epenthesis which may be an areal feature shared by Mordvin and Baltic. The relationship of Mari *šüštö* and the Mordvin, Finnic and Baltic words remains ambiguous, as the Mari vocalism cannot be easily derived from Baltic \**i*.

as \*šāšta. It is possible that *k*-epenthesis is then a uniquely Erzya development, as no trace of *k* is found in Moksha. Regarding the etymology of E *kšna*, M *šna*, Keresztes does not comment on the secondary nature of *k*, which he reconstructs already for Proto-Mordvin \*šukšna, but not for Proto-Finno-Volgaic \*šišna.

A more problematic assumption is Carpelan & Parpola's idea to consider this word an *Indo-Aryan* loan; they argue that the word is missing from Iranian. It remains uncertain whether the Indo-Aryan root has a cognate in Iranian or not. Parpola (2017: 261, footnote 9) notes that there are varying views about the etymology of the Indo-Aryan root, with a possible connection to Avestan *siiazd* suggested by Hintze (1994) and Werba (No. 195 s.v.), although Lubotsky (2004) and Cheung (2007) have rejected this explanation.

If we assume that the idea of Hintze is correct, then we can safely consider this a Proto-Indo-Iranian loan. A formal parallel to *šišta-* is not found in Iranian, but the verbal adjectives in *-ta-* (< PIE \*-to) were, of course, a productive and frequent word-formation type in the Iranian languages too. LIV<sup>2</sup> (s.v. \*k'eis-) mentions the PIE root \*k'eys- as a possible pre-form of the Indo-Aryan verbal root, but this source does not give any parallels in other languages, so this connection remains hypothetical. It is thus impossible to state for certain that this is not an Indo-Aryan loanword, but formally and phonologically the word could have been borrowed from Proto-Indo-Iranian, provided that PIA \*šayš- reflects PII \*ćayš-.

Also other central terms of apiculture were borrowed from Indo-Iranian into Uralic, notably \*meti 'honey' and \*mekši 'bee', which belong to the oldest layer of borrowings (see the entries on these words). \*šišta is probably a later borrowing, but theoretically it could be as old as the other words mentioned above, as it is uncertain what the development of PIE \*k' was at the "Pre-Indo-Iranian" stage and how this would have been substituted.

This word shows the RUKI development of PIE \*s (with \*mekši, \*wiša as possible parallels) with PU \*š as the substitute for Indo-Iranian \*š.

An Indo-Iranian etymology

**Fi.** *salko* 'long pole or rod'; **SaN** *čuolggu* 'rod for pushing nets under ice', **Erzya** *šalgo* 'stick, thorn, stinger'; **Komi** *šuw*, **Udmurt** *deđ'i-šul* 'sled runner' (< PP *šul*); **East Khanty** *sayᶇa* 'lath' (< PKh \*sāyᶇa); **North Mansi** *sāyla* 'lath' (< (PMs \*sīylā); **Hungarian** *szál-fa* 'lumber wood; long log'

< PU \*ćilka-j ~ \*šilka-j/\*ćilka-w ~ \*šilka-w (UEW: 460–61, No. 929; Zhivlov 2014 132; Aikio 2015nb: 60)

← ? PII \*ćakala- > OI *śákala-* M 'a chip, fragment, splint, log, piece, bit' (EWAia II s.v. *śákala-*)

(Munkácsi 1901: 555–556; Korenchy 1972: 66–7 No. 22)

or

← PIE/Core IE \*g<sup>h</sup>algho-, \*g<sup>h</sup>alghah<sub>2</sub> > PB \*žalgas, \*zalga > Lith *žalgas*, *žalgà* 'long thin pole', PG \*galgan- > OHG *galgo* 'pole of a draw-well; gallows';

? **Armenian *jatk* ‘rod, stick, staff’** (Kroonen 2013 165, s.v. *\*galga-n*; Derksen 2015: 511, s.v. *žalga*)

(Koivulehto 1983: 113; 1999c: 210; 2001: 238 = 2016: 93, 212, 280; SSA s.v. *salko*)

This word has two competing Indo-Iranian etymologies. The earlier Indo-Iranian etymology (see Korenchy 1972: ) requires metathesis and is thus less likely: the traditional view was that the origin of the Uralic word was Indo-Iranian *\*čakala-* (> OI *šákala* ‘a chip, fragment, splint, log, piece, bit’), which is etymologically unrelated to the Indo-European word that is attested in Balto-Slavic, Germanic and possibly Armenian. It is even uncertain whether the Indic word goes back to Proto-Indo-Iranian, as EWAia states that no secure cognates exist (Lithuanian *šakalyš* ‘Holzplan, Splitter’ is mentioned as an uncertain cognate). This connection is not mentioned by Derksen (2015 s.v. *šakalyš*). Korenchy’s etymology would require metathesis on the Uralic side. Semantically this is not a very good etymology, as it can explain the semantics of the Ugric words but not the meaning ‘pole’ of the more western languages.

Koivulehto’s newer etymology does not leave room for a possible Indo-Iranian origin, as this Indo-European word is attested only in Germanic, Balto-Slavic and possibly Armenian. It has actually been noted by Kroonen (2013 s.v. *\*galgan-*) and Derksen (2015 s.v. *žalga*) that the Northwest Indo-European word does not have a good Indo-European etymology and is only shared by the two northern branches, Germanic and Balto-Slavic. Neither of them gives any further Indo-European etymology for the word, but the Uralic etymology is mentioned by both. The possible pre-form of the Germanic and Balto-Slavic words is reconstructed with PIE *\*o* by both Derksen and Kroonen, but the vowel could equally well be *\*a* here, as *\*a* and *\*o* merged in both Baltic and Germanic. The relation of the Armenian word to the Germanic and Balto-Slavic forms is considered as uncertain by Martirosyan (2010 s.v. *jatk*).

Kallio (ms.) has assumed that the Uralic word is borrowed more specifically from Balto-Slavic *\*žalgo-* and not from an earlier Proto-Indo-European or Proto-Northwest-Indo-European form. Kallio’s suggestion is clearly the best etymology suggested so far. The Uralic sibilant *\*ś* can be easily explained from Proto-Balto-Slavic *\*ž*. Most of the cases where Uralic *\*ś* could reflect PIE *\*g’* or *\*g<sup>h</sup>* involve various problems (Holopainen 2018a), but Uralic *\*ś* from Balto-Slavic *\*ž* is a completely plausible sound substitution. Also the vowel substitution is unproblematic, if we assume that the Balto-Slavic origin had *\*a*, for numerous parallels for the substitution of Indo-Iranian *\*a* by Uralic *\*i* exist, and it is possible to assume a similar development in Balto-Slavic loanwords, as Proto-Balto-Slavic likewise had a very simple vowel system compared to that of Uralic.

Regarding the etymology of Armenian *jatk* ‘rod, stick, staff’, Martirosyan (2010: 439, s.v. *jatk*) considers it possible that the word is related to the Germanic and Baltic words, but he also gives an alternative etymology which connects *jatk* with Armenian *joł* ‘pole’ and Lithuanian *žúolis* id.

Korenchy mentions (referring to MSzFE) that Hungarian *szál* cannot be related to *szalag* ‘Band, Streifen’, *szalu* ‘Deichsel’ or the dialectal verb *szil* ‘schneiden, spalten’, as has sometimes been assumed. Also SSA mentions that Finnic *salko* could be derived from the verb *salia*. Hungarian *szil* and Finnish *salia* both reflect a PU verb *\*šili-*, and theoretically it would be possible that the Uralic word is a derivation from this verbal root. This has also been suggested by Niklas Metsäranta in personal communication. Katz (2003: 246–247)

also mentions this possibility, and assumed that OI *śálka*- ‘Span, Abschnitzel’ is borrowed from the Uralic word. EWAia (II s.v. *śala*-) considers *śálka*- a derivative of *śala*- ‘Stab, Stock’, which makes the etymology of Katz improbable.

The labial vowel in Finnic and Saami is due to a suffix *-j* or *-w*. It is interesting that the primary function of such a suffix is uncertain, as this appears also in some kinship terms like *\*kăliw* and *\*natiw*. If the Uralic word was borrowed from the *o*-stem *\*g<sup>h</sup>alg<sup>h</sup>o-*, the suffix *-w* might have been used to substitute for the labial vowel of the donor form. However, it is possible that by the time the word was borrowed from Balto-Slavic, the short *\*o* had become *\*a* (cf. above), and the suffix here has nothing to do with the vowel of the donor language.

Not an Indo-Iranian etymology (another IE etymology is possible)

**Fi** *solki* ‘fibula’, **Estonian** *sõlg*; **SaN** *čulgum* (< **PSa** \*); **Mordvin** *śulgamo*, *śulgam*; **Mari** *šolkama*, *šălkama* (< **PMA** \*) (UEW s.v. *\*śolke*; Sammallahti 1988: 553; SSA s.v. *solki*; Zhivlov 2014: 139)

< ? **PU** *\*ćolki*

← **PII** *\*ćalHā-ka-* > **OI** *śalā-ka-*, derived from *\*ćalHā-*, see below (EWAia II: 615–616, 618 s.v. *sáru-*)

**or**

← **PIE** *\*kólh<sub>2</sub>-o-* > *śalā-* ‘rod, stick’ (EWAia II: 615–616, 618 s.v. *sáru-*)

(Koivulehto 2001b: 243 = 2016: 285)

This very complicated etymology. Although the PU word may be borrowed from PII *\*ćalā-ka* > OI *śalā-ka-* ‘rod’, the IE etymology of this word is unclear (EWAia II: 620) and it is not necessarily from earlier *\*k<sup>h</sup>ólh<sub>2</sub>-o-*. The irregular relationship of the various Uralic words makes a Proto-Uralic reconstruction improbable.

Although Koivulehto’s etymology looks simple, this etymology involves numerous problems, both on the Uralic and especially on the Indo-European side. First of all, it is not at all clear that the Indo-Aryan forms that Koivulehto mentions are actual cognates of the other Indo-European forms which he lists. This means that we might not be able to consider this word a possible Indo-Iranian borrowing, as Koivulehto has done. It has to be noted that if the Indo-Aryan words are not derived from this Indo-European source, there is no reason to assume that the word had a laryngeal and can be reconstructed for a form like *\*ćalHa-*.

Also, the *\*l* speaks against an Indo-Iranian origin. Most of the earliest loanwords have *\*r* from PIE *\*l*, even the Pre-Indo-Iranian loans such as *\*kekrä*, and the few counterexamples can be explained as false etymologies (see the entry on *\*śali* above). There is thus no reason to suppose that Proto-Indo-Iranian had a phoneme *\*l* retained during the times when it was in contact with Uralic. This also fits the general views on Proto-Indo-Iranian reconstruction (Mayrhofer 2004).

It is further unclear whether Finnic *solki* and the Saami, Mari and Mordvin words are even related. The vowel correspondences are irregular (Saami *u* cannot reflect earlier \**o*), and the Saami, Mari and Mordvin forms reflect a trisyllabic stem. Also Koivulehto notes this irregularity, but he only states that irregularity points to borrowing. Zhivlov (2014: 139) only considers the Finnic and Mordvin words cognates and he does not comment on the Saami, Mari and Permic words in any detail. Sammallahti (1988: 553) reconstructs the proto-language forms as \**śolkama*, but this does not explain the Finnic *i*-stem nor the Permic forms. The words are not mentioned in the list of Uralic back-vocalic words compiled by Aikio (2013). Theoretically the words could be parallel borrowings from Indo-European, but no good original can be found. Formally a satem form akin to Latin *culmus*, Latvian *salms* and Germanic *halma-* that reflect PIE \**k'olh<sub>2</sub>-mo-* or to Greek κάλαμος from \**klh<sub>2</sub>-mo-* (Beekes2010: 621–622; Kroonen 2013: 204–205; de Vaan 2008: 150; Derksen 2015: 235–236, 548) could come into question, but semantically they are not very convincing, as these words mean ‘straw’ or ‘reed’, and it would be difficult to assume that the words in Mari and Mordvin would be borrowed from Germanic, let alone Greek or Italic. Balto-Slavic \**śalma-* could formally be the best candidate, but it is likewise far from the Uralic meanings semantically, and due to the irregularities of the Uralic forms, we should assume very early parallel-borrowings in this case. It is not even certain that all these words indeed reflect the same Indo-European root. The Indo-Iranian word is not mentioned by any of the Leiden dictionaries mentioned above, making the Indo-European background of the word uncertain. On the whole, even if the root \**k'elh<sub>2</sub>* itself can be reconstructed for Proto-Indo-European, many of the cognates mentioned in earlier sources are on shaky grounds.

The UEW notes that Chuvash *śalGeme* ‘Brustspange der Frauen’ is borrowed from Mari. According to Bereczki (2013) the Mari word has not been borrowed into Chuvash, as is usually assumed, but both stem from a substrate language in the early modern era. It would be very difficult to prove this statement correct.

Problematic is the relationship of the Mari forms *śolkama*, *śolkama* and *śarkama*, *śarkämä*. The UEW mentions the forms with *r* with a question mark, but the reason for this *r* remains mysterious. Rather ambiguously the UEW also note that Mari should regularly show \**-ly-* instead of the attested *-lk-*, and that the *-k* might result from the influence of *śarkama*. This explanation does give any information on the cluster *-rk-* in the first place.

The Uralic stem vowel is also unclear, as Finnic points to an *i*-stem, and the rest of the forms are some kind of obscure derivatives. Uralic *i*-stem would be unexpected, as most borrowed Indo-Iranian *a*-stems show \**-a* in the Uralic side too. Zhivlov reconstructs the word as an *i*-stem, but does not comment on the stem-vocalism of the Mordvin word in any detail. This alone would not be an obstacle, but combined with the other uncertain issues, this also speaks against the etymology.

To sum up, the more precise origin and relationship of the Uralic words remains uncertain, but there is no reason to assume that these words are borrowed from Indo-Iranian. It is possible that the Uralic words represent parallel borrowings from some unknown source. This is admittedly not a very precise explanation for the origin of the word, but as there are many irregularities in the shared vocabulary of the so-called Finno-Permic languages, possibly resulting from early borrowings from unknown sources, this explanation is better than the Indo-Iranian etymology.

Not an Indo-Iranian etymology

Fi *suka* ‘haircomb; Borste; Bürste, Striegel, Kamm’, Karelian, Veps, Est *suga* ‘Bürste, Kamm’ (cognates in all Finnic languages); SaN *čohkut* ‘to comb; kämmen, bürsten, striegeln’ (has cognates in all Saami languages except Akkala < PSa \**čoko-*); ? Mo *šuva* ‘husk of grain; Granne, Achel, Spreu’; Mari *šu* ‘husk of grain; Kleie, Granne, Achel; Schweinborste’ (< PMa \**šu*); Ko *šu* ‘rye; grain; Roggen, Korn, Getreide, Getreidekörner’ (< PP *šu*)

< ? PU \**šuka* ‘awn’ ~ \**ćuka*

← PII \**ćuka-* > Av *sūkā-* ‘spike, needle’, OI *šūkā-* ‘awn, stangle’ Oss *syg* ‘awn’; cf. also Ved. *suci-* (< ? earlier \**šuci-*) ‘needle’ (EWAia III: 494–5 s.v. *šūkā-*; Lubotsky 2001b: 309)

or

← PBSI \**šukā* ‘comb’, cf. Lith pl. *šùkos* ‘comb, woolcomb’ (Frankel 1962–65: 1031, s.v. *šùkos*)

(SKES s.v. *suka*; Joki 1973: 315–316; Katz 1985: 268–269; 2003: 215–216; Redei 1986: 59–60; Lushnikova 1990: 261–262; SSA s.v. *suka*; Kallio 2009: 32–33; Junttila 2012)

The Indo-Iranian etymology is an early one but complicated because of the false cognates that have been lumped under one Uralic entry in many earlier sources. Although the Finnic (Fi *suka*) and Saami words are connected to the words in Volgaic and Permic by Joki, Kallio (2009: 32–33) separates them from these and considers them Proto-Balto-Slavic loanwords. Semantically the connection between the Finno-Saamic words and the rest is odd, and different origins for these apparently similar words is a better explanation. Phonologically there would not be a compelling need to derive them from different sources (all could be borrowed from PII \**ćukā*), but semantically it is better to separate these words like Aikio does. Fraenkel (1962–1965: s.v. *šùko*) notes that the background of the Baltic word is unclear, but the Lithuanian word could regularly reflect Balto-Slavic \**šukā*.

The Indo-Iranian etymology for Volgaic and Permic words is convincing. The word has to be borrowed from Proto-Indo-Iranian and not from some later Iranian form because of the palatal sibilant in Uralic, so the loanword must be ancient despite its limited distribution in the centre of the language family. In the UEW the donor form is classified as “uriran.” but this view is outdated.

The words in Mordvin, Mari and Permic might be parallel loans from Indo-Iranian, as the relationship between them is not entirely regular. In Mordvin, earlier \**u-a*-stems should develop *o* in the initial syllable, compare PU \**muna* ‘egg’ > Mordvin *mona*, PU \**kuma-* ‘face down’ > Mordvin *koma-* (Sammallahti 1988: 537–538). However, the borrowing has to be early in all the languages because Permic and Mordvin clearly show reflexes of the PU palatal sibilant, which could not result from later Iranian forms.

Katz (1985: 268–269; 2003: 215–216) is uncertain of the connection of the Finnic, Saami and the various Mari forms. There are several identical Mari entries in Bereczki 2013, two of which are relevant for this: s.v. *šū* 1 ‘Spreu, Klein’ = Mo *šuva*, *šiva*; ? Ud *šu*,

Ko *śu* < ? PP \**śu* (also Fi and Sa forms mentioned with caution); and *šū* 2 ‘Borste; Stachel’ = Ud, Ko *zu* (< PP \**zu*). Bereczki does not mention the Indo-European/Indo-Iranian loan etymology. It is probable that only the word *šū* 1 ‘Spreu, Klein’ is borrowed from the same source as Mordvin *śuva* and the Permic forms, and the rest of the words are of unknown origin.

There are a few other etymologies with such “Central Uralic” distribution as well, but on the other hand there are also many that have been preserved in the western extreme of the family too, pointing to replacement of cultural lexicon in both the centre and the periphery (it is well known that a great influx of Germanic, Baltic and Turkic loanwords have replaced many older words in these languages).

An Indo-Iranian etymology (Mari, Mordvin and Permic words)

**Ko, Ud *śum̐š* ‘Riemen’ < PP \**śum̐š* (Csúcs 2005: 381)**

← ? PII \**syúHman-* > OI *syūman-* ‘Band, Riemen, Gurt’, root *syū* < PII \**syuH* < PIE \**syeh<sub>1</sub>* (EWAia II s.v. *šiv̐y*; Werba 382 No. 473 s.v. *siv/syū*)

(KESK: 274; Joki 1973: 318. No. 151; Gamkrelidze & Ivanov 1984: 927; Katz 1985: 340; Rédei 1986c: 77; Lushnikova 1990: 43, 263; Helimski 1997b: 121; Zhivlov 2013: 220)

This is a convincing etymology, but the dating of the borrowing is problematic because of the Permic vocalism and the *ś* sibilant. Because of the Permic *u* corresponding to the Indo-Iranian *u*, the loan cannot be older than Proto-Permic, but the Permic sibilant points rather to Proto-Indo-Iranian or Indo-Aryan as the donor language, because in a late Iranian loan we would expect *h-* instead of *s-*. It has been argued by Helimski (1997b: 121) and Zhivlov 2013: 220) that the Permic word is borrowed from the Andronovo Aryan language. Helimski also notes that only the verb ‘to sew’ is attested in Iranian, but not the noun corresponding to Old Indic *syūman-*.

Joki (1973) and Rédei (1986c) argue convincingly that the *ś* is due to the substitution of Indo-Iranian *sy-*. Rédei considers the borrowing to be early Proto-Iranian or Middle Iranian and he does not take a stance on the problem of the retained *s*. Rédei’s categorization is followed by Csúcs (2005). Joki assumes that this is an Indo-Iranian or Proto-Iranian loanword, but this assumption seems to be an educated guess rather than based on actual evidence.

It is important to note here that the Sogdian forms CSogd. *šum*, MSogd. *šum* (denomin.) ‘to sew’, CSogd. *šumqy* ‘sewing’ that are connected to OI *syūman* and the verb *sīv-* by Rédei are not necessarily reflexes of the same Indo-European root. Cheung (2007 s.v. *čaiH*) lists them as possible cognates of the Iranian root *čaiH* ‘to sew’. This root is of an unclear origin, but according to Cheung perhaps attested as the latter part of the words Skt. *sū-cī-* ‘needle’, YAv. *sū-kā-* (f.), Oss. I *su-žin*, D *so-žinæ* ‘sewing’. Tentatively one could assume that the Sogdian form with *š-* is the source of the Permic words, either as a Wanderwort or as a direct borrowing from an Iranian form that is close to the Sogdian one. This could solve the problem with the Permic sibilant.

However, Cheung also mentions the Sogdian words under the entry of \**hyauH* ‘to sew’. Here he notes that the *š* in Sogdian and *ž* in Yigdha *žū y-*, M. *žiy-* ‘to sew’ and Wakhi *švan*

‘string, rope’ result from the RUKI rule in compounds. Even if the Permian word is borrowed from a true reflex of this Iranian root, there is no obstacle to deriving the Permian form from a RUKI word with \*š-. In front of the glide -y-, \*š could have as easily been substituted by \*ś. Especially in Permian where \*ś is phonetically more “hushed” (see Kochetov & Lobaonva 2007), this substitution would actually make sense. Although we have no parallel examples among the loans into Permian, this could solve the problem with the sibilant. Admittedly, the absence of the noun \*šyú(H)man- in Iranian is still problematic, but as the root existed there anyway and since -man- was a productive suffix in Proto-Iranian (see Schneider 2010), we have to assume that this noun could have existed in Iranian but it has not been attested.

An Indo-Iranian etymology

**Hu *szekér* (: *szekere-*) ‘wagon’**

< Pre-Hu/Pre-Ug \*śākārū-?

← **Indo-Iranian, cf. OI (Vedic) śakata- N, śakaṭī́ - F ‘Karren, Planwagen’ < ? \*śakarta-** (EWAia II: 601–602 s.v. śakaṭī́-)

(Korenchy 1972: 73–74; Joki 1973: 317; Katz 1985: 314–315; 2003: 251; Ligeti 1986: 151; Lushnikova 1990: 253–254)

see also Khanty *līkər* ‘(Zug)schlitten; Narte’

This is an old etymology but a very problematic one, and it is considered uncertain by sources like Ligeti (1986: 151). The Indo-Aryan word does not have a convincing Proto-Indo-European background which makes the idea that Hungarian *szekér* is from Indo-Iranian uncertain. In EWAia (II: 601–602 s.v. śakaṭī́) it is noted that the Indo-Aryan word might be from the root \*śak- ‘laufen’ which is attested also in Old Persian and Avestan.

Katz explains the Indic form from earlier \*śakarta-. This could explain the *r* found in Hungarian, but this explanation is not mentioned in EWAia, and it is not at all certain here that Indic -ṭ- reflects an old \*-rt- cluster. Katz notes that the -ṭ- could be a Prakritism, as this kind of development can be assumed some forms of Middle Indo-Aryan, from which the form śakata- could have spread into the language of the Rigveda.

The relationship between the Hungarian word and Khanty *līkər* is irregular (Khanty points to an old \*s or \*ś), and already Korenchy rejects the old idea that these two words could be derived from Proto-Ugric. As the word has no regular cognates in Iranian from which the Pre-Khanty form \*s- could be explained, the Indo-Iranian origin of the Khanty word remains even more uncertain.

? Indo-Iranian etymology

Fi *hata* ‘Keim, aufkeimende Saat’, *hadas* : *hata(h)alla* ‘aufkeimend’; *hatu* ‘Keimwurzel’, *hatajat* (cognates in Karelian and Estonian; SSA s.v. *hata*)

< ? Pre-Fi \*šatas

← PI ? \**dzaHta*- verbal adjective from the root \**dzanH*- ‘to be born, to grow’  
< PII \**žanH*- < PIE \**ǵenh<sub>1</sub>*- (PI \**zanH*- in Cheung’s 2007 reconstruction; Werba 1997: 288–289 No. 286 *jan<sup>i</sup>*)

See also \**sentü*-

(Koivulehto 1999a: 225 = 2016: 227)

This is one of Koivulehto’s Proto-Iranian loans based on the palatal criterion (the substitution of Iranian \**ts* and \**dz* by the Uralic affricate \*č). The etymology is semantically convincing, but there are difficulties in the sound substitution and also competing etymologies exist, and one cannot totally rule out the Germanic etymology that is mentioned by LÄGLOS (s.v. *hata*).

A point of interest is the exact sound substitution that has operated in this word and the other cases where PU \*š seems to correspond to PI \**ts*. Koivulehto presented the substitution as PI \**c* → PU \*č, but since Aikio (2015a) we know that this cannot be correct, as \*č develops to \**t* in all environments in Finnic. Kallio (unpublished) has thus reformulated Koivulehto’s rule in a way that suits the modern view of Finnic historical phonology: in the word-initial position the substitution was \**ts* → \*š, while medially Koivulehto’s original idea of \**ts* → \*č still applies (this is supported by a number of examples, such as \**mača* and \**počaw*).

However, this kind of substitution poses some intriguing questions. First of all, why were the PI affricates \**ts*, \**dz* not substituted by \*č also in word-initial position? Phonetically this would seem to be the most logical substitution, and there were certainly no phonotactic limitations that forbade \*č- at the beginning of the word in Uralic. Or were the Iranian and Finno-Ugric affricates so different that \*š was felt to be closer to the Iranian phoneme by speakers of Uralic? Proto-Iranian (as reconstructed by Mayrhofer 1989) possessed a number of affricates (\**ts*, \**dz*, \*č and \*dž, the last two resulting from the second palatalization), while Proto-(West-)Uralic only had two.

There is another possibility, although it is not easier to prove correct than the one presented above. In most daughter languages of Proto-Iranian, \**ts* and \**dz* are reflected as sibilants (e.g. Av *s* and *z*), and if there were sibilants \**s* and \**z* in an Iranian language which was in contact with Western Uralic, it would not be impossible to suppose \**z* being substituted by Uralic \*š, as the two Iranian sibilants could have had a wide scale of phonetic realizations, and as there was no sibilant \**z* in Uralic, it could be possible to assume this. Koivulehto has cautiously suggested a possibility for a similar substitution of PII \**s* (which, of course, continues the PIE \**s* and is of different origin than Iranian *s*, but the case is anyway similar) in the case of PU \**siŋka* ‘old’ ← PII \**sana-ka*-id. (see the entry). However, as there are examples of \**z* reflected by Uralic sibilant \**s*, the substitution \*š ← \**z* remains hypothetical.

The second syllable seems to point to the substitution U \*-*as* ← PI \*-*as*, which is attested in a number of convincing etymologies, but as the form *hadas* is scarcely attested

(mostly in inflected forms such as *hataalla*), it is possible that the word is an original *a*-stem, and -s is a derivational suffix here.

Regarding the alternative etymologies, LÄGLOS considers it possible that the Finnic word was borrowed from Germanic *\*sāda-* ‘seed’ (> ON *sað*), with the well-attested substitution PFi *\*š* ← (P)G *\*s*. As this etymology remains a notable possibility, one cannot consider the Indo-Iranian etymology certain.

? Indo-Iranian etymology

**Mansi North *sāt*, West *šot* ‘luck’ (< PMs *\*šāt*); Kh North *sot* (< ? PKh *\*sot/sot*) < ? POÜg ? *\*šat-* ‘luck’**

See PP *\*šud* for discussion and references

An Indo-Iranian etymology

**Finnish *herä-*, *heräjä-* ‘to awaken’, *herätä*, ‘to wake up’ (cognates in all Finnic languages)**

< ? Pre-Fi *\*šerä-*

← ? Pre-II or “Pre-I” *\*Hdžer-* (= PI *\*Hdžar-*) > Av perf. *jayara* ‘is awake’, OI *jarase* ‘he wakes up’; < PIE *h<sub>1</sub>ger*, > Greek *ἐγείρω* ‘I wake up’ (Cheung 2007: s.v. *\*Hgar-*; EWAia I: 574–575 s.v. *jar-*; LIV<sup>2</sup>: 245–246, s.v. *h<sub>1</sub>ger*; Werba 1997: 187–188 No. 71)

(Koivulehto 1999a: 221; 2001b: 249 = 2016: 223, 291)

Semantically this Indo-Iranian etymology is convincing, but various phonological problems make it uncertain. This word is one of the few suggested examples of the substitution of Indo-Iranian *\*dž* (from PIE *\*g-* before front-vowels). The Indo-Iranian verbal root has an Indo-European etymology (with a cognate in Greek *ἐγείρω* ‘I wake up’ etc. LIV<sup>2</sup>: 245–246, s.v. *h<sub>1</sub>ger*). Koivulehto’s assumption that Iranian *\*dž* can be substituted by Pre-Finnic *\*č* is totally plausible, but due to the new reconstruction of Aikio (2015a: 4–5) we have to reconstruct the Pre-Finnic word as *\*šerä-*, rather than *\*čerä-*, as Aikio has shown that *\*č* develops into *t* and not *h* in all positions in Finnic. If the etymology is correct, a substitution *\*š* ← *\*dž* then has to be assumed. The case is similar to the substitution of Iranian *\*dz* by Uralic *\*š*, and as there are some convincing examples of this substitution, one can accept that also the affricate *\*dž* can be reflected by *\*š* in loanwords.

Because of the *\*e*-vocalism, the borrowing should be very early, despite its limited distribution in Finnic only. Koivulehto mentions both Pre-Indo-Iranian and Pre-Iranian as possible donor languages, but as the idea of Pre-Iranian loanwords is not very convincing (see *\*teksä*), the Pre-Indo-Iranian etymology would be more probable.

However, both Kulonen (2010: 142, 216–217) and Junttila (ms.) have connected the Finnish word *herätä* and its cognates to a larger family of related Finnish words such as Fi *herkkä* ‘sensitive’ (= Estonian *erk* ‘watchful’), which is unlikely to be of PI origin due to

semantic reasons. This is an old idea that is found also in SKES (s.v. *herätä*) and EES (s.v. *erk*). This makes the background of the Finnic word more complicated. Referring to Koivulehto (1976: 248, footnote 1), Junttila also mentions that the Finnic word has been borrowed from another Indo-European word, from a reflex of the verbal root *\*ser-* ‘to keep an eye on, to protect’ (LIV<sup>2</sup>: 534, s.v. *\*ser-*).

Koivulehto’s (1976) original suggestion was that an unattested Germanic reflex of this Indo-European root could be the origin of the Finnic word, but LÄGLOS (I: 98–99, s.v. *herätä*) does not find this convincing. According to (LIV<sup>2</sup>: 534, s.v. *\*ser-*), the root *\*ser-* is reflected in Indo-Iranian (*Av ni-šarharatū* ‘has to watch over’) and Greek (Homeric ὄρονται ‘watch over’, 3PL.MED, Mycenaean *o-ro-me-no* ‘watching over’<sup>44</sup>) and possibly in Anatolian (Lydian *sareta-* ‘protector’, *katared-* ‘keeps watch’). In a Proto-Indo-European loan a substitution Uralic *\*š* ← Indo-European *\*s* might occur, as a similar substitution is known from Germanic loans (LÄGLOS I: 98–99), but it is unlikely that the Finnic word is borrowed from the Indo-Iranian reflex of PIE *\*ser-*, as the substitution *\*š* ← *\*s* does not occur in any convincing Indo-Iranian etymologies.

? Indo-Iranian etymology

### Mansi North *sirej*, West *širi* ‘sword’

< PMs *\*širyV*

← Iranian (Alanic) *\*ciry* ‘sword’ > Ossetic *cirǰ* ‘sword’; cf. Ossetic *cyrǰ*, *cirǰ* ‘sharp’, YAv *tiyra-* ‘sharp’ < PI *\*tiyra-* (Abaev 1958–1989 I: 325; Cheung 2002: 177)

(Munkácsi 1903; Korenchy 1972: 68–69, No. 24; Lushnikova 1990: 196)

The Mansi word is probably a late loan from an Ossetic-type Iranian language. The word for ‘sharp’ has a good Indo-Iranian etymology, but the noun for ‘sword’ is found only in Ossetic. Also the affricate from *\*ti-* points to an Ossetic type language. The Mansi *\*š* could regularly reflect an older *\*č*, as in inherited PU words, meaning that the word was probably borrowed before the sound change *\*č* > *\*š* in Mansi, with the substitution *\*č* ← *c*. Although there is no way to date the sound change absolutely, it cannot be very late as it predates Proto-Mansi, and that means that the change *\*ti* > *\*ci* in Alanic has to be relatively early too.

An Indo-Iranian etymology

<sup>44</sup> According to Beekes (2010: 1095–1096), the Greek word is rather from PIE *\*wer-*, although he admits that the absence of the reflexes of *ʷ-* (< PIE *\*w-*) in some forms is problematic.

**Mansi East *šišwə*, West *šešwə*, South *šisəŋ* ‘hare’ (< ? PMs *šVšwəŋ*)**

← PII \**ćasa-* > Av *saŋha-*, OI *śasá-* ‘hare’ < PIE \**kás-ó-* (EWAia II s.v. *śasá-*; NIL: 410, s.v. \**kás-*; Kölligan 2018: 2265)

(Katz 1985: 136; Blažek 1990: 40, Helimski 1997b: 121)

OI *śasá-* is assimilated from regular \**śasa-* (EWAia II: 624–625, s.v. *śasá-*; Kölligan 2018: 2265). The Avestan word is regular in this respect. This Indo-Iranian etymology by Blažek is an interesting possibility, but it suffers from unclear consonantal developments. Mansi *š* would go back to PM \**s*, which in turn reflects PU \**ś*, meaning that the borrowing has to be rather old and a late Iranian etymology is unlikely, as Pre-Mansi/PU \**ś* could reflect PII \**ć* but certainly not Proto-Iranian \**ts*. However, it remains unclear why there is a cluster *-šw-* in the Mansi word, as this cannot be explained from the Indo-Iranian forms. This is the weakest point of the etymology.

There is another word for ‘hare’ in Mansi that bears some resemblance to *šəšwé*, namely N *sáwér* ~ *sáur* ~ *saur* [sōwər]. It would be tempting to connect the two words and assume that they are parallel borrowings from somewhere. The Iranian etymology remains uncertain.

? Indo-Iranian etymology

**Fi *hyvä* ‘good’, Li *jevā* id. (cognates in all Finnic languages); SaN *savvit* ‘to heal a wound’ (cognates in all Saami languages except Akkala and Ter < PSa \**save-*); Mo E, M E *čiv*, M *čiva* ‘hospitable’**

< ? \**šiwä* ~ \**čiwä* ‘good, hospitable’

? ← PI \**tsiwá-* < PII \**ćiwa-* < ? PIE \**kíwó-*, cf. OI *śivá-* ‘auspicious, propitious, gracious, favourable, benign, kind, benevolent, friendly, dear’; deity name *Śivá-* (EWAia II :640, s.v. *śivá-*; Lubotsky 1988: 95)

(Koivulehto 2009: 85–87)

As can be seen from the two-fold reconstruction above, this etymology is a problematic one. It is not possible to derive the Finnic, Saami and Mordvin forms from a common proto-form regularly. Koivulehto rightly refuted the reconstruction \**šeŋä* found in FUV (99), the UEW and SSA (s.v. *hyvä*) as impossible, noting that the suggested Permic and Mari cognates do not belong here, but also the Finnic, Saami and Mordvin forms involve irregular correspondences. The irregular correspondence of Saami *s*, Finnic *h* and Mordvin *č*, *š* was mentioned by Koivulehto when he suggested the etymology, but he tried to explain this through supposing that the Saami verb was a very early Finnic loan. However, this explanation is not enough, since the Mordvin and Finnic forms are also irregular – the Finnic goes clearly back to pre-Finnic \**š*, while the Mordvin forms have to reflect pre-Mordvin \**č*. This anomaly was not such a big obstacle yet in Koivulehto’s era, as the reconstruction of PU \**č* and \**š* and their reflexes in Finnic and Mordvin were not totally

clear at the time, but Aikio (2015a: 4–5) has now convincingly shown that \*č is always reflected by *t* in Finnic, and all the inherited forms with *h* have to go back to \*š.

However, the Finnic and Saami forms might eventually be connected, if the Mordvin forms are left out, as both could regularly reflect \*š. Koivulehto tried to explain the Saami word as a possible Finnic loan, but his explanation was complicated because of the \*č ~ \*š problem. If the Saami word were a loan from Finnic, it would have to be very early because of Saami *s*; even in early loans, Saami *š* often corresponds to Finnic *h*, and also the word *hyvä* has been borrowed into some Saami languages very early as Kildin Saami *šuvvǎ*, Pite Saami *šuvies*, Lule Saami *šjuvēs* (SSA s.v. *hyvä*).

Also the vowel in the reconstruction \*šiwä ~ \*čiwä is problematic. An old \*i-ä stem should give *e* in both Mordvin languages (Bereczki 1988: 320), and the Mordvin forms point rather to Pre-Mordvin \*čewä. As is well known, Saami \*i and \*e merged early, so the Saami forms give no evidence here but could reflect either \*šewä or \*šiwä (or \*šüwä, as \*ü likewise merged with \*e and \*i in Saami). Although the Finnic forms are apparently as ambiguous as the Saami ones, they cannot reflect \*šewä but only \*šüwä or \*šiwä. Kallio (2016: 49) explains the Livonian cognates (Curonian) *jōv*, (Salats) *jua* as regular developments from earlier \*šüwä, through the intermediary stage \*iva. So we have \*šiwä or \*šüwä as the Finno-Saamic form and \*čewä as the preform of the Mordvin words. Neither of these fits the Indo-Iranian original very well.<sup>45</sup>

Kallio (p.c.) has hinted at a possibility that the Moksha word could be a loan from Erzya, and this would explain the \*č in both languages (in Erzya \*č is a regular reflex of \*š, but not in Moksha). This is of course possible, but hard to prove, and it would be good to find parallels for such Erzya loans in Moksha before taking this thought forward.

The Iranian etymology itself is also not an unproblematic one. As Koivulehto himself states, a cognate of OI *šiva* is not attested in Iranian. As the OI word has an IE etymology, and since the Old Iranian corpus is much smaller than the Old Indo-Aryan one, it might be possible to consider that Iranian has simply lost the cognate, as Koivulehto argues. However, it is always problematic to derive words from unattested forms (that is, forms that are not attested in a particular branch and cannot be reconstructed in a proto-language on the basis of daughter languages). What is more, it is unclear why Iranian \*ts should be reflected by \*š and not \*č in the word-initial position in Uralic – this was not a problem for Koivulehto’s idea, as he tried to derive both the Finnic and Mordvin forms from \*čiwä, but as is shown above, this cannot be correct. There are only a few examples of the substitution U \*š ← PI \*ts (always in the word-initial position), so it is dubious whether this is an actual substitution rule (see the entry on \*šukta for a more detailed analysis of this problem).

Koivulehto does not refer to Lubotsky (1988: 865), who has discussed this Indo-Iranian root and notes that the etymological relationship between OI *šiva-* and *šewa-* is impossible, as the latter reflects \*čeyH-wa, from the *seṭ* root \*čeyH-, and there is no trace of the laryngeal lengthening in \*čiwä-. As the etymology of the noun *šiva-* thus seems to be on a more shaky ground even within Indo-Aryan, it is more and more doubtful to assume that an exact cognate would have existed in Iranian. It has to be admitted, though, that Lubotsky’s notion of a PIE root \*keyH- has not been followed by all later etymologists,

---

<sup>45</sup> I am grateful to Santeri Junttila for discussing these phonological developments with me.

as Kroonen, for example, (Kroonen 2013: 227, s.v. \**hīwōn-*) derives OI *séva-* and Germanic \**hīwōn-* from PIE \**key-weh<sub>2</sub>-*.

All in all, there are too many unclear points in this etymology to support Koivulehto's explanation. The case remains unclear, and I would see the evidence weighing against the Iranian etymology instead of for it. It is also clear that the Mordvin and Finnic forms cannot be derived from the same source. Maybe this is a substrate borrowing, as many of the "Volgaic" substrate words have irregular consonant relations in Finnic and Mordvin, and \**š* is one of the typical sounds in this layer (Aikio 2015b: 44–47; Zhivlov 2015).

Not an Indo-Iranian etymology

**Fi *hoiva* 'care', *hoitaa* 'to take care of, to nourish' (cognates in all Finnic languages except Veps and Lude) < ? PFi \**hoji-***

< Pre-Fi \*? *šoji*/\**šaji*/\**šji*

← PI \**stsāyáH-* < PII \**ścaH-yaH-*, OI *chā-yā-* 'Schatten; Widerschein, Abbild', YAv *a-sāiia-* 'the one who has no shadow', MP *s'dk' /sāyag/ id.* (EWAia I: 559, s.v. *chāyā*; Lubotsky 2001a; Kobayashi 2004; Gotō 2013: 51)

(Koivulehto 2009: 80–82 = 2016: 410–412)

Koivulehto's recent etymology is quite convincing, even though some of his arguments have to be reweighed in the light of newer results in historical phonology. Koivulehto considers the Finnic stem *hoi-* primary here, arguing that both the verb *hoita-* and the noun *hoiva* are derived from this primary stem. This *hoi-* would reflect earlier \**šoji-*. Koivulehto considers also *šōji* and *čōji* as possible pre-forms, but it is now known, thanks to Aikio (2012a), that the Finnic long vowels are secondary. It is likewise known that Finnic *h* cannot reflect Pre-Finnic \**č* (Aikio 2015a: 4–5), so if the stem *hoi-* indeed reflects a Pre-Finnic stem, \**šoji* is the only option among the possibilities mentioned by Koivulehto. This also means that Koivulehto's idea that this loan manifests the substitution \**ō* ← \**ā* has to be rejected. However, also a pre-form \**šaji* would be possible: here the Finnic vowel development could be compared to that of Finnic *voi* < Proto-Uralic \**waji* 'butter' (Aikio 2012a: 236). And finally, it has to be added that due to the merger of \**a* and \**i* in Proto-Finnic, also a pre-form \**šji* would be possible.

All of this has relevance for the vowel substitution in this etymology. The Iranian word \**stsāyā-* or \**stsaHyaH-* had a long vowel (probably at least at the phonetic level, even if the laryngeal was still present in some form), and this is one of the relatively small number of loans where one can attempt to determine whether the short and long \**a* in Indo-Iranian had different substitutions. Here the problem is that numerous pre-forms for the Finnic word can be assumed. As the vowel substitution \**a* ← \**a*, \**ā* is the most frequent, it is safest to assume that the same substitution took place here, although both \**o* and \**i* remain possibilities.

The substitution of the stem vowel is of particular interest here, as if the noun \**šaji* can indeed be reconstructed, it points to an old \**i*-stem (= *e*-stem in traditional terminology), whereas a large part of the Indo-Iranian loans have been borrowed as Uralic \**a*-stems.

Koivulehto pays attention to the same issue, noting that *i*-stems often appear after heavy syllables. As \*šaji hardly had a heavy syllable according to modern Uralic reconstruction, and as the same can be said of the predecessor of Finnic *juoni*, which Koivulehto mentions as a substitution parallel (see the entry \**jini*). It seems that there is no correlation of syllable weight and the substitution of stem vowels. Here one can perhaps assume that the Uralic word was borrowed from a nominative form \**stsā́HiH-* (< PIE \**skéh1-yh2*; Gotō 2013: 51), and here Uralic \*-*i* (probably phonetically *ə*, as reconstructed by Kallio 2012a) would correspond to Indo-Iranian *i/y*. Phonetically this would be expected, but it would be essential to find parallels to this substitution. It is troubling that often we have no evidence on what case form was the donor form of the loans into Uralic, but here we could perhaps argue that it was the nominative. The other Uralic \*-*i*-stem loan mentioned above, \**jini*, cannot be explained similarly, as it is an *a*-stem in Indo-Iranian, and there are no case forms ending in *-i*.

The substitution of the word-initial consonant probably shows that the word was borrowed from Proto-Iranian, although here again Koivulehto's views have to be revisited based on newer evidence. As noted above, PFi *h* has to reflect earlier š, and this seems to be the substitution of the PI affricate *ts* in a number of loans (see \**šukta*), but here the Proto-Iranian word probably had \**sts* (from Proto-Indo-Iranian \**sć*), as Koivulehto also notes. There is no reason to assume that *sts* would have been substituted differently from \**ts* in the word-initial position.

Semantically the etymology is possible, although not very obvious. Based on the evidence of Indo-Iranian and the related languages, the primary meaning of the word on the Indo-European side was 'shadow' (compare Tocharian B *skiyō*, Greek σκιά 'shadow'). Koivulehto gives the meaning 'shelter' for some of the Iranian forms he cites, referring to Monier-Williams for the meaning of the Vedic word ('shadow; shelter, protection'). If we assume that the meaning 'shelter' can be reconstructed for Proto-Iranian already, the etymology can be considered as convincing, also semantically. Koivulehto also gives the meaning 'shadow; defence, shelter' for the New Persian word. The meaning is not found in Manichean Middle Persian dictionary of Durkin-Meisterernst (2004, s.v. *sy'g* 'shadow, pavilion'), nor in Gharib's (1995) Sogdian Dictionary (s.v. *sayag* 'shadow, canopy'), but in Christian Sogdian the meaning 'shelter' is found (Sims-Williams 2016, s.v. *sy'q*). In New Persian the meaning 'protection, auspices' indeed appears (Hayyim, s.v. *sayeh*), but it is uncertain whether this meaning can be reconstructed for earlier stages of Iranian. One can argue that based on the secondary meanings in Vedic, Sogdian and New Persian, the meaning 'shelter' can be reconstructed for Proto-Iranian as well, but this is far from certain.

According to Koivulehto, Finnic *suoja* 'shelter' and its cognates represent a later borrowing from the same Iranian source (see the entry below). If both of these etymologies are correct, they provide evidence that the word meant 'shelter' quite early in Iranian.

An Indo-Iranian etymology

**Fi *huhta*, *huhta*<sup>46</sup> ‘burned patch in slash-and-burn agriculture’ (cognates in Karelian, Ludic and Estonian); Mordvin E *čuvto*, M *šufta* ‘tree’**

< WU ? \**šukta*

← PI \**tsuxta-* > Av *upa-suxta* ‘burned’, root *saoč-* ‘to burn’, OAv *saocaiiaṭ* ‘inflames’ < PI \**tsauč-* (Cheung 2007: 338–340, \**sauč-*; EWAia II: 655–656, s.v. *SOC-*; Werba 1997: 243–244, No. 196)

(Koivulehto 1999a: 225–226; 2001b: 256–257)

Koivulehto (1991: 32) first had the idea that this Finno-Mordvinic word is a loan from Baltic \**šukta*, but later he changed his mind and argued for a Proto-Iranian origin. Although reflexes of the root \**k* *ewk-* ‘burn’ is attested widely in Indo-European, this particular formation is not present in Baltic, but in Iranian it is widely attested. The etymology is thus plausible: there is no competing etymology for this word, and other agricultural terms have also been borrowed from Iranian (maybe roughly simultaneously, as \**očra* is also a Proto-Iranian borrowing). The Baltic etymology cannot be totally ruled out for certain, as there are other examples of Indo-European loans where the origin of the word has not been preserved in the branch in question. The Baltic etymology of the Uralic word is also mentioned by EWAia, which was published before Koivulehto’s Iranian etymology.

A problem with the Iranian explanation is the fact that contrary to what Koivulehto believed, the Finnic word has to reflect an earlier \**šukta*, not \**čukta*. The old assumption was that both \**š* and \**č* can yield Finnic *h*, at least word-initially, but Aikio (2015a: 4–5) has shown this assumption to be incorrect: \**č* yields regularly Finnic *t* in all positions, whereas \**š* always develops to *h*. Also the Mordvin consonantism can be explained from \**š*, as this develops secondarily to *č* in Erzya and is retained in Moksha.

This means that if the word indeed is borrowed from Proto-Iranian, the substitution the word-initial position was \**š* ← \**ts*. While there is nothing wrong in this idea *per se*, it remains unexplained why in medial position the substitution was \**č* ← \**ts* (this is supported by a number of examples), and why this same substitution could not be applied in the word-initial position as well. \**š* and \**č* were clearly different phonemes in this position in Proto-Uralic: a number of examples show that both were allowed in all positions, and there is no reason to assume that the situation had changed in Proto-West Uralic. If the word was borrowed from Baltic, \**š* ← Baltic \**š* would be a phonetically expected substitution.

The word-internal consonant cluster \**-kt-* has to result from the substitution of Iranian \**xt*, as \**k* became a spirant in such clusters in Proto-Iranian (cf. the Avestan form *.suxta-*). Koivulehto’s Proto-Iranian reconstruction \**tsukta-* is thus slightly erroneous. Of course one cannot know for sure when the spirantization happened exactly, but at least before Proto-Iranian split up the change had to be complete.

A further problem lies in the vocalism. Koivulehto (2001b: 57) states a bit ambiguously that the Mordvin irregular vocalism “itself points to borrowing”. A regular development for

---

<sup>46</sup> The long *uu* in eastern dialects of Finnish and in Karelian is secondary (T. Itkonen 1987).

\**u* would be Mordvin *o*, giving a form \*\*čovto etc. The fact that the Finnic and Mordvin cognates are irregular points to the scenario that the word is a parallel borrowing, borrowed into Mordvin at a time when the change from \**u* to \**o* was no longer productive (or that the word was borrowed with a different vowel that then developed to *u* in Mordvin). There are not many examples that point to relatively late, separate borrowings in Mordvin and Finnic.

It is interesting to compare the situation with Iranian loans to other irregular cognates of Finnic and Mordvin. Aikio (2015b: 44–46) has pointed out that in Finnic, Mordvin and Mari and more rarely in Saami there are many irregular cognates, which are probably borrowed from an unknown substrate language after the westernmost branches of Uralic had already divergent from the proto-languages (examples include Fi *lehmä* ‘cow’ ~ Mo E *lišme* ‘horse’ < ? \**lešmä*, Fi *vehnä* ‘wheat’ ~ Mo E *viš*, Mari E *wiste* ‘spelt’ < ? \**wešnä*). It would be tempting to assume that the Proto-Iranian loanwords were borrowed roughly at the same time as these substrate words, into the already differentiated West Uralic dialects. This could explain the irregularities in vocalism.

Semantically this Iranian etymology shows an interesting development in Mordvin, as there the more original meaning of ‘burned’ has developed into a more general meaning of ‘tree’. The UEW even reconstructs the meaning of the Finno-Mordvin word as ‘a kind of tree’, but here it is more natural to suppose that the word was borrowed into Pre-Finnic and Pre-Mordvin in a meaning akin to ‘burn’, and the Mordvin simply shows a later semantic development. Koivulehto also mentions a parallel to this in Estonian *kask* ‘birch’, which has developed from Proto-Finnic *kaski* ‘slash-and-burn agriculture’ in a very similar way.

An Indo-Iranian etymology

**Ud, Ko šud ‘luck’ < PP \*šud** (Csúcs 2005: 384)

← Old Iranian \**šāti-* or Middle Iranian \**šād/šāt-* < PI \**čyaHta-*, root \**čyaH-*; OAV *šiiāta-*, *šāta-* ‘luck’, root *šiiā-* ‘to enjoy’, MP *šād* ‘glad, joyful’, Sogd *šāt id.*, Oss *æncad* ‘quietness, tranquility’ (< PI \**ham-čyāta*) (Cheung 2002: 160; 2007: , s.v. \**čīaH*)

(Joki 1973: 320–321; Katz 1985: 194; Rédei 1986c: 78; Honti 1999: 129; Csúcs 2005: 384)

The word has to be a relatively old borrowing into Permic because of the *u* vocalism, which can point to earlier \**a*–*a*. Rédei notes that *d* as the substitution for *t* is irregular here, and mentions also the possibility that the word was borrowed from a form which already had *d* (< \**t*), cf. MP *šād*. It is unlikely that the Permic word is borrowed from a language closely related to Middle Persian, but it could be acquired from an Eastern Iranian form which went through a similar development, such as a form akin to Sogdian *šāt*, if we assume that *d* developed there early enough. As the Permic word needs to be borrowed before the vowel change \**a* > \**u* that led into Proto-Permic, the word could really be a Middle Iranian loanword, as Rédei assumes. If the word was borrowed around the beginning of the Common Era, it would be early enough for the word to be able to

participate in the vowel changes. On the other hand, *d* between vowels regularly disappears in Permic, so the word cannot be as early loan as Permic *gu-*, for example.

The correspondence Ossetic *c* ~ Avestan *š*, *š*' is due to development of Proto-Iranian *čy-*. The Permic word is borrowed from a word which had already developed *š*. The affricate was retained in some East Iranian languages, at least in Khotanese *tsāta-*: in Ossetic *čy* it developed into *c*, meaning that the Permic word cannot be borrowed from a form of Alanic. A similar prefixed form as in Ossetic *\*ham-čyaH-* is reflected also by Sogdian *'nc(')y* 'to rest, dwell, cease'.

In addition to the Permic words, similar words are also found in Khanty and Mansi: West Mansi *šot*, North *sāt* (< PMs *\*šāt*), Kh North *sot* (< ? PKh *\*sot/sot*) < ? POUg ? *\*šat*- 'luck'. There is no consensus on whether the Ob-Ugric words are independent loanwords from Iranian or borrowed from Permic (in no case can the Permic, Khanty and Mansi words be derived from an unitary PU form). Rédei mentions both possibilities and does not take a clear stance, Joki (1973) considers the Permic and Mansi words to be cognates but notes that the Khanty word (which is only found in North Khanty) is possibly borrowed from Mansi. Honti (1999: 129) mentions this loan etymology briefly when discussing the reconstruction POUg *\*š* and its background, but he does not discuss the etymology in more detail. It is also worth noting that Honti presents no precise origin for the word, he simply states that it is a loanword from "Iranian *\*šāt*". Honti assumes that the word had the phoneme *\*š* already in Proto-Ob-Ugric, and we have to assume a substitution Iranian *\*š* → POUg *\*š* > Khanty *s*, Mansi *š*. The Iranian etymology of the Ob-Ugric words is not mentioned by Korenchy (1972) or Helimski (1997b).

The whole issue of reconstructing Proto-Ob-Ugric *\*š* is based on very limited evidence from Mansi (in Khanty POUg *\*š* and *\*s* would have merged), and as this word could equally well have been borrowed from Permic as Iranian, it remains doubtful how much evidence this etymology can have for Proto-Ob-Ugric historical phonology. Honti (1982: 106, No. 608) still reconstructs POUg *\*sōt* > PMs *\*šāt*, PKh *\*sot/sot*. Zhivlov (2006: 82) only mentions the Proto-Mansi word in his discussion on the long vowels in Proto-Mansi, but he does not comment on its status as a possible inheritance from Proto-Ob-Ugric. It is thus uncertain whether the Ob-Ugric words are true cognates. The Khanty word has a very limited distribution in northern Khanty (Obdorsk and Kazym dialects), making it likely that the word is borrowed from Mansi, like Joki has suggested. The Mansi word is attested more widely, and the relationship of the forms with *š* and *s* appear similar as with the inherited Proto-Mansi words. The Mansi vowel (Proto-Mansi *\*ā*) is hard to explain from the Permic word. We can assume here that the Permic and Mansi words are indeed independent borrowings from Iranian: in Pre-Permic, the Iranian vowel *\*ā* has been substituted by Permic *a*, while in Mansi it has been substituted by *\*ā*. Formally the Proto-Mansi word could also reflect a Pre-Mansi/Uralic *\*a-i*-stem. The Northern Khanty word is then borrowed from Mansi, with North Khanty *o* substituting for the *o* vocalism of the Northern Mansi words.

Like the dating of the borrowing into Permic, the dating of the Mansi loan is likewise difficult. Here it is even more troubling that the Iranian form *\*šyāti-* or later *\*šāt* could produce either Pre-Mansi *\*šāti* or later Proto-Mansi *\*šāt*. In theory, Mansi *š* could also reflect earlier *č*, so the word can also be borrowed from an Iranian variety which preserved the affricate *č*. The question of the emergence of Mansi as a separate branch is very poorly understood, making the dating of the borrowing even more difficult.

An Indo-Iranian etymology (Mansi and Permic separate borrowings)

**Mansi South *šerkeš*, West *šarkš*, East *šarkš*, ‘eagle’**

< PMs \*šārkVśV

← ? Iranian (Alanic) \**cargas*, compare Ossetic *cærgæes* ‘eagle’, Chwarezmian *krkys* [\**kargīs*] id., Av *kahrkāsa-* ‘vulture’, Sogdian *kiryī*, *karyə* ‘small falcon’, *crks* /čarkas/ (Cheung 2002: 176; Gharib 1995 s.v. *crks*)

(Munkácsi 1903; Jacobsohn 1922: 229; Korenchy 1972: 69 No. 25; Teubner 1974: 304; Lushnikova 1990: 196; Helimski 1997b: 120)

Helimski mentions that this is one of the clearly Ossetic-type loans in Ob-Ugric. Cheung (2002: 176) considers the Ossetic word and its suggested cognates Wanderwörter. The Avestan word has sometimes been analysed as ‘chicken-eater’ (with *-āsa-* reflecting PII \**aśa-* ‘to eat’ which does not exist in Iranian as a separate root; see \**aśnas*), but this is probably folk etymology (de Vaan 2000; Lubotsky 2001b: 315). If the Iranian etymology is correct, the Alanian form with affricate *c-* would be the most probable source. Also Helimski (1997b: 120) is supportive of the Mansi words’ Ossetic origin. As Mansi has other Alanic loans, it is more probable to assume such an origin for this word as well, even though theoretically the Mansi and Iranian words could be parallel borrowings from some unknown source.

An Indo-Iranian etymology

**Khanty South *täyər*, North *äyər*, East *əyər* etc. ‘Panzerhemd’**

< ? PKh \**äiyər*

← “Middle Iranian” \**zyar-*, compare Ossetic *zyar* ‘metal, PL armour’, Pashto *zyə’ra*, *zyər* M ‘armour, chain mail’ [Korenchy: “Pashto *zyarah*”], Chwarezmian *zyr<sup>u</sup>k* (Morgenstierne 2003: 101, s.v. *zyə’ra*)

(Munkácsi 1901: 630; Jacobsohn 1922: 228; Katz 1985: 346; Korenchy 1972: 72–73 No. 31; Joki 1973: 323; Helimski 1997b: 120)

This is a good etymology, although the background of the Iranian word is not clear. The etymology appears already in the works of Munkácsi and Jacobsohn. The word has to be a relatively early borrowing, as it was acquired before the Proto-Khanty ~ Proto-Ugric sound change \**s* > \**ś*. Mainly because of this feature Korenchy states that the word goes back to Proto-Ob-Ugric, and this is actually a very good example of a loan whose age cannot be determined on the basis of its distribution. Although Korenchy considers this a Proto-Ob-Ugric word, it is obvious that this word was acquired into a language whose sibilant system was still identical with Proto-Uralic (in Proto-Ob-Ugric the change \**s* > \**θ* would

have already taken place, so a borrowing into that particular stage would have resulted in modern Khanty forms *\*\*săyr* etc.).

As mentioned above, not much is known of the background of the Iranian words. EWAia (III: 205, s.v. *jagara-*) mentions that the Sanskrit word *jagara-* ‘armour’ is an Iranian loan (see also Baghibidi 2005) and compares it to the Ossetic and Pashto words, but no information on the background of the Iranian word is given. Also Morgenstierne (2003) gives the same explanation, rightly noting that the borrowing into Indic has to be very early – the Indic affricate *j-* could here be a substitution of a similar Iranian sound, which already before Old Iranian times became *\*z*.

The early borrowing from Iranian into Indic (as well as into Khanty/Uralic) and the word’s attestation in Chwarezmian mean that the Iranian word is an old one and probably can be derived from Proto-Iranian. Regarding the origin of the Khanty word, an early Alanic (Pre-Ossetic) source is probably the best possibility, although Jacobsohn (1922: 228) was critical of the Ossetic-type origin of the word. Munkácsi did not even mention the Ossetic word but compared the Khanty word only with Pashto. Joki notes that the word is borrowed from an “Eastern Iranian language of Siberia”, but this does not tell much, and as the borrowing is rather early, it is in no way clear that the word has been borrowed somewhere in Siberia (a more western area of contact is also possible). Jacobsohn stated that the word is probably not from Alanic/Ossetic but from some other variety of Eastern Iranian. Korenchy’s idea that the chronological stage of the donor form is “Middle Iranian” can hardly be correct, as the Middle Iranian period of the attested Iranian languages started in the last half of the first millennium, and the sibilant change of Proto-Khanty must have happened considerably earlier because it is shared with the other “Ugric” languages, as well as with Samoyedic.

Honti (1997: 35–37) has tried to argue that the word could be a later borrowing from a Persian-type language, where PIE *\*ǵ* was reflected by *\*θ*, and the Proto-Khanty word-initial consonant (? *\*θ*) would be a simple substitute of this. However, there is absolutely no reason to support Honti’s scenario, as there is no proof of this kind of development in the Iranian languages of the steppe, which were not closely related to or affected by Persian. The Iranian word is also attested from Eastern Iranian varieties only, although the word’s existence in more western languages, such as Persian, is also possible, if the word indeed goes back to Proto-Iranian. Honti’s argumentation stems from the attitude that it would be unlikely for an early Iranian borrowing to be present only in Khanty, but as stated above, inherited words do get lost anyway, and there are numerous examples of an early borrowing being preserved in one Uralic branch only, so it is best to reject his objection. The modern Khanty forms probably continue an earlier *\*s*, which was a substitute for Iranian *\*z*.

#### An Indo-Iranian etymology

**Khanty East *likər, ikər* ‘(Zug)schlitten; Narte’ < ? PKh \**likər*  
< ? Pre-Kh/Pre-OUg ? \**säkärä***

← **Indo-Iranian, cf. OI (Vedic) *śakaṭa-* (neutr), *śakaṭī-* (feminine) ‘Karren, Planwagen’ < ? \**śakarta-*, *śakartī-* (EWAia II: 601–602 s.v. *śakaṭī-*)**

As is the case with the Hungarian word *szekér*, connected to the Khanty word in earlier research, the Indo-Iranian origin is very uncertain. See the entry \**szekér* for more details on the problematic background of the Indo-Iranian word.

The Khanty word should be a specifically Iranian borrowing, from a form with \*s-, as Proto-Khanty \**l* could not be explained from Proto-Indo-Iranian \**ć*. As is the case with the Khanty word for ‘Panzerhemd’, Honti (1997: 35–37) has argued that the word could be borrowed from a Persian-type language, where PIE \**ǵ-* was reflected by \**θ*, this sound being substituted by Proto-Ob-Ugric/Pre-Khanty \**θ-*. The idea is not convincing, as there is no other evidence for Persian-type reflexes of the PIE palatovelars in the Iranian languages of the steppe, but all other loanword evidence points to \*s and \*z as in the majority of the Iranian languages.

Although the etymology for the Khanty word is uncertain for the above reasons, semantically the idea of deriving ‘sledge’ from a word for ‘chariot’ would be totally plausible.

? Indo-Iranian etymology

**Hu *titok* ‘secret’, der. *titkos, titkolni, ? tilt* ‘to forbid’; Ms North *tūjt-*, West *tujt-, tujt-, tunt-* (< PMS \**tuj-*)**

? < PUg \**taja* (?)

← PI \**tāya-* ‘heimlich; Diebstahl, Dieb’ > Yav *tāia-* ‘Diebstahl’, OAv, YAv *taiia-* (adj.) ‘(diebisch) verstorhen, heimlich; Dieb; Diebstahl’; cf. OI (Vedic) *stāyúh-* ‘Dieb, Taschendieb’, *stāyát* ‘verstorhen, heimlich’, root (s)*tā* < PII \*(s)*taH* < PIE \*(s)*teh<sub>2</sub>* ‘to steal’ (EWAia II: 759, s.v. *stāyát*; Werba 1997: 442, No. 590)

(Korenchy 1972: 72, No. 30; Katz 1985: 288–289; 2003: 230–231; UEW s.v. *taj<sub>3</sub>-tt<sub>3</sub>*; WOT: 1290)

The Iranian etymology of these Ugric words is complicated by the numerous Hungarian words connected here by the UEW. Some of the forms mentioned by Korenchy have alternative etymologies, and listing all kinds of Hungarian word forms under this etymology makes the whole etymology look poor. In principle, the Indo-Iranian etymology for this Ugric word is possible. Semantically there are no obstacles, and the Mansi word could reflect an old \**a-*stem, like Korenchy argues. The Uralic diphthongs have various reflexes in Hungarian, so it seems possible that *titok* comes from earlier \**taj-* and can be considered a cognate of the Mansi word. Also *tulaj* (obs.) ‘allein’ and *tulajdon* ‘eigen; Eigentum, Besitz; Eigenschaft; (veralt.) der Nächste; nur, allein’ are connected here by the UEW, but it is impossible to connect these etymologically to *titok* and *tilt*, and they have to be of different origin.

WOT (1234, s.v. *tilt*) notes that also a Turkic origin (West Old Turkic *\*tiyil-*, a cognate to East Old Turkic *tidil-* ‘to be obstructed’) has been suggested for the Hungarian word family *tilt* ‘to forbid’, *tilos* ‘forbidden’, but WOT considers this unconvincing due to “chronological and semantic difficulties”. It is more difficult to understand what the chronological difficulties are, but semantically ‘to be obstructed’ and ‘to forbid’ are not very close, although the connection does not seem impossible to me. WOT also notes that the idea that Mansi *tūjt-* is a cognate to the Hungarian words is implausible. However, that Hungarian *titok* ‘secret’ and Mansi *tūjt-* are cognate is supported by WOT (1050, 1290), which follows the UEW here in deriving these words from *\*taja-*. The Iranian etymology of this Ugric word is supported by WOT.

The Hungarian noun *tolvaj* ‘thief’ which is connected to PUG *\*taj-* by UEW is probably an early borrowing from Proto-Mansi *\*tūlmak* ‘thief’, which is derived from *\*tūl-* ‘to steal’ < PU *\*sala-* id. (Barczy 1943, s.v. *tolvaj*; Helimski 1982: 54–55).

? Indo-Iranian etymology

## Fi *taivas* ‘sky, heaven’ (cognates in all Finnic languages)

### < Pre-Fi *\*tajwas*

← PI(I) *\*daywá-* > OAv *daeuiua-* ‘pre-Zoroastrian god’, YAv *daēuiua-* ‘demon’, OI *devá-* ‘heavenly, divine; god’; cf also the later Indic vrddhi formation *dāiva-* ‘belonging to or coming from heaven’; < PIE *\*deywo-*, root *\*dyew-* (EWAia I: 742–743, s.v. *devá-*; Gotō 2013: 53–54)

(Joki 1973: 323; Rédei 1986c: 60; Koivulehto 1999a: 228, 232 = 2016: 230, 234; SSA s.v. *taivas*; Katz 1985: 90; 2003: 81)

This is a good etymology, first suggested by Diefenbach (1851: 607) and almost universally accepted, as both the semantics and phonology are obvious. The word reflects the same mythological influence from the Indo-Iranians as the loans *\*juma*, *\*sampas* etc. The Indo-Iranian word is from a root that is widely attested in Indo-European, with also formal parallels that continue the PIE vrddhi formation *\*deywo-* in many languages, such as Lithuanian *dievas* and Germanic *\*teiwaz* ‘name of a god’.

It is also an old idea of Thomsen (1869: 73; Kalima 1936b; 1950) to derive the word from Baltic (from *\*deiwas* > Lith *dievas* ‘god’), but the Indo-Iranian vowels fit better the Finnic ones, and it is much more likely that this is an Indo-Iranian borrowing (Larsson [1984]; Koivulehto [1999b: 80]). Kalima (1936b; 1950) argued that the Finnic diphthong *ai* could result from earlier *\*ej*, but this view is impossible (see now Kallio 2014 for the development of Finnic diphthongs).

Katz argued that the question of whether the word is borrowed from Indo-Iranian or Baltic is “formal ... nicht lösbar”, but an Indo-Iranian origin is more likely because of the Baltic loanwords “semantisch schlichten Sphären angehören” – this is a “loanword layer argument” in Junttila’s terminology. While I agree with Katz in that the Indo-Iranian etymology is more convincing, his argument is not very strong, as words relating to so many semantic spheres have been borrowed from Baltic that *taivas* would easily fit in. SSA

remarks that semantically the Indo-Iranian etymology is better than the Baltic one, as in Baltic no meanings referring to ‘heaven’ are attested. Nevertheless, EES mentions that the Baltic origin would be more likely, although detailed arguments in favour of this decision are not given.<sup>47</sup>

Schmid (1979: 268) has argued that the Iranian origin is unlikely because of the negative semantics that are connected to the Avestan word *daēvō* and its later Iranian cognates, but this is not a very strong argument, as these negative semantics are the result of the Zoroastrian religion, and there is no reason to suppose that in Proto-Iranian the word already had acquired a pejorative meaning ‘demons’ or the like.

The ending *-as* probably reflects the Indo-Iranian thematic masculine ending. As there are no variants without *-as* of this word in Finnic, this is one of the more convincing cases of the ending being substituted. According to Helinski (2000), these loans belong to a later (Finnic-Permic) layer of loanwords.

According to Koivulehto (2003), the Finnic verb *toivo-* ‘to wish’ is borrowed from a reflex of the same Indo-Iranian root (see the entry).

An Indo-Iranian etymology

### **Fi *takra* ‘piece of meat’ (cognates in Karelian, Ludic and Veps) < PFI *\*takra***

← PII/PI *\*daHtrá-* > OI *dātrá-* ‘gift’, Av *dāθra-* ‘Gabe, Geschenk, Almosen; festgesetzter Lohn, Sold’, root PII *\*daH* ‘zuteilen’ < PIE *\*deh<sub>2</sub>(y)-* (Hintze 2000: 33–37; Werba 1997: 294 No. 297; NIL: 62–63, s.v. *\*deh<sub>3</sub>-*)

(Koivulehto 1999a: 232 = 2016: 234; SSA s.v. *takra*)

The etymology is convincing. The semantic development has parallels in other Indo-European languages listed by Koivulehto. The substitution of the cluster *\*-tr-* is argued convincingly by Koivulehto: as there was no *\*-tr-*, the existing cluster *\*-kr-* was the expected substitute. Koivulehto also supports this by presenting similar cases of the substitutions of a stop-liquid cluster from the well-researched layers of early Germanic borrowings in Finnic, such as *\*nekla* ‘needle’ (> Ka *niekla*, Fi *neula* id.) ← G *\*nēdla-*. It is interesting that in some other loans the substitution has been different, at least in Permic *burd* < Pre-Permic *\*pertä* ← *\*pettro-* ‘wing’ (see the entry), but these words can reflect different substitution patterns in different parts of the Uralic language family.

The vowel substitution is expected, as most of the loans show Uralic *\*a* from PII *\*a*. Also the Finnic *a*-stem is expected, as most of the Indo-Iranian *\*a*-stems are reflected like this.

The Vedic word has a short vowel, but also a long-vowel form *dātra* is attested in both Vedic and Avestan, although its meaning in the Rigveda is unclear (possibly ‘Gabe’

<sup>47</sup> This question has a long research history, which is referred to by Joki (1973). Kalima (1936; 1950) has defended the Baltic origin both by assuming that the origin ‘heaven’ might have been present in Baltic earlier and by considering the origin of the Finnic variation of *\*ei* and *\*ai* diphthongs unclear, assuming that Finnish *taivas* could continue earlier *\*ei*. Koivulehto rightly states that this view is now outdated, as Finnish *\*ei* can reflect earlier *\*ai*, but not vice versa (see now also Kallio 2014 and 2018 for a detailed discussion of the development of these Finnic diphthongs).

according to NIL: 62–63). NIL reconstruct the pre-form of the *\*datra-* words as *\*dh<sub>3</sub>-tRo-* (> *\*datra*) and *\*doh<sub>3</sub>-tRo-*, although it is noted that the etymologically unrelated form *dātrá-* ‘Zugeteiltes, Anteil’ (root *\*deh<sub>2</sub>(y)-* ‘to share’) could also have influenced the vowel length of the Vedic word (Anm. 80). NIL also mentions a Vedic form *dāttra-* N ‘Geschenk’.

It is difficult to date the borrowing: either PII, PI or even some later Iranian language could come into question. The limited distribution does not help much.

An Indo-Iranian etymology

**Fi *talas* ‘shelter’ (cognates in Estonian and Livonian); *talo* ‘house’ (derivation, with cognates in all Finnic languages); Ud *tīliš* ‘hut’; Mansi *tul* ‘pool; shed’**

< ? PU *\*tala-sV*

← PII *\*tala-m* > OI *tala-* ‘surface, level (Fläche, Ebene)’ (EWAia I: 637)

(UEW s.v. *tala*; SSA s.v. *talo*; Korenchy 1972: 74–75; SSA s.v. *talas*)

This is an old but problematic etymology. Both the etymological affinity of the Uralic words discussed here and the various Indo-European loan etymologies are uncertain. Sammallahti (1988: 550) considers Finnic *talas*, Udmurt *tīliš* ‘a hut’ and Mansi *tul* ‘pool; shed’ as regular cognates, deriving them from PU *\*tala*. The same comparison is found also in the UEW, although there the Mansi word is connected with a question mark. Recently Aikio (2015b: 56) has considered the Finnic and Permic words to be regular cognates.

LÄGLOS (III: 268–9, s.v. *talas*) notes that Koivulehto (in an unpublished handout) has considered *talas* as a borrowing from early Proto-Germanic *\*stāla-s* (> ON *stóll* ‘chair’), and that Hofstra (1985) supposes the word was borrowed from Germanic (cf. *\*stalla-z* (> Old Norse *stallr* ‘stand; (pagan) altar; stable, manger’). LÄGLOS considers both etymologies plausible, and states that the Finnic word is possibly (but not certainly) a loanword from Germanic. Also a Baltic origin has been assumed (Proto-Baltic *\*talas* > Lith *pā-talas* ‘Bed’) but according to LÄGLOS this is semantically less suitable than the Germanic and Indo-Iranian words. Both possibilities, as well as the Indo-Iranian etymology, are mentioned by SSA.

For a Uralic word with reflexes in Permic and possibly in Mansi, the Indo-Iranian source would be more credible, as no Germanic or Baltic loans have such a distribution within Uralic. However, the Indo-Iranian etymology suffers from various problems. Finnic *-as* that is often found in loanwords usually corresponds to Indo-Iranian and Baltic *-as* or Germanic *\*az* (< PIE *\*-os*), but in this case it has to be a later suffix, because the Indo-Iranian word is a neuter and the Finnic suffix cannot be explained from that. It is uncertain what the relationship of the Udmurt *-s* with the Finnic *-as* is here.

The etymology of the Indo-Iranian word is uncertain according to EWAia. Although it is assumed in IEW (1061) that the word has cognates in various Indo-European branches, such as Slavic *\*tbo* ‘ground’, Baltic (Lith *pā-talas*, Latvian *patali* [pl.] ‘bed’, Old Prussian *talus* ‘floor’), Germanic (German *Diele* ‘floorboard’) and Latin (*tellūs* ‘earth’), EWAia is less certain of the Indo-European etymology. Derksen (2015: 465) reconstructs the Indo-

European root as *\*tlh<sub>2</sub>-* and connects the Baltic words *\*patalas* and *\*tiles* ‘bottom of a barge, flooring’ to the Slavic and Germanic words, but does not mention the Indo-Aryan word.

The *\*l* that is found in the Uralic forms make the etymology improbable, as most of the loans reflect the results of the Proto-Indo-Iranian sound change *\*l > \*r*. The possible Indo-European etymology of the Uralic words remains open, but it is not very likely that they are borrowed from Indo-Iranian.

Not an Indo-Iranian etymology

**Hu *tál* ‘Schüssel’; Mansi West *tūl, tul*, North *tūlāna* (deriv.) ‘Speiseschüssel aus Holz...’ (also used as a derivational suffix; Riese 2001: 118) (< PMS *\*tūl*)**

< ? PUG *\*tala*

← ? “Middle Iranian” *\*tāl-*, compare Modern Persian *tāl* ‘a bell-cymbal...’

(Munkácsi 1901: 586–587; Korenchy 1972: 75–76, No. 34.; Helimski 1997b: 121)

The Hungarian and Mansi words are regular cognates. However, an Iranian origin seems unlikely because of the scarce attestation of the Iranian word. The etymology is considered uncertain by Helimski (1997b: 121). Also, a Middle Iranian origin for an inherited word in Mansi and Hungarian would be an anachronism, for although it is difficult to determine the exact beginning of the “Middle Iranian” period in the unattested Iranian languages of the steppe, it can hardly have begun before 500 BC, and almost certainly after that, and Hungarian and Mansi must have separated long before that date (the views that date the split of “Proto-Ugric” to 500 BC are not very well-founded, as has been recently noted by Aikio 2014d: 30). Therefore, this is either a Wanderwort or a chance resemblance.

Not an Iranian etymology

**Fi *tarna, taarna* ‘sedge (*Cladium mariscus*)’; Ko *turun* ‘grass, hay’, Ud *turiṅ id.* (< PP *\*turiṅ*; Csúcs 2005: 388)**

< PU *\*tarna* or *\*tīrna* ‘grass’ (UEW s.v. *tarna*; Sammallahti 1988: 554; Zhivlov 2014: 143; Aikio 2015b: 65)

← PII or PI *\*tīna-* ‘grass, blade of grass, herb’, Khot *ttarra-* ‘grass’, MP <tlk> /*tarrag*/ ‘Schnittlauch’ (EWAia I: 664, s.v. *\*tīna-*)

(Setälä 1928: 305–306; FUV: 138; KESK: 287; Joki 1973: 325; Katz 1985: 310; 2003: 121; Rédei 1986c: 60; Lushnikova 1990: 266; SSA s.v. *tarna*)

This is an old and convincing etymology. The Finnic and Permic words are regular cognates, so the borrowing is old. The original could be either PII or PI. There are no known cognates outside of Finnic and Permic. Because no other ancient Uralic term for ‘grass’ is known (the Ugric languages reflect *\*pūmV*, which could in principle be ancient, if

cognates in other branches have been lost), it is hard to judge whether the cognate of these words has simply been lost in other branches, or whether this is a more recent (“Finnic-Permic”) loan. Borrowing of a word denoting ‘grass’ is perhaps connected to the influence of the pastoral culture of the Indo-Iranians. Also another word denoting a kind of grass was borrowed from (Indo-)Iranian into Pre-Permic: *\*antV* > U *ud-*, *urt-*, Ko *od*, ← *\*ándhas-* ‘Kraut, Grün, Somakraut, Somatrank’, see the entry for that word for more details.

The Indic word has exact formal parallels in Germanic *\*þurna-* (> ON *þorn-* ‘thorn, briar’) and in Slavic *\*tǫrnǫ* ‘thorn’, all of which reflect a PIE *-no-* stem *\*tr̥no-*. The Uralic word could be borrowed from the Proto-Indo-European form as well, but semantically the Indo-Iranian loanword is clearly the best option, as the meaning of ‘grass’ is found only there. Because of the vocalism, as well as semantics, the Germanic and Slavic forms are out of question as donor forms (also the distribution in Finnic and Permic would be difficult to explain if the word was borrowed from these branches).

The reconstruction of the Uralic vowel is unclear, as the Permic and Finnic words can point to *\*a* or *\*i̯*, (the two sounds merge in most positions in Finnic and in *\*a*-stems in Permic), which makes determining the vowel substitution here rather difficult. The Finnic long vowel *aa* results from a well-known sporadic lengthening before *r*. The substitution *\*i̯r* ← *\*r̥* is phonetically expected, but it is actually not found in many loans. This substitution could be a sign of a more recent loan, as *\*er* seems to be the substitution in loans that have a wide distribution. In Proto-Indo-Iranian *\*tr̥na* the *r̥* is accented, which could have played a role here. It is not probable that the two words are later, parallel borrowings, at least not from a later Iranian form, as there is no evidence that the word was reflected in Alanic (there is no reflex in Ossetic), and the attested Middle Iranian forms show the assimilation *\*rn* > *rr*, meaning they could not be fitting originals for the Finnic and Permic words.

An Indo-Iranian etymology

**Md E, M *tarvaz* ‘sickle’**

< **Pre-Md *\*tarwas***

← **PI ?, cf. OI *taravāri-* ‘one-sided sword’ or (unattested) *\*darga-*, cf. Lith *dalgis* ‘scythe’**

(Joki 1973: 325; Rédei 1986c: 60–61; Katz 1987: 452, footnote 9; Helimski 2000: 192; Blažek 2003: 94–95)

This word is confined to Mordvin. The hypothetical Indo-Iranian etymology (= *\*darga-*) is doubted by Joki (1973) and KEWA and it has been criticized by Burrow (1976: 61), who notes that for Lith *dalgis* ‘scythe’; no cognate is attested in Indo-Iranian, the etymology has to be considered uncertain (“cannot be considered as having much value”). Here one has to agree with Burrow. The Mordvin word also cannot be a direct loan from the Baltic form, as Mordvin *r* could not be explained from Baltic *\*l*. Derksen (2015) considers Baltic *dalgis* a cognate to ON *telgja* ‘to cut’ and argues that the word is a North European (substrate?) word. Rédei refers to IEW, where a PIE noun *\*dolgho-* is reconstructed. The Indic word

*taravāri-* that Rédei mentions cannot possibly continue such a form, and this is noted also by Katz (1987: 452, footnote 9), who is critical of the Indo-Iranian etymology. Also Blažek (2003: 94–95) is critical of the etymology and suggests that a suitable Iranian origin for the Mordvin word could be found in the word *dargaz* ‘scythe’ in the Iranian Talysh language. As the background of the Talysh word is uncertain, this etymology remains likewise hypothetical.

Although Rédei does refer also to the remarks in KEWA that considers the whole Indo-Iranian comparison problematic, he nevertheless supports the Indo-Iranian origin of the Mordvin word, mentioning that it can be borrowed from a form akin to OI *taravāri-* ‘one-sided sword’. It remains uncertain what Rédei considers to be the exact origin of the Mordvin word, as he also mentions the possibility that the word is borrowed from the Nuristani branch. EWAia does not mention *taravāri-* or the alleged Nuristani cognates (Prasun *trāz*, *tərwaz*, Kati *tərwōc*) at all.

The hypothetical Indo-Iranian etymology has since been supported also by Helimski (2000: 192), who does not provide any further evidence. Because of the very hypothetical nature of the etymology, it is best to reject it altogether.

Katz (1985: 272) has suggested an alternative etymology, a loan from unattested Indo-Iranian noun *\*drapa-* ‘Sichel’, which would be derived from PIE *\*drep-* ‘abschneiden’ (LIV<sup>2</sup>: 128–129, s.v. *\*drep-*). Because of the hypothetical nature of this etymology, it cannot be considered more convincing than the traditional etymology.

Not an Indo-Iranian etymology

### Mansi (North) *tas (xum)* ‘Fremder’

← PII *\*dāsyu-* ‘stranger’ > OI *dāsyu-* ‘Götterfeind, Feind, Dämon, Angehöriger für Barbarenstämmen, die als unglaublich und nicht-opfernd gelten’, OAv *daxīiu-*, YAv *daŋhu-* ‘Land, bewohner eines Landes’, OP *dahyu-* ‘Land, Provinz, Gau’ (EWAia I 711–712, s.v. *dāsyu-*)

(Katz 1985: 106; 2003: 93; Lubotsky 2001b: 309)

Katz refers to the early explanation by Munkácsi that the word is borrowed from Iranian *\*dasa-*, but in this case the word should be a very late loan, after *\*ś > s* had taken place in Mansi. Both Katz and Lubotsky assume that the Mansi word (Pre-Mansi/Uralic *\*tāśV?*) is borrowed from an Indo-Iranian form *\*dasyu-* (*\*dāsyu-m* in the reconstruction of Katz, who apparently assumes that the word has to be borrowed from the accusative; the Indo-Aryan word is of masculine gender, the Avestan and Old Persian words feminine), and in this case the Pre-Mansi palatal sibilant could be explained as a substitution of the *-sy-* cluster. This is more specifically argued by Katz, Lubotsky does not pay attention to the development of Uralic sibilant, but notes only that the word cannot be a late loan from Iranian because of the *\*s*.

There are not many parallels to the substitution of a *\*-sy-* cluster by Uralic *\*-ś-* (see *\*śum̐s* for a possible parallel, which however has its own problems). This, together with the isolation of the Mansi word (it is found only in North Mansi), makes the Indo-Iranian

etymology very uncertain. Munkácsi & Kálmán (s.v. *tas*<sub>2</sub>) mention that the word is borrowed from Komi, but it remains uncertain what Komi word is meant here.

Not an Indo-Iranian etymology

**Fi *tiine* ‘pregnant (of animals)’**, cognates in most Finnic languages; **SEst *tiinõh* (cognates also in Veps, Ludic, Votic, Izhorian and Estonian)**, < PFI *\*tijněš*; **Mari *tüž*, *tüjüž* ‘pregnant (of animals)’** < Proto-Mari *\*tjuž* (TschWB 846–847)

< PU ? *\*tejniš*

← PIE/Pre-II? *\*d<sup>h</sup>eh<sub>1</sub>inyah<sub>2</sub>-*/*\*d<sup>h</sup>eh<sub>1</sub>inu-s* > OI *dhenú-* ‘cow, milking cow’, Av *daēnu-* ‘female animal’, from PIE *\*d<sup>h</sup>eh<sub>1</sub>-in-*, root PII *\*d<sup>h</sup>aH* < PIE *\*d<sup>h</sup>eh<sub>1</sub>-* ‘to suck mother’s milk’ (EWAia I: 797, s.v. *dhenú-*; LIV<sup>2</sup>: 138, s.v. *d<sup>h</sup>eh<sub>1</sub>-*; Werba 1997: 299 No. 308 s.v. *dhā/dhī(dhe)* ‘saugen’; Garnier et al. 2017: 296; Derksen 2015: 127–8)

(Kalima 1936b: 169; Joki 1973: 329; Katz 1985: 235; SSA III s.v. *tiine*; EES s.v. *tiine*; Aikio 2014b: 90–91)

The Indo-Iranian etymology for the Finnic word was suggested already by Kalima (1936). Although it seems to have fallen out of favour by the early 2000s (see SSA etc.), Aikio (2014) has convincingly rehabilitated this old Indo-European etymology on the basis of his new – and convincing – ideas on the development of Proto-Uralic vowels in Finnic (see also the entry for *piimä*). The Mari developments are also neatly explained by Aikio: Proto-Mari *\*ü* reflects PU *\*e* in a number of cases, and there are also parallels for the loss of intervocalic *\*n*, which makes Finnic *tiine-* and Mari *tüjü-* ~ *tüü-* regular cognates.

The Indo-Iranian origin of the Finnic and Mari words is plausible, but the borrowing should be very early because of the vowel *\*e* that has to be reconstructed on the basis of the Finnic and Mari words. However, also a Baltic origin has been suggested for the Finnic word already by Lõo (1911: 86), and this is mentioned as a possibility by SSA and EES too: the cognate of the Indo-Iranian word is Proto-Baltic *\*deini* > Lith *dieni* ‘with young’ (Derksen 2015: 127–8). Liukkonen (1999: 142–144) has supported the Baltic origin, and actually argued for a similar sound change *\*ej* > *\*ij* as Aikio. Junttila (2012: 278) has rejected Liukkonen’s arguments, but thanks to Aikio we now know that Liukkonen was, in fact, right, although Liukkonen’s arguments were not based on as solid evidence as Aikio’s.

There are no phonological reasons to argue for the superiority of either Baltic or Pre-Indo-Iranian etymology, but semantically the Baltic words are better. The meanings are derived from relationships to the sucking of milk, and the Baltic meaning of pregnancy has clearly developed from this. The attested Indo-Iranian forms denote simply (milch-)cows, and no meaning related to pregnancy is attested. Semantically the Baltic words are thus closer to the Uralic ones, and there are no obstacles to deriving the Mari and Finnic word from Baltic/Balto-Slavic. In Finnic there are numerous Baltic loanwords, and there are some in Mari as well, although many of the Baltic etymologies for Mari words are not very credible (Mägiste 1959). If the Mari and Finnic words are derived from a unitary PU form, it is more natural to assume that the borrowing is from Proto-Balto-Slavic, as a Baltic loan shared by Finnic and Mari would be unique. If Finnic and Mari reflect a Proto-Uralic word,

the borrowing could also be from Proto-Indo-European or some very early daughter language, such as Core Indo-European. The words can also be parallel borrowings from Baltic.

The nature of the original non-initial syllable of this word is difficult to explain, and receives no comments in Aikio's otherwise careful treatment. The Finnic *-eš* looks like a derivational suffix (this is a productive suffix, cf. Fi *ihme* 'wonder' < \**imeš*). However, also the Mari word-final *-ž* looks suspiciously like a reflex of a PU sibilant. The regular reflex of any PU sibilant would be Mari *š*, but between vowels *ž* is the expected outcome. Here it has to be assumed that the Mari word has been trisyllabic, and *ž* reflects a sibilant that has become voiced. It would be tempting to think that the PU (?) *-š* would reflect either the nominative *-s* or the laryngeal, but since both the Mari and the Finnic words look like derivations, it is probable that the sibilants in both languages/branches have arisen independently.

Hungarian *tej* 'milk' is a later Iranian loan from ultimately the same Indo-Iranian root (EWUNg s.v. *tej*; WOT: 1336).

? Indo-Iranian etymology (can also be a Balto-Slavic loan)

**Fi *kahdeksan* '8', *yhdeksän* '9'; (North) Saami *gav-cci*, *ov-cci* id.; Mordvin E *vejkse*, M *vejχsa*; Mari *kandaš*, *indeš* id.**

**< West Uralic/Finno-Permic ? \**-teksä* ~ \**-teksa***

← “Pre-Iranian” \**détsa-* ‘ten’ (= PII \**dátsa-* < PII \**dáća-*) < PIE \**dék̑m-* ‘ten’ > Av *dasa-* ‘10’ (Szemerényi 1960; 1977b: 129–130; Parpola 1999: 198–199; Koivulehto 1999a: 220; 2001b: 255 = 2016: 222, 297; Bereczki 2013 s.v. *indeš*, *indeše*, s.v. *kandaš*, *kandaše*)

The idea of deriving the element \**-teksä* from the Indo-European word for ‘10’ is old. Already Euroepaeus 1853: 12–13) argued that the element in Finnish *kahdeksan* and *yhdeksän* is related to the Indo-European words for ‘10’. Jacobsohn (1922: 171–172) suggested the possibility that the word is borrowed from Indo-European. Joki (1973: 25–26) only mentioned the old ideas that the word is a loanword in his survey of research history but did not include the etymology in the Wörterverzeichnis because of the implausibility of the etymology. Joki's decision was criticized by Szemerényi, who supported the Indo-European origin of \**-teksä* (see also Szemerényi 1960). Joki's reluctance to support the etymology is easy to understand, as none of the earlier explanations is able to argue convincingly from which Indo-European form the Uralic words were borrowed.

The Iranian etymology in its final form was presented by Parpola (1999: 198–199) and Koivulehto (1999a: 224–225), who were the first to argue for a more precise Iranian origin for the word instead of more ambiguous comparisons between \**-teksä* and the various Indo-European words for ‘10’. This idea was based on Koivulehto's argument for a substitution Uralic \**-ks-* ← PI *-ts-*. However, this etymology is problematic for numerous reasons.

The word *\*\*teksä* does not exist as an independent word in any of the languages mentioned above, but is simply a postulated part of compounds ‘one-to-ten, two-to-ten’. While the lack of a simple word *\*teksä* is not an insurmountable problem, it is also suspicious that in languages other than Finnic the “compounded” forms are so reduced that it is almost impossible to say anything precise about their background. It is also troubling that the Mari dialects show a number of different forms, which do not point explicitly to a reconstruction *\*-teksä*.

The literary language forms of Mari are actually suspiciously similar to the word for ‘10’ in the Permic languages (*das*) which is a well-known (late) borrowing from Iranian (cf. Ossetic *dæs*). One could almost suppose that the Mari words are indeed compounds of some kind, consisting of some forms of ‘one’ and ‘two’ and the word *daš*, borrowed from Iranian (or from Permic). However, this also does not explain the strange dialectal forms, which point to a more complicated pre-form of the element *-daš* in Mari. But one could nevertheless suspect that the form *-daš* has been secondarily influenced by contacts with either Permic or Alanic.

The etymology of this word has been also explained differently already by E. Itkonen (1973: 336–339), according to whom the words ‘8’ and ‘9’ derive from words *\*kakti e-ksä* ‘zwei existieren nicht’ and *\*ükti e-k-sä* ‘ein existieren nicht’, respectively. This explanation is not totally convincing either, and does not remove the phonological difficulties of the reduced forms of Saami and Mordvin or the strange dialectal forms *kändäkšə*, *kändäjšə* in Mari, but it is more compelling than the Iranian etymology.

The Iranian etymology also suffers from the assumption that Indo-European *\*e* would have been retained until the Iranian period (here the consonantism points clearly to Proto-Iranian). Koivulehto and Parpola have both argued that the merger of non-high vowels was so late that these kinds of consonantal change could have already taken place before it in Iranian, but their linguistic evidence for this is very meagre (see also the entry for *\*kečrā*). It rather seems that this kind of idea cannot be supported by the etymologies they present. Together with these chronological problems and the Uralic problems mentioned above, it is best to reject this etymology.

Bereczki (2013) notes that Koivulehto’s etymology is problematic for the Mari forms because in his explanation the auslaut *-n* of the reconstructed protoform is due to the influence from *kymmenen*, which is absent from Mari and looks like a Finno-Mordvinic explanation. The phonological problems of the loan etymologies themselves are not discussed. While this is an interesting point, the *-n* can be a later development from analogy independently in Finnic and it does not have to reflect a proto-form with *\*-n*. In any case the Mari word is difficult to analyse and it is unlikely that the word ultimately derives from Iranian.

Hungarian *tíz* ‘ten’ is borrowed from the same Iranian word (with irregular *\*z* from Iranian *\*s < \*ts*).

Not an Indo-Iranian etymology

**Fi *terni* (stem *terni*-) ‘milk of a cow that has recently given birth’, also in Est *ternes*, *ternespiim*, Li *ter-sēmd’a* id.**

← ? Pre-II \**terun*- > PII \**taruna*- > OI *táruṇa*- ‘young, fresh’ Av *tauruna*- ‘young; son’, Oss *tæryn* ‘son’ (EWAia I: 632, s.v. *táruṇa*-)

(Setälä 1928: 304–305; Joki 1973: 328; Katz 1985: 144; 2003: 121; Rédei 1986c: 61; SSA III s.v. *terni*; Lushnikova 1990: 268; Hahmo 1994: 84–85; EES s.v. *ternes*)

This is an old etymology, originally suggested by Setälä. If the etymology is convincing, the borrowing has to be rather old despite the limited distribution because of the retained \**e*. Setälä’s etymology was criticized already by Jacobsohn (1933: 138–139 MSFOu LXVII), who notes that it is unlikely that the Finnic disyllabic word could be derived from trisyllabic Indo-Iranian \**taruna*-. Also the form *tárṇa*- is attested in Old Indic, and this is cited by Setälä, but this is a later form that shows a Middle Indo-Aryan development (EWAia), and it is impossible to derive the Finnic word from this kind of form. It is possible to assume that the word was simply borrowed as disyllabic into Uralic, as most of the Uralic word-stems were disyllabic.

EWAia lists a number of possible cognates for the Indo-Iranian word, Greek τέργν ‘soft, delicate’, *τέρυ* ‘weak, soft’, possibly metathesized Latin *tener* ‘soft, delicate’ cf. de Vaan 2008: 613). EWAia considers the connection of *táruṇa*- and the adjective *turá*- ‘sick; tender’ possible. Both could be derived from the PIE root \**ter*- ‘to grind’, *turá*- reflecting earlier \**tr̥h<sub>3</sub>-ó*-. Theoretically the Finnic word could be a borrowing from some other archaic branch of Indo-European, although assuming a loan from some other branch of Indo-European would not solve the phonological problems either, and one has to note that the word does not appear in the branches such as Balto-Slavic or Germanic that have had most contact with Finnic, and the meaning ‘soft’ attested in other branches does not really fit the meaning of the Finnic words.

The Finnish (non-altering) *i*-stem is secondary and results from derivation, cf. \**kota-j* > *koti* ‘home’. The Estonian form probably continues PFi \**terneš* (-*s* here can be generalized from sandhi forms, as it is, of course, not a regular reflex of \*-š). The original stem vowel should have been either \*-*ä* or \*-*i*, but it is difficult to determine from the modern Finnic forms.

Setälä (1928: 304) connected *terne* with *terhen*, *terendama*, Fi *terhentää*. This has been criticized by SKES; Hahmo mentions the idea and agrees with SKES. This equation is not plausible because of semantics. Some expressions like *terneet* ‘clouds’, *terneh-pilvi* could have been derived from *terhen* via metathesis, but these are etymologically unrelated to the other word *terni* ‘milk of a cow’. This explanation is still mentioned by EES, however. *terendama* (Fi *terhentää*), has a possible Baltic etymology (from hypothetical Proto-Baltic \**ster*-), suggested by Vaba (1997).

? Indo-Iranian etymology

**Fi *terve* ‘healthy, whole’, Est *tere* id. and regular cognates in all the other Finnic languages < PFI *\*terveh***

**< Pre-Fi *terwV***

← PII (or PI) *\*drva-*, *\*drvā* > Av *druua-* ‘healthy’, OP *duruva-* ‘solid, firm’, New Persian *darōd* ‘health, bloom’, OI *dhruvā-* ‘solid, firm, fixed, secure’ (EWAia I: 798–799, s.v. *dhruvā-*)

(Setälä 1928: 300; Koivulehto 2003; EES s.v. *tere*)

Koivulehto’s etymology was presented in a presentation and its handout. While the etymology has not been published as such, it is referred to by EES, K. Häkkinen (2004) and LÄGLOS (III: 291, s.v. *terve*), all of whom accept the etymology. LÄGLOS considers Koivulehto’s Indo-Iranian etymology better than the Germanic etymologies that have been suggested: Katz (1990: 14ff.) has derived the Finnic word from PG *\*trewwas* (> ON *tryggr* ‘faithful’), and Hofstra (1992: 59–60) from PG *\*derbaz* (> ON *djarfr* ‘brave’). In SSA, the relationship of *terve* to *terva* ‘tar’ (originally presented by Kiparsky [1952: 94–99]) is considered the most viable option: *terva* is originally a Baltic loan from *\*derva* ‘tar’. Koivulehto’s etymology is semantically more convincing than Kiparsky’s explanation, but it was not yet known to the editors of the SSA.

Both semantically and phonologically the etymology suggested by Koivulehto is convincing. The meaning ‘healthy’ seems to be attested exclusively on the Iranian side, although also there the more original meaning of ‘solid, firm’ is attested in Old Persian. According to EWAia, the adjective is derived from the verbal root *\*dhar-* ‘to keep, maintain’, which is likewise not far from the semantics of ‘healthy, whole’. The vowel *e* results here from the substitution of the cluster *\*drv-*. Here the Indo-Iranian form does not have syllabic *\*r*, so the word-initial consonant cluster of a stop and *r* has been solved in the Uralic side by adding an epenthetic vowel. This is a very rare substitution among earlier loans, which casts some doubt on the etymology, but Uralic/Volgiac *\*mura* ‘sing’ from Indo-Iranian *\*mrawati* serves as a parallel (the vowel is different, but the substitution strategy is the same).

Setälä (1928: 298–308) had presented another Indo-Iranian source for the word *terve*, namely Pre-II *\*dhermen-* (> PII *\*dharman-* > OI *dharman-* ‘dharma; order’), but this etymology was rejected already by Jacobsohn (1933: 139), who argued that the etymology is phonologically impossible (there is no reason for the Indo-Iranian cluster *\*-rm-* to be substituted by *\*-rv-* on the Uralic side). Jacobsohn’s argument is still valid, and deriving the Finnic word from *\*dhermen-* is impossible.

A Slavic etymology for the Finnic word has been suggested by Ahlqvist (1857), who derived the word from Russian *здоровый* ‘healthy’ < PS1 *\*s̥d̥orvъ*), which is ultimately from the same Indo-European root *\*d<sup>h</sup>er-* as the Indo-Iranian word (the Slavic word continues PIE *\*h<sub>1</sub>su-d<sup>h</sup>or-uo-*, cf. Derksen 2008: 478–479, s.v. *\*s̥d̥orvъ*). It would be very difficult to derive the Finnic vocalism from either the Proto-Slavic or Russian word.

**An Indo-Iranian etymology**

## Hu *tíz*, dial. *téz* ‘10’

← ? Iranian *dasa-* (see *\*teksä* for references)

(Joki 1973: 329–330; Harmatta 1977: 171; Ligeti 1986: 148; Honti 1993: 159; WOT: 1356)

This word is probably not borrowed from Alanic, but an earlier loan, belonging to the same layer as *tej* and *tehén* (see above). TESz considers the word an early loan because of the unvoiced *\*t-*. Ligeti and WOT consider the vocalism problematic, as the Iranian form had only *\*a*. In later Iranian (Alanic) loans in Hungarian the Iranian *a* vowel is sometimes reflected as Hu *é*. There should not be any problems in deriving Hu *í* from earlier *\*e*, cf. Hu *víz* ‘water’ < PU *\*weti*, and here we could perhaps assume that the vocalism *é* that is retained in some dialectal forms is primary. If the word was a Pre-Indo-Iranian loanword, there would be no problems with *\*í*, but there are no known examples of a Pre-Indo-Iranian loanword that would be retained only in Hungarian.

The etymology is convincing, but the voiced *z* cannot be explained, as there are no other examples of this sound substitution. Harmatta reconstructs the Pre-Hungarian (or Proto-Finno-Ugric in his terminology) form as *\*teše*, assuming a borrowing from a form that retained PII *\*ć*, but assuming a Pre-Hu *ś* will not make *z* more regular. Joki notes that the loan is from Middle Iranian, or more precisely “West Middle Iranian”, but this does not help much. One possibility would be to assume a loan from a Persian-type language and consider a substitution Pre-Hu *\*t* (which would yield *z*) ← Persian *\*θ*, but there are no parallels to this. Note that the Hungarian word *bűz* ‘smell’ (Ligeti 1986: 149) is not a parallel, as its *z* has been explained as a dissimilation<sup>48</sup> from *d* that is retained in the adjective *büdös*, and in any case here the Iranian original has *d* that is derived from PII *\*dh*.

Ligeti discusses the word in his listing of Iranian loans of Hungarian and lists the various problems that are mentioned by Joki and TESz, but he does not present any concluding remarks about the precise Iranian source of the Hungarian word.

Permic *\*das* ‘ten’ is a transparent Iranian loan from the same origin, probably from Alanic (see Appendix I).

An Indo-Iranian etymology

**Fi *toivoa* ‘to hope, wish’, *toivo*, *toive* ‘hope, wish’, Est *tōotama* ‘to promise’, cognates in all Finnic languages, Votic *toivoa* ‘to wish, to foresee’ probably borrowed from Ingrian (SSA s.v. *toivoa*; EES s.v. *tōotama*)**

< Pre-Fi *\*tojwa-*

← PII (or PI) *\*daywa-* or *\*dāywa-*, *\*dāiwyā-* > OI *devá-*, *daíva-*, *dáivya-* ‘heavenly, divine’, (EWAia I: 742–743, s.v. *devá-*)

(Koivulehto 2003; ms.)

<sup>48</sup> A similar dissimilation has been assumed for Hu *édes* ‘sweet’, which is apparently a derivative of *íz* ~ *éz* ‘taste’; in a word with *z* – *s*, the *z* becomes *d*. This rule is supported by very few examples only, which makes it dubious.

Koivulehto's etymology for this Finnic word family is convincing. Koivulehto never formally published the etymology, but it is referred to by Häkkinen (2004) and EES. The Indo-Iranian original is the same word from which the noun *taivas* (see the entry) has been borrowed; the semantic difference between 'to wish' and 'heaven' is rather wide, but both can be derived from the semantics of the Indo-Iranian word, as Koivulehto lists also the meanings 'divine will, faith, happiness' among the meanings of the Indic word.

It is uncertain whether the word is confined to Finnic, as North Saami *doaivut* (with cognates in most Saami languages) is either a cognate of the Finnic word or borrowed from it (Kuokkala 2018: 32). In SSA the words are considered to be cognates, EES mentions the possibility of Finnish loan into Saami, although no further arguments in favour of this are given. The relationship of the Saami and Finnic words remains unclear for the time being. Cognates in other related languages have not been suggested.

It is difficult to see why PII \**a* was substituted differently in the same environment in these two words, \**tajwas* and \**tojuV-*. One possibility is that the words reflect two different layers of borrowing, but it is very difficult to prove this. On the other hand, Koivulehto notes that for the word in question an old vrddhi formation \**dāywa-* (PIE \**dēywo-*) can be reconstructed (cf. also Gotō 2013: 53–54).

There are also examples of Uralic/Pre-Finnic \**aj* being reflected as \**oi* in Finnic, at least \**waji* 'butter' > *voi* is a convincing example of such a development, although here the word-initial \**w-* has probably played a role too (Aikio 2012a: 236).

One can suggest as a possible solution that in *taivas* a short *a* is reflected, whereas in *toivo* the vowel *o* is the substitute for a long *ā*. Admittedly there is not much phonetic substance in this explanation. Another hypothetical reason could be a difference in accentuation: in the Vedic vrddhi-forms *daivya-* 'divine', fem. *daivi-*, *dāiva-* *id.* (but note also *daivá-*) the accent is on the first syllable, but the accent is on the last syllable in *devá-*.

Despite the uncertain vowel substitution, the etymology can be accepted. No competing loan explanation for the Finnic word exists.

An Indo-Iranian etymology

**Fi *tora* 'fight, quarrel', *torata* 'to fight', Estonian *tõra*; SaN *doarru-* 'to fight' (< PSa \**doaro-*); TN *taro-*, Mat *toro-* 'to wrestle' < PS \**tāro-***

< PU \**tora-* (UEW s.v. \**tor3*; SW: 12; Sammallahti 1988: 554; Kulonen 2010: 212)

← PII (?) \**dhārāya-* > OI (Vedic) *dhārāyati* 'maintains', root *dhṛ* 'halten, festigen, stützen, tragen', PII \**dhār-* > PI \**dar-*, < PIE \**dher-* (EWAia I: 778–779, s.v. *DHAR*; Werba 1997: 200; Cheung 2007: 57–59, s.v. *dar<sup>1</sup>*)

(Aikio 2015b: 62)

The recent Indo-Iranian etymology for this word family is good, but Aikio expresses his view on the connection of Saami-Finnic and Samoyed words rather ambiguously, as he states that "at least the Saami and Finnic words" can be derived from the aforementioned Indo-Iranian source. However, as the Samoyed word is a phonologically regular cognate of the Finnish and Saami forms, there would be no problem in supposing a Proto-Indo-

Iranian source for a PU word ancestral to all the forms mentioned here. Nevertheless, semantically the assumption of an Indo-Iranian loan etymology is not clear, as it is unclear why Aikio assumes the Uralic words to be derived from an Indo-Iranian word with a rather different meaning. While ‘to maintain’ and ‘to fight’ are meanings that can be related, at least some semantic parallels should be sought to support this development. Aikio’s etymology remains thus uncertain but possible.

There is also another Indo-Iranian word which might be a source for this Uralic word: the OI (Vedic) root *tar<sup>i</sup>-*, present *táratī* ‘kommt durch, überwindet’, *i*-reduplicated present form *tirati* ‘bringt (sc. durch Hindernisse) hindurch’, (EWAia I: 629 s.v. *TAR<sup>i</sup>-* < PIE *\*terh<sub>2</sub>-*, see LIV<sup>2</sup>: 633; for a detailed account on this Sanskrit word, see Tichy 2004). This etymology is also not semantically straightforward, but the meanings ‘to overcome’ and ‘to fight’ can be connected at least equally well as the words suggested by Aikio. In Young Avestan the meaning ‘fought’ is also attested in *titarat* ‘bekämpfte, drang an’. Also in Hittite the reflex of the IE root is connected with ‘fighting, overcoming’, as the name of the god Tarḫunnas ‘the overcomer’, the thunder-god who defeated the great serpent Ilyunkas, is derived from it (Watkins 1995: 450; Watkins 2011, s.v. *\*terə-*).<sup>49</sup>

The vowel substitution *\*o* ← *\*a* is here similar as in Aikio’s etymology, although the Uralic forms could also reflect the zero-grade present form (PII *\*tr̥H-á-ti*) equally well. If the Uralic word is borrowed from a form with *\*a*, this is one of the many examples where PII *\*a* was substituted by PU *\*o* before *\*r*.

Aikio also mentions the Komi verb *dorjī-* ‘to defend, protect’, which he considers a separate borrowing from the same Indo-Iranian source. Actually, the Komi word’s semantics suit the Indo-Iranian *dhārayati* ‘to maintain, bear, support’ better than the other Uralic forms, and it suits less well the other Indo-Iranian word which is offered here as the source of the Proto-Uralic *\*tora*. One could tentatively present that *\*tora* is borrowed from *\*tr̥H-* ‘überwinden’ and Komi *dorjī-* from *dhārayati* ‘to maintain, bear, support’. Komi *o* could reflect old *\*a* in Uralic *\*a*-stems. According to Metsäranta (personal communication), the Komi word could be derived from *dor* ‘edge’.

Aikio notes that the Mordvin verb *E t’urēms*, *M t’urēms* ‘to fight’ has also been connected here (UEW s.v. *\*tor<sub>3</sub>-*; Kulonen 2010: 212), but that is problematic because of the palatal consonants. One option would be to consider the Mordvin word a separate borrowing too, but this does not remove the phonological difficulties. As palatalization of consonants is very frequent in the Mordvin languages, one could perhaps assume an analogical development here. It is also noted by Aikio that SaN *doarjit* and Finnish *torjua* have been compared to the Mordvin words; semantically they could also be borrowings from either of the Indo-Iranian words mentioned above, and the vowel could represent either PII/PI *\*a* or *\*r̥(H)*.

Mari *torle-* ‘rügen, schelten’ is also mentioned as a cognate by the UEW, and Hungarian *dorgál-* ‘rügen, tadeln, schelten, zurechtweisen’ is connected here with a question mark. *-gál* is a productive suffix in Hungarian, and the same is said of Mari *-l-*. The Hungarian word is considered to be either of unknown origin or to be related to the Uralic words mentioned above by Zaicz (2006, s.v. *dorgál*). Semantically it is not clear whether the Hungarian word can be derived from the same Uralic source as the other words, but there

<sup>49</sup> An opposing view about the etymology of the Hittite word has been expressed by Lindeman (1987: 46, footnote 27), who considers the Hittite verb unrelated to PIE *\*terh<sub>2</sub>-*.

are no phonological problems (Hungarian word-initial *d-* instead of *t-* occurs in some inherited words, although the conditions for this development are unknown; Zaicz mentions *daru* ‘crane’ [< PU \**t̥rka*] as a parallel, and this is also accepted in Aikio’s 2015b wordlist). Zaicz mentions an onomatopoeic origin as one possible etymology, and the UEW considers the Uralic word to have arisen through onomatopoeia. This is not a very precise explanation, because we know very little of the birth processes of onomatopoeic words at the Proto-Uralic level, so it is very difficult to prove this kind of origin. A parallel onomatopoeic Urschöpfung in all the different Uralic languages is out of the question.

To return to the Indo-Iranian etymology of the word \**tora-* that is found in Finnic, Saami and Samoyed, this is a particularly important etymology because it is a possible example of a loanword that has reflexes both in the traditional “Finno-Ugric” and Samoyedic languages. It has long been argued that the distribution of Indo-Iranian loanwords is one of the key pieces of evidence for the early binary split of Proto-Uralic into Samoyedic and Finno-Ugric branches. Although counter-evidence has been presented in form of early *Indo-European* loanwords that are present in both “branches” (Koivulehto 2001b), the Proto-Indo-European loan etymologies are less convincing and more open to different interpretations. Solid Indo-Iranian etymologies for Uralic words with a distribution stretching to Samoyed provide more convincing evidence against the traditional assumption of the early binary split.

J. Häkkinen (2009) has argued that Indo-Iranian loans into *East Uralic* (= Proto-Ugro-Samoyedic) prove the idea of an early split wrong, but this is right only for those who accept Häkkinen’s idea of another binary split, into Finno-Permic and East Uralic. The problem here is again that there are no borrowings with an Ugro-Samoyedic distribution, and all the evidence for contacts with Ugro-Samoyedic comes from the Ob-Ugric languages.

To sum up the arguments, there are two possible Indo-Iranian etymologies for the Proto-Uralic \**tora* word, and it is a probable example of an early borrowing into Proto-Uralic, the reflexes of which are found in both Samoyedic and the other Uralic languages. A methodological note on the existence of two competing etymologies is in order here: Klaas Ruppel (2015) has argued that the alternative loan etymologies from the same source language do not actually strengthen an etymology but rather weaken it. Keeping this in mind, the Indo-Iranian origin of \**tora* cannot be considered certain.

As a side note it should be mentioned here that Hyllested (2014: 83–85) derives the Celtic word for ‘pig’ from this same Uralic word, more precisely from Finnic, where words for ‘pig’ and ‘boar’ have developed from the word originally meaning to ‘fight’; the boar’s fang is called *torahammas*, and the name of the boar is obviously derived from this. However, as these boar names do not even go back to Proto-Finnic, it is very unlikely to postulate a contact situation where this word was borrowed into Proto-Celtic (!), and this situation is simply impossible to imagine between some later Finnic language and some Celtic idiom. It seems very improbable that there were any kinds of contacts between Celtic and Finnic at any time, and Hyllested’s other examples of such lexical contacts are also rather unconvincing.

Grünthal (2002) has derived Hungarian *tart* ‘to keep, to maintain’ from the same Indo-Iranian verb *dhāráyati*. Here the semantic relationship is more convincing than with Aikio’s etymology. Rédei (1963) connects *tart* with Finnic *tarttua*, but this is probably erroneous. Instead, *tart* could be borrowed from Alanic, as in Ossetic a stem *dard-* ‘halten,

besitzen' is attested (Bielmeier 1977: 139), although the unvoiced stops in the Hungarian word point to an earlier origin (in Alanic loans, the Iranian voiced stops are usually reflected by voiced stops in Hungarian). The Iranian origin of the Hungarian word remains probable but uncertain.

An Indo-Iranian etymology

**SaN *doaris*; E *troks, truks, turks*, M *tarks, turks* 'via, through, across, over'; MariE, W *toreš* 'across, crosswise'**

< *\*toras, \*toraksi* 'crosswise' (UEW s.v. *tora-ks3*)

← PI *\*tarás* (or *\*tr̥Hás*) < PII *\*tr̥Hás* > YAv *tarō* 'sideways, to the side, through, over', OI *tirás* 'through, over' (also a preposition with Accusative) < PIE *\*tr̥h<sub>2</sub>as* > Old Irish *tar* 'across'; root *\*tarH-* < PIE *\*terh<sub>2</sub>* (EWAia I: 646–647, s.v. *tirás*; Gotō 2013: 149; Dunkel 2014: 799–; Pinault 2016: 101)

(Aikio 2015b: 62)

This recent etymology of Aikio is convincing, and there is no earlier loan etymology for this “Finno-Volgaic” word. However, the Mordvin, Mari and Saami words cannot really be regularly derived from the same source due to irregularities in the vocalism, so they might present parallel borrowings from the same Indo-Iranian source. The first-syllable vowel substitution PI *\*a* > PU (“PFW”) *\*o* is well-attested, and the environment before *\*r* seems to often trigger this substitution (Gamkrelidze & Ivanov 1984; Heikkilä 2014: 46). However, as the Indo-Iranian laryngeal *\*H* (from PIE *\*h<sub>1-3</sub>*) was probably still retained in Proto-Iranian, this word might be a slightly later borrowing with *\*r̥H* already vocalized as *ar* (in Indo-Aryan, the outcome was *ir*, as can be seen from the Sanskrit cognate). Proto-Indo-Iranian *\*tr̥Has* might also be a possible source, as *\*o* would equally well work as a substitute for the syllabic resonant.

Regarding the etymology of the Indo-Iranian word, Gotō (2013: 149) writes that PII *\*tr̥Has* is probably an old ablative form of a root noun. Stüber (2002: 30–31) writes that the preposition *\*tr̥h<sub>2</sub>-és* can be reconstructed already for PIE, and this is reflected in the Indo-Iranian branch by Vedic *tirás* 'über ... hin' and Avestan G *tarē*, Y *tarō* 'seitwärts, durch, hin' and also in Celtic (Old Irish *tar* 'durch, über... hin' < PC *\*tares*). Stüber (referring to Nowicki 1976: 52) argues that this preposition is *not* a locative of a postulated PIE *s*-stem *\*térh<sub>2</sub>-os* but a genitive form of a PIE root noun. In Vedic there exists an *s*-stem *táras-* (seemingly reflecting the PIE *\*e*-grade), but according to Stüber this is a later Indo-Aryan Neubildung.

Also the substitution of the second syllable is without problems, as there are many examples of the PII or PI masc. nom. ending *\*-as* being reflected by *\*-as* in loanwords in Uralic. It is, however, interesting that here the word is borrowed into a stem type *\*-ksi*. It is unclear why this should be the case, but perhaps the stem type has been productive enough to cause loanwords ending in *\*-as* to be absorbed into it. It would be interesting to find more examples of this. Mordvin *-ks* might be a derivational suffix. Mari *š* can be a lative

suffix here, especially as we are dealing with an adverb. The Saami has to be either an *\*o-a* or *\*a-i*-stem.<sup>50</sup>

The Indo-Iranian root here is ultimately the same as the one behind the origin of Uralic *\*tora-*.

An Indo-Iranian etymology

**Fi *touko*; Mordvin E *tundo*, M *tunda* ‘im Frühling, Frühling’; K *tuvış*, Ud *tuliş* ‘Frühling’;**

**Hu *tavasz* ‘spring’; Kha (eastern) *toy id.*; M (northern) *tūja id.* (UEW: 532–3, No. 1069)**

< PU ? *\*tawkV*

← PII *\*taukman-* ‘germ, germed seed’ > OI *tókman-* ‘Schössling’, Av *tauxman-* ‘Samen’, OP *ta-u-ma-a-* ‘Familie’; *toká-* ‘Nachkommenschaft’, *túc-* ‘posterity, children’; root *\*tawk-* (Schindler 1972: 22; EWAia I: 651, s.v. *túc-*, 670, s.v. *toká-*; Lubotsky 1988: 78; 2001b: 314)

(Katz 1985: 269–270, 2003: 216; Koivulehto 1991: 11, footnote 4; Koivulehto 2003; Blažek 2003: 96; Holopainen, Junttila, Kuokkala 2017: 114–115)

This is an old Indo-Iranian etymology, but it has not been commented on by recent sources in detail. Koivulehto discussed the etymology briefly in 1991, and also in the 2003 handout. The Uralic words have been recently subjected to a critical analysis by Holopainen, Kuokkala & Junttila (2017), who state that it is impossible to derive all the Uralic words from the same PU source, and thus an Indo-Iranian borrowing into Proto-Uralic has to be ruled out. However, some of the individual Uralic words could theoretically have been borrowed from Indo-Iranian. This was already observed by Katz, who only mentions the Finnic loan as a borrowing from Indo-Iranian and considers the relationship with the other Uralic words denoting ‘spring’ unlikely. Also Koivulehto (ms.) separated the Ugric words from the Finnic and possibly the Mordvin words and derived only the latter from Indo-Iranian.

As observed by Holopainen, Junttila & Kuokkala (2017) Finnic *touko* cannot reflect PU *\*tawkV*, as the Uralic diphthong *\*ow* regularly yields Finnic *\*uu* (Aikio 2012a: 243; 2013: 11). The Hungarian word could reflect a form *\*tawkV*, but it could equally well be derived from a form with simple *\*o* or *\*u*.

The origin of the Finnic labial vowel *-o* is unclear (usually the labial vowels in non-initial syllables of Finnic are secondary, but Aikio [2015b] has recently suggested that in some cases Finnic *-o* could reflect PU *\*-o*; this idea is disputed). One could perhaps suppose that the *-m* in the non-initial syllable of the PII word could have influenced the labial vowel on the Uralic side, but it is difficult to prove this. Although Aikio has argued

---

<sup>50</sup> The arguments presented here are based on my discussion with Riho Grünthal and Niklas Metsäranta. We intend to deal with this etymology in more detail in a future publication on Indo-European lexical influence on Mari.

that Uralic \**a*-*o* words have reflexes that differ from \**a*-*i* and \**a*-*a* words, we have no convincing examples of \**o*-*o* words, which makes it practically impossible to determine whether the Finnic labial vowel here is old (derived from PU) or not.

While Koivulehto (1991) considers the word an Indo-European but not necessarily Indo-Iranian loan, this is unlikely as EWAia considers the Indo-European background of the Indo-Iranian words unclear. The Uralic words cannot be borrowed from any other branch of Indo-European, as the word \**taukman*- is confined to Indo-Iranian according to Lubotsky (1988: 78). Lubotsky (2001b: 314) considers this as one of the inherited words in Avestan and Sanskrit, but no cognates outside Indo-Iranian can be presented with certainty.

Lubotsky (2001b: 314) notes that in OI also a root noun *túc*- F ‘posterity, children’ is attested (cf. Schindler 1972: 22 who considers this root noun a Proto-Indo-Iranian formation without further Indo-European cognates). This has to reflect PII \**tuk-e*-. If the Uralic word was borrowed from this form, this could explain the lack of the reflex of \**-m*-, but the zero-grade \**u* would be less fitting than the full-grade diphthong. However, as the cluster \**-km*- would have been impossible in Uralic anyway, it is difficult to determine with certainty which is the exact donor form of the Uralic word.

Semantically the etymology seems possible. Uralic \**ow* could easily reflect the PII diphthong \**au*; \**o* is a frequent substitute of PII \**a*, and here in a labial environment it may even be expected. This is one of the rare examples of Indo-Iranian diphthongs in the loanwords.

It is obvious that the alleged Uralic cognate set in reality represents several different words, and the connection of Finnic *touko* to the Indo-Iranian words seems unlikely. The Ugric words might be Indo-Iranian borrowings, but this is also uncertain.

? Indo-Iranian etymology

### **Hu *tehén*, dial. *tején* ‘cow’**

< Pre-Hu \**täjVnV*

← PI \**daynu*- ‘cow’ (see \**täjniš* for references)

(Joki 1973; Harmatta 1977: 171; Ligeti 1986: 146–147; Lushnikova 1990: 45; EWUNg s.v. *tehén*; WOT: 1336)

Although this can be a late loan because of its limited distribution, the word is discussed here because of the principle that distribution is not a very well-working criterion in the stratigraphy of loanwords. This is also obvious regarding the loans that are found in the Ugric branch/Sprachbund, as there are possibly very early loans that are found in only Khanty or Mansi, and the words that can have been borrowed into Proto- or Pre-Ugric are rarely found in all the languages. There is no question that the Indo-Iranian etymology of the word is correct.

The word is probably a rather early loan because of the disyllabic form of the Hungarian word, which can be best explained from \**dainu*- with a substitution \**täjVnV*, rather than a later Iranian form with vowel contraction. WOT lists the word among the Alanic

loanwords, but there is nothing that would make this specifically Alanic. The \**ä* substitution is found in early loans in Ugric, so the word can be an early borrowing, perhaps belonging to the same layer as Ugric loans like \**sīraña* (> Hu *arany*). Also the voiceless stop *t*- points to an early origin, as in later loans Iranian voiced stops have been substituted by Hungarian voiced stops (Ligeti 1986: 146). Ligeti also refers to an old idea that words like *tehén* and *tej* ‘milk’ could be borrowed from an Iranian language which had went through a devoicing of \**d*-, but this idea is ad hoc, and the Hungarian voiceless stop can easily be explained by Uralic/Pre-Hungarian phonotactics.

The *h* in standard Hungarian is a hiatus-filler, in some dialects another hiatus-filler, *j* is found.

Uralic \**tejniš* is a parallel loan from the same Indo-Iranian source of its Balto-Slavic/Baltic cognate (see the entry).

An Indo-Iranian etymology

### Hu *tej* ‘milk’

< Pre-Hu \**täji* or \**täjä*

← **Alanic, cf. Oss *dæjyn*, *dæyun* ‘(Brust) saugen’** (Bielmeier 1977: 140; Garnier et al. 2017: 298–299)

(Harmatta 1977: 171; Ligeti 1986: 147; Lushnikova 1990: 44; EWUNg s.v. *tej*; WOT: 1336)

Like the word *tehén* above, this word is found only in Hungarian but can still be a rather old loan. The word shows the \**ä* substitution that is found in Ugric early on. Semantically the etymology is convincing, although there is no noun meaning ‘milk’ that would correspond to the Hungarian word very well. Here Iranian *d*- has been substituted by Hungarian *t*-, which means that the loan has to be relatively early, as in late loans Hungarian voiced stops correlate with Iranian voiced stops (see the discussion in the entry *tehén* above).

An Indo-Iranian etymology

**Nga *tagai* ‘knife’, TF *tšixə* ‘hook’, Slk (Pr) *täqa* ‘sword’, (Tym) *tēya* ‘spear’**

< PS \**täjkå*

← **Iranian \**tayga*-**, > **Av *taēya*- ‘scharf, Schärfe’** (Cheung 2002: 231)

(Janhunen 1983: 124; Joki 1973: 322; Katz 1985: 329)

This is one of the few loan etymologies attested in the Samoyed languages only. Despite its limited attestation, this word cannot be a very recent borrowing, as the Samoyed word clearly reflects a diphthong, pointing to an archaic Iranian source language (although PIE/PII diphthongs were retained longer on the Iranian side than on the Indo-Aryan side,

by Middle Iranian times the Iranian diphthongs had turned into long vowels, and this is reflected also in the loans of the Ob-Ugrian languages, at least). It might be plausible to assume that the word was acquired from an “Old Iranian” language (if this kind of term is suitable for the unattested Iranian languages) sometime during the first millennium BC, before Proto-Samoyed split up. Probably the words for ‘bear’ and ‘flour’ belong to the same layer. All of them being important cultural words (the bear being important in mythology), they do not point to any intensive contact between Iranian and Samoyed but can be results of sparse trade contacts.

An Indo-Iranian etymology

**FT** *tšiet* ‘aus dopnelt gelegten, gekochten Birkenrindenscheiben zusammengenähte Zeltdecke’, **Kam** *t’è* ‘Zelt aus Birkenrinde (die Rinde zusammengenäht)’

< PS\**tāt* ‘tent roof made of birch bark’ < Pre-S \**tāši*

← PII or PI \**ta(é)šya-* ‘to be formed (out of wood etc.)’, root \**tačš-*, **OI** *takš* ‘zimmern, behauern, hämmern, härten’ (EWAIa I: 612–613, s.v. *TAKŠ*, 613–614, *takšan-*)

(Katz 2003: 239; Kallio 2015b: 82, footnote 4)

The etymology of Katz is convincing on both semantic and phonological grounds, and it has been supported by Kallio (2015b). The word is clearly an old loan, predating the sound changes that led to Proto-Samoyed, so it probably belongs to the same layer as PS \**wātā-* ‘to grow’ and perhaps \**jäā* ‘flour’. As is the case with \**wātā-* (see the entry), Samoyed \**t* could reflect either \**s* or \**š*, but because of Indo-Iranian \**š*, it is probable that the Pre-Samoyed form had \**š*.

An Indo-Iranian etymology

**Md E** *ukso*, **M** *uks* ‘*Ulmus campestris* od. *scabra* od. *foliacea*’; **Ma W** *oškâ* ‘Pappel’, **E** *oško* ‘Schwartzpappel, *Populus nigra*; (Ramst.) Esche’

< PU ?

← ? Indo-Iranian (unattested), cf. ? PG \**aska-* (> ON *askr* ‘ash’), < PIE \**h<sub>3</sub>es-ko-*, Albanian *ah* ‘beech’, Armenian *haci* ‘ash’ < PIE \**h<sub>3</sub>es-k-*, Lith *úosis* id. < PIE \**Heh<sub>3</sub>-s-io-* (Kroonen 2013: 38, s.v. *aska-*)

(Jacobsohn 1922: 15–16, 54; FUV: 138–139; Joki 1973: 333; Katz 1985: 130; 2003: 114; K. Häkkinen 1987; Bereczki 1992: 93; Blažek 2018: 25–27)

This etymology is included here only because it has been repeatedly been claimed that the word is borrowed from Indo-Iranian, although this claim does not have any substance.

Joki's statement "Entlehnung au seiner altertümlischen iran. (skyth.) Sprachform wäre ebenso möglich..." (based on Jacobsohn) is absurd, as the word is not attested in a single Iranian language. Katz mentions that the reflex of the IE word has disappeared from Indo-Iranian (he assumes that the word is an Indo-European tree name), and considers the Mari and Permic words to be borrowed from an IE form that is formally similar to the Germanic word. It remains unclear whether Katz assumes that the Uralic words are borrowed from a form of Indo-Iranian or from some other Indo-European branch.

Kroonen (2013: 38) calls the Indo-European word a "European tree name", meaning clearly that the distribution of the word is European, though he reconstructs the root for Indo-European. If the word indeed reflects a PIE root, then it is possible that it once was reflected in Indo-Iranian as well, but it is impossible to prove this.

It might be possible that this word represents a parallel substrate borrowing in Mari, Mordvin and several branches of Indo-European.

Not an Indo-Iranian etymology

### Ud oš 'ox, bull', Ko eš, eška id. < PP \*oška

< Pre-P \*uška

← PII \*ukšan- or PI \*uxšan- > OI ukšan-, nom.sg ukšā, OAv uxšan-, nom.sg. uxš-ā 'Stier' (EWAia I: 210, s.v. ukšan; Gotō 2013: 38)

(FUV: 134–135; KESK: 213; Joki 1973: 334; Katz 1985: 242; 2003: 195–196; Lushnikova 1990: 43, 270; Bereczki 2013 s.v. üškâž)

This is a well-known etymology which can hardly be doubted. While the word for 'ox' is attested in numerous Indo-European languages, the RUKI development \*š makes it clear that the Permic word was borrowed from Indo-Iranian. The word is not attested in Balto-Slavic, the other "RUKI language" in close contact with Uralic.

A more precise dating of the borrowing is difficult. The word has to be early, as it was borrowed into Permic before the characteristic vowel changes and the apocope of last syllables. The lack of cognates in other Uralic languages could point to a relatively late borrowing from Proto-Iranian or some "Old Iranian"-type language, but it is interesting here that the cluster PII \*kš/PI \*χš has retained the velar stop in Permic: a borrowing from a spirantized Iranian form \*χš might have resulted in a form \*\*uša or the like, as a velar spirant could easily have been left unsubstituted.

According to Lushnikova, the word has to be borrowed from the Indo-Iranian nominative form which ends in a vowel, as otherwise we would not expect a Uralic vocalic stem. This is true also of other *n*-stems borrowed into Uralic. The nominative is probably the most obvious donor form, but the lack of *-n* can also be explained with Uralic phonotactics, as *n*-stems did not necessarily exist in the era of Pre-Permic into which this word was borrowed, and in any case \**a*-stems were much more frequent.

It is a well-known fact that Mansi *üška* was borrowed from Komi, but it has only recently been argued by Metsäranta that also Mari *üškâž*, *üšküž* (PMa < \**üškəž*) is also a borrowing from Permic. Here both the Mari vocalism and the metathesis of the consonant

cluster point to a Permic origin. For parallels for an Iranian loan being borrowed from Permic into Mari, see \**ačwa* ~ \**očwa*.

FUV also mentions Fi *ehko*, *hiehko* here with a question mark, but any relationship between them and the Permic words is completely impossible.

An Indo-Iranian etymology

**Fi *udar* (: *utaret*), *utare* (cognates in all Finnic languages) ‘udder’; Md E, M *odar* id.; ? Ma *waðar*, *woðar* id. (< PMa \**wáðar*)**

< PU ? \**utarV*

← PII \*(H)*ūdhar*-‘udder’, > OI *ūdhar*- (~ *ūdhan*-) ‘udder’ (EWAia I: 240–241, s.v. *ūdhar*-)

(FUV: 138; Joki 1973: 332–333; Katz 1985: 245–246; 2003: 198–99; Lushnikova 1990: 269; 2000: 178, 193; Bereczki 1992: 95; 1994: 98–99)

Helinski (2000: 193) lists this word as one of the Uralic (Finno-Permic?) words with a consonantal stem (see also \**sasarV* above). He considers the dialectal Finnish forms such as *utare*, *utara* and *utaro* secondary.

There is no reason to reconsider our ideas of possible Uralic stem shapes on the basis of this word. What is more, there is no certainty that this is a genuine Indo-Iranian borrowing, as also a Germanic origin (PG \**ūdara*-, \**ūdira*-, Kroonen 2013) has been suggested, and since cognates of the Germanic and Indo-Iranian words for ‘udder’ are widespread among the Indo-European languages, it is difficult to determine the exact origin of this word. One (not decisive) argument in favour of the Indo-Iranian origin would be the word’s semantic sphere: there are relatively many (?) Indo-Iranian loanwords dealing with cattle terminology, but on the other hand, the Germanic loans in Finnic connect to so many different semantic spheres and cultural aspects, that this is not a very convincing argument.

SSA considers the Uralic words cognates and postulates an Indo-Iranian etymology. (SSA cites the OI form and states that the word was borrowed from Proto-Iranian. It remains unclear why we should point exactly to PI as the source language, but perhaps that is because of the limited distribution). The UEW (No. 1671) considers the word either a PII or PI borrowing.

Also a Germanic etymology has been suggested, but LÄGLOS (s.v. *utare*) is reluctant to accept it because there are only few tentative Germanic etymologies that have a distribution stretching from Finnic to Mari and Mordvin. But LÄGLOS’s reluctance is difficult to understand, because it would be a very viable option to differentiate between the word in these three branches. They cannot go back to a Uralic proto-form, and it seems obvious that we are dealing with several parallel borrowings here. The Finnic word may well have been borrowed from Germanic. The Mordvin word could be a cognate of the Finnic one, but there is no obstacle to assuming that it represents a separate borrowing from Indo-Iranian.

Bereczki (1992: 95; 1994: 98–99) has considered the Mari word a separate borrowing from Indo-Iranian, and he is most certainly right as the Mari and Finnic words cannot be derived regularly from a unitary Proto-Uralic form. Helimski (2000: 172) notes that the Uralic word has to be reconstructed with *\*t* because of the Mari word (PU *\*δ* would regularly disappear in Mari), but if the Mari word is a parallel loan, the Finnic and Mordvin words could be derived also from a West Uralic form *\*uδarV*.

Mari *\*w-* is unexpected here, as there is usually no prothetic *\*w-* in Mari words inherited from Proto-Uralic. It is difficult to explain the Mari *\*w-* on the basis of the Indo-European original – perhaps the glide + *u* sequence is a substitution for the long *\*ū* due to the lack of long vowels, or perhaps the Indo-European pronunciation of the glide *u/ū* has been so consonantal, that *\*w* was a suitable substitution. Katz has argued that *w-* is a substitute for a laryngeal, but there are no convincing parallels to such substitution and it would not make much sense phonetically. Bereczki (1992: 95) vaguely notes that the explanation for the *\*w-* has to be sought on the Iranian side, as it cannot be explained as a secondary Mari development.

Katz also pays attention to the divergent forms in the Uralic languages, which make the reconstruction of a unitary proto-form impossible.

? Indo-Iranian etymology (probably parallel borrowings from various Indo-European languages)

**Fi *vasara* ‘hammer’; SaN *veahcir* id. (< PSa *\*veacērē*; cognates in all Saami languages; Lehtiranta 2001, No. 1367) id.; Md E *užer*, *vižir*, M *užar* ‘axe’**

< ? PWU *\*wačara*

← PII *\*wážra-* > Av *vazra-* ‘club’, OI *vájra-* ‘thunderbolt, Indra’s weapon’ (EWAia II: 492, s.v. *vájra-*)

(FUV: 139; Joki 1973: 339; Gamkrelidze & Ivanov 1984: 928; Katz 1985: 336; 2003: 267; Lushnikova 1990: 33, 283–284; SSA s.v. *vasara*; J. Häkkinen 2009: 23–24; 2012: 7; Parpolä 2012: 161; 2015: 63)

This word is a well-known Indo-Iranian etymology, and there is no reason to doubt the ultimate Indo-Iranian origin of the Finnic, Saami and Mordvin words. J. Häkkinen (2009: 23–24) has claimed that this is one of the words which show the rare, but in his opinion regular, correspondence of Fi *a* and Mo *i*. This is, however, a ghost-correspondence, as Häkkinen’s other examples of the same vowel correspondence are not convincing: *marras* and *mird’e* are not true cognates but parallel loans, and *\*wäškä* is a Wanderwort which cannot be regularly reconstructed for Proto-Uralic. In another article (J. Häkkinen 2012: 7–10) it is claimed that the words can represent parallel borrowings. It seems, however, that the Erzya dialectal form *vižir* can be explained from the same Pre-Mordvin form *\*wačara*, and also in Saami the development *\*a* > *\*ea* is more or less regular before *\*ć*. There is thus no reason not to reconstruct a West Uralic *\*wačara* and assume a borrowing into this or some earlier proto-language. Because of *\*ć*, the word cannot be a later Iranian loan, but it has to be borrowed from Proto-Indo-Iranian.

As a side note it can be mentioned that this etymology is an interesting indication of cultural influence from Indo-Iranian on (West-)Uralic: in Indo-Iranian the word *\*waǰ́ra-* is clearly associated with mythology, and it can be assumed that the word has been borrowed into Uralic along with other words relating to mythology (despite its current status as a more ordinary term for a tool), such as *\*juma*.

Joki (1988) notes that in Finnic the meaning ‘axe’ is lost because of a later borrowing from Baltic, *kirves*. Mordvin would have retained the earlier meaning, according to Joki. Parpola (2015: 63) argues that the meaning ‘hammer’ in Finnic and Saami can be due to later Scandinavian influence, as the hammer was an important weapon of the Scandinavian storm god Thor.

Tocharian A/B *wāšir* is also borrowed from the same Indo-Iranian source (Zimmer 1999: 145 notes that the word is ultimately from Indic, but perhaps borrowed via Iranian).

An Indo-Iranian etymology

**Fi *vatsa* ‘stomach’, Ka *vattša*, Li *vatsā* id. (cognates in all Finnic languages);  
? Ms North *was* ‘stomach’**

< ? PU *\*waćća* (UEW s.v. *\*waća*)

← PII *\*watas-* ‘year’, *\*watsá-* ‘calf’ > OI *vatas-* ‘year’, (der.) *vatsá-* ‘calf’ < PIE *\*wet-só-*, *\*wet-os-*

(Parpola 2017)

This etymology involves the same Indo-Iranian donor form as the etymologies for *\*wasa* and Mansi *\*wěšəγ* ‘calf’ (see the respective entries below). Koivulehto (1999a: 218) had also derived Fi *vuosi* and its Uralic cognates from the same Indo-Iranian origin, but this has been convincingly refuted by Aikio (2012a: 233–234).

Parpola’s etymology is not very convincing, as it involves semantic problems and phonological substitutions that are not very probable. Also the relationship with the forms in various Uralic branches is irregular, and it is not at all probable that a word *\*waća* or *\*waća* can be reconstructed for Proto-Uralic. Parpola rightly refutes the UEW’s reconstruction *\*waća*, as this is an impossible pre-form for Finnic *vatsa* (*\*ć* would yield Finnic *s*). The connection of the Mansi word *wać* ‘stomach’ with the Finnic word is very doubtful. Parpola is confident about their relatedness because of the semantic nearness, even in spite of the fact that the Mansi word is found only in one dialect. The Mansi word’s sibilant is irregular: usually PU *\*ć* gives Mansi *s*, although there are apparent exceptions to this in the material of the UEW (see the entry *\*ićä*). The re-evaluation of the development of *\*ć* in Mansi and the other Ugric languages would be an important task for Uralic historical phonology.

Parpola does not pay attention to the Mansi sibilant, but he notes that the vowel development is uncertain, as various reflexes of PU *\*wa-* are found in Mansi, with PU *\*wa-* > PMs *\*u-* being the most common. Parpola also mentions examples of other reflexes of assumed *\*wa-* (> North Mansi *ā, ə, o*) but these are based on the evidence of the UEW, and when one takes a look at the listing of Uralic *\*a-a* and *\*a-i*-stems in Aikio (2015b), it is

easy to see that these developments are based on incorrect reconstructions or improbable etymologies. Mansi *wāyl-* ‘to descend’ is from Proto-Mansi *\*wēyl-* which ultimately reflects PU *\*wīlka-*. PU *\*wajši* ‘water-bird’, allegedly the pre-form of North Mansi *wās*, is missing from the listings of Uralic *\*a-i-*stems in Aikio (2015b) and Zhivlov (2014), which means that the UEW’s etymology is probably not correct, at least in the form that it is presented. And finally, North Mansi *at-wās* ‘silver’ is the reflex of PU *\*wāškä*, which has recently been considered as an early Wanderwort by Aikio (2015b: 42). It is thus probable that *\*wa-* > *\*u-* is indeed the regular development, which means that *wac* cannot be a cognate to the Finnic word.

Parpola (footnote 11) also mentions the possible Mordvin cognates *E vača*, *M vačo* ‘hunger, hungry’, noting that they cannot regularly reflect a form with *\*ć*. The relation of the Mordvin words to Finnic *vatsa* remains uncertain (the resemblance of the words could be simply accidental), but it does not have much relevance for the possible Indo-Iranian etymology.

Parpola argues that the Indo-Iranian/Indo-Aryan cluster *\*-ts-* was substituted by Uralic geminate *\*-ćć-*. It is true that no *\*-ts-* cluster can be reconstructed for Proto-Uralic, but it is uncertain whether *-ćć-* would be the expected substitution here. Proto-Uralic also possessed the affricate *\*č*, and one can assume that this would have been phonetically closer to the sequence *\*-ts-*. Although Parpola is confident about the reconstruction of the geminate affricate *\*-ćć-* for Proto-Uralic, there are actually not many examples supporting this. Sammallahti (1988) lists no cases of Proto-Uralic or Proto-Finno-Ugric *\*-ćć-*. Parpola refers to Koivulehto (1999a: 219), Kallio (2007: 233) and Lehtinen (2007: 95–97) who argue that Finnic *-ts-* can reflect the earlier geminate affricate *\*-cc-*. While this is correct, it does not automatically mean that this Pre-Finnic affricate goes back to Uralic *\*-ćć-*. More evidence would be needed to back up this substitution.

Semantically the etymology is a bit far-fetched. Parpola is aware of the semantic difficulties, but he argues that as the Sanskrit word means also ‘child’, and the Finnic word denotes also ‘womb’, the difference between the meanings is not so big after all. This explanation would be more convincing, if parallels for the semantic shift ‘child’ > ‘womb’ or vice versa could be presented. Parpola also mentions the semantic development from ‘year’ to ‘pregnancy’ as a possible connection between the Indo-Iranian and Finnic words. Here one has to note that because of the very large corpus of various types of Old Indo-Aryan texts, it is relatively easy to find several meanings for one and the same word. As a specialist on Indo-Aryan, Parpola does refer to the meanings found in the Vedic texts already: the meaning ‘child’ for *vatsá-* is found already in these earliest texts, but is still doubtful whether we can reconstruct it for Proto-Indo-Iranian. *DatSemShift* (s.v. *womb*) does not list any example of the development ‘child’ > ‘womb’.

Parpola’s idea that the word would be borrowed specifically from the Indo-Aryan branch is based mostly on his views on the archaeological cultures connected with the spread of early Indo-Iranian speakers, but he also argues that *\*ts* could only result from Indo-Aryan *\*ts* or its Iranian cognate *\*s*. This is probably correct, but as there are various problems with the etymology and it is not at all clear that the Uralic words in question point to a borrowing from a form akin to Indo-Aryan *\*watsá-*, this etymology does not bring valuable new evidence for the idea of Indo-Aryan borrowings in the Uralic languages. Kümmel (2018a) actually argues that *\*ts* should still be reconstructed for Proto-Iranian and the change *\*ts* > *\*s* is a later (Common Iranian?) development, and thus this word

could also be an Iranian loan. Kümmel does not comment Parpola's etymology in more detail.

Not an Indo-Iranian etymology

**SaN *váža* 'reindeer cow' (cognates in all Saami languages except S; < PSa \**wāćem*)**

< ? Pre-Sa \**waćim*

← PII \**waćā-* 'cow' > OI \**vaśā-* 'cow' (EWAia II: 528–9, s.v. *vaśā-*)

(Koivulehto 1999a: 217 = 2016: 219; Sammallahti 2001: 407; Holopainen 2018b: 166–167)

This etymology is one of Koivulehto's Indo-Iranian etymologies from the late 1990s/early 2000s and one of the alleged early loanwords that have a distribution confined to Saami. As has been shown by Holopainen (2018b: 166–167), this etymology involves phonological problems, and the same can be said of many of the other Indo-Iranian etymologies which are found only in the Saami languages. The arguments will be repeated here.

The main problem with the Indo-Iranian etymology for this Saami word is the Saami vowel combination PSa \**a–e*, which is not found in inherited words, and according to Aikio (2009) this means that the word cannot go back to a period older than Proto-Saami. As Indo-Iranian loanwords have to be older than the relatively recent Proto-Saami, there is no possibility to consider this word an Indo-Iranian borrowing. To avoid circular reasoning, it has to be stressed that there are no other possible examples for an Indo-Iranian loanword with this vowel combination in Saami, and the general views of Saami linguistics prehistory (see especially Aikio 2012b) situate Proto-Saami in Southern Finland and Karelia, whereas Proto-Indo-Iranian was spoken in a territory that was far from this in both space and time. Furthermore, this word cannot be a later Iranian word because of the palatal affricate.

Regarding the etymology of the Indo-Iranian word, it is unclear whether the word goes back to Proto-Indo-European. Latin *vacca* 'cow' is usually considered a cognate, but because of the irregular geminate this cannot be considered certain. De Vaan (2008: 649, s.v. *vacca*) also points out that the Latin vocalism does not fit the Indo-Iranian vowel (the Latin should reflect \**o* and the Indo-Iranian word \**e*, although it has to be said that both could also reflect Indo-European \**a*, although this is not apparently accepted by de Vaan as a follower of the Leiden school of Indo-European reconstruction). It is possible that Latin and Indic words are parallel borrowings from same or related sources. The Indo-Iranian word could still be the origin for the Saami word in spite of these remarks, but the limited and irregular Indo-European distribution makes this assumption more unlikely.

It must be mentioned that Koivulehto also noted that the word could have been borrowed already from a Proto-Indo-European or Pre-Indo-Iranian form which still had \**k´*. However, there is simply no reason at all to support this view, and since the word is probably not an Indo-European borrowing at all, this word offers no evidence for the substitution of \**k´* by Uralic \**ć*. For general criticism of the loans that show alleged reflexes of PIE palatal stops, see Holopainen (2018a).

Not an Indo-Iranian etymology

**Fi** *voi* ‘butter’; **SaN** *vuodja* ‘butter, grease’ (< **PSa** *\*vōdje*); **Mo** **M** *vaj* ‘fat, butter, oil’; **Hu** *vaj* ‘butter’; **Mns East** *vāj* ‘fat’; **Kh East** *voj* ‘fat, butter, oil’

< **PU** *\*waji* ‘butter’

← **Pre-II** *\*(H)ágyas-* > **PII** *\*(H)ázyas-* > **OI** *áñjas-* ‘Salbung’, root *añj* ‘salben, schmieren’ < **PII** *\*Hang/j* < **PIE** *\*h<sub>3</sub>ong<sup>w</sup>* (EWAia I: 54, s.v. *AÑJ*; Werba 1997: 333–334, No. 379)

(Koivulehto 1999a: 217–18 = 2016: 219–220)

This etymology has been convincingly refuted by Aikio (2012a: 236): Aikio argues that the initial *w-* and the reconstructed PU vocalism make the Indo-Iranian etymology extremely unlikely. Also the “Pre-Indo-Iranian” form postulated by Koivulehto is problematic because of the reconstruction that involves an “intermediary” form of the second palatalization: *\*g’*. This means that the word should have been borrowed before *\*g* became an affricate, but at a time when it was already palatalized, and this sound would have been substituted by Uralic *\*j*. There are no parallels to the substitution of this intermediary stage, and also the substitution of PIE *\*g’* by *\*j* is supported by very few convincing examples (see *\*juyi-*, *\*jäyi-*).

Czentnár (2015: 9–17) has argued that the Uralic word has been borrowed into various Turkic languages (such as Turkish *may*, Crimean Karaim *may*, Uzbek *moy*).

Not an Indo-Iranian etymology

**Md E** *vano-*, **M** *vanâ-* ‘to watch; observe, watch over’

< **PU/PWU/Pre-Md** *\*wajna-*

← **PII/PI** *\*wayn-a-* > **Av** *vaēna’ti* ‘sees’, **OP** *vaināmiy* ‘I see’; **OI** *vénati* ‘watches, observes’, root *\*wayn-* (EWAia II: 582–83, s.v. *VEN*)

(Koivulehto 1999a: 229, 230 = 2016: 231, 232)

This is a convincing etymology, and clearly a relatively early loan despite the limited distribution of the word in only Mordvin, as both the vowel changes leading to Proto-Mordvin and the retained Iranian diphthong mean that the word cannot be a late loan into Proto-Mordvin. There is no other competing etymology for the Mordvin word.

Koivulehto reconstructs the Pre-Mordvin word as *\*wa(j)na*, but it is most logical to assume that the word was borrowed with the diphthong, which then became lost in Mordvin (this is the assumption of Koivulehto as is evident from the text, but his choice of reconstruction could mean otherwise). For the development of the Mordvin *j*, see also the entry *\*pejmä*.

Koivulehto notes that SaN *oidnit* is a parallel borrowing from the same Indo-Iranian source. This is convincing, and the loans have to be parallel since the Saami word reflects \**o* (or in Aikio's newer reconstruction \**a-i*; in any case the Saami and Mordvin words have to be parallel borrowings).

An Indo-Iranian etymology

**Fi *vanka, vanko, vanku* 'beim Schwenden als Werkzeug verwendete lange Stange mit einem Haken an der Spitze'; Ko *vug* 'Griff, Handgriff, Handhabe, Henkel, Ohr', Ud *vug* 'Bogen, Henkel (an Gefäßen)' (< PP \**vug*)**

< ? PU \**wan̄ka* ~ \**wīnka* (UEW s.v. *wan̄ka*; Zhivlov 2014: 143)

← PII/PI \**wanka* > OI *vānka-* 'crooked', OI root *vānc/vank* 'wanken' < PII \**wank-* < PIE \**wenk-* (EWAia II: 492–493, s.v. *VAÑC*; Werba 1997: 232 No. 172)

(FUV: 139; KESK: 69; Joki 1973: 336–336; UEW s.v. *wan̄ka*; Katz 1985: 344; 2003: 273; SSA s.v. *vanko*; Koivulehto 1999a: 227 = 2016: 229; ms.)

The idea of the Indo-Iranian etymology of this word is old, and phonologically convincing, although in Indic no meaning of 'hook' as a noun is attested. The meaning attested in Finnic and Komi can probably be derived from 'crooked'. As the root \**wenk-* is widely reflected in Indo-European (such as Germanic \**wan̄ha-* > OE *wōh* 'crooked'), theoretically some other donor language could also come into question, but because of the distribution in Finnic and Permic the Indo-Iranian source is the most plausible. As can be seen from the reconstruction of Zhivlov, the Uralic form could reflect either \**wan̄ka* or \**wīnka*.

Estonian *vang* '(Fluß)biegung' and Livonian *vanga* 'Bachwiese' have been connected to Finnish *vanko* in earlier research, and this connection is still mentioned by the SSA. LÄGLOS separates these and gives a Germanic etymology, \**wanga-z* (> ON *vangr* 'Feld, Flur, Wiese') (LÄGLOS III s.v. *vang*). The Ob-Ugric word \**wā̄nkVrV* 'Haken' was possibly borrowed from the same Indo-Iranian word as is assumed here (see the entry).

? Indo-Iranian etymology

**Fi *vuori* 'mountain, rocky hill'; Ko *ver* 'forest', Ud *vir* 'hill, highlands' < PP \**vór*; Ms East *wōr* 'evergreen forest'; Kh East *wor* 'ridge along a river'**

< PU \**wari* 'forest; hill' (Aikio 2012a: 233–234, 236)

← PII (unattested) \**áras-*, cf. Greek ὄρος 'mountain' < ? PIE \**h<sub>3</sub>eros-* (Beekes 2010: 1109–1110, s.v. ὄρος)

(Katz 1985; Koivulehto 1999a: 218 = 2016: 220)

The Indo-Iranian etymology for this Uralic word has been suggested independently by both Koivulehto and Katz. Koivulehto's etymology has been convincingly rejected by Aikio (2012b). Aikio's main argument is that the Proto-Uralic form requires the reconstruction

of initial *\*w-*, which cannot be explained by the Indo-European etymology. Here it is important to note that Aikio's new reconstruction of the word was not yet known to Koivulehto, who operated with different (now outdated) rules of Uralic phonology and was thus able to explain the etymology by the standards of his time (the late 1990s). However, as also Koivulehto himself noted, the Indo-European word for 'mountain' does not have a cognate in Indo-Iranian and Koivulehto only postulates a Proto-Indo-Iranian form of this word to account for the loan in Uralic.

A further argument against any kind of Indo-European origin for the Uralic word is that according to Beekes (2010) the Greek word is a Mediterranean substrate word, although Beekes also mentions the possibility that the word is from the IE root *\*h<sub>3</sub>er-*, an explanation which is mentioned by earlier dictionaries (Frisk 1954–1972 ὄρος; Chantraine 2009 s.v. ὄρος). This root has a possible Indo-Iranian reflex in OI *r̥sua-* 'skin', but this is of course semantically and phonetically far from the Uralic word.

Not an Indo-Iranian etymology

**Ko *vurun*, *virin* 'wool' < PP *vurun* (Csúcs 2005: 398)**

< Pre-Permic *\*warna* ~ *wirna*

← PII/PI *\*Hwr̥naH* > OI *ūr̥nā* F 'wool', Av. *varənā-* 'animal wool', Chwarezmian *w'n* 'wool'; < PIE *\*h<sub>2</sub>wl̥-h<sub>1</sub>no-* / *\*h<sub>2</sub>wl̥-h<sub>1</sub>neh<sub>2</sub>* (Cantera 2001: 15; Peters 1980: 23, footnote 18; EWAia I: 243, s.v. *ūr̥nā*; Pinault 2016: 252–253)

(Joki 1973: 348–9; KESK; Rédei 1986c: 81; Lushnikova 1990: 277; Csúcs 2005: 398)

Although this word is limited to a single Uralic language, the Komi vocalism shows that the borrowing must be relatively old, that is, it must predate the Proto-Permic stage. Based on solely phonological arguments, the Komi word could reflect a Proto-Uralic word. In Udmurt wool is called simply *yš-gon* 'sheep-hair'. It is difficult to explain the loss of the word for 'wool' from Udmurt, if it has not been replaced by a loanword but a transparent compound. Joki reconstructs the pre-form (? Pre-Permic = PU) of the word as *\*varne*/*\*vorne*, but this represents outdated views on Permic vocalism. The Komi word could reflect an earlier *\*i-a* or *\*a-a*-stem (Sammallahti 1988: 553–554). Thus it could reflect a similar development as the word *\*tirna* 'grass', another well-known Indo-Iranian borrowing. Uralic *\*ir* is a regular substitution for PII/PI *\*r̥*. Gamkrelidze & Ivanov (1984: 928) list the word among the *o*-loans, assuming a substitution of Iranian *\*a* as Uralic *\*o* before *\*r*. However, it is uncertain whether the donor form had *\*a* or *\*r̥*, and the Komi word cannot reflect Uralic *\*o*. Gamkrelidze & Ivanov are thus clearly mistaken, and this word offers no evidence for their *r*-rule. In connection with this word the Avestan *a* and Indic *ūr̥nā* point to a zero-grade form according to Cantera (2001: 15): in Old Indic the *w* has been vocalized, whereas in Avestan the epenthetic *ə* before *r* + consonant has become *a* when accented. Cantera (2001: footnote 19) also notes that Peters (1980: 23, footnote 18) assumes that the last syllable of *\*h<sub>2</sub>wl̥-h<sub>1</sub>neh<sub>2</sub>* was stressed, but the Indo-Iranian forms contradict this.

It is difficult to determine the exact origin of the word: the Indo-Iranian etymology is inevitable, as the semantic match is complete and the change of *\*l* to *\*r* in this Indo-European word can only point to Indo-Iranian, but it is unclear from which stage this word has been borrowed. As the word-initial *\*w* has been vocalized in Indo-Aryan, this would not be a suitable donor form, but one would hardly expect an Indo-Aryan borrowing in any case. Both PII and PI forms are suitable origins, though based solely on the distribution, one could opt for an Iranian borrowing, from PI or some early (Old) Iranian language. A later, Middle Iranian form is clearly out of the question because of the archaic appearance of the Komi word. The attested Middle Iranian words, such as Chwarezmian *w'n*, show that the *r* element has disappeared altogether. No reflex of the Indo-Iranian word is attested in Ossetic, making its existence in prehistoric Alanic uncertain.

The determination of the precise source and dating of the etymology are also made difficult by the complicated reconstruction of the Proto-Indo-European word for wool, for which various different proposals have been made. The latest attempt to reconstruct the PIE word is by Pinault (2016), who suggests PIE *\*h<sub>2</sub>wl̥-h<sub>1</sub>no-* and argues that the word was originally a compound with the meaning 'yearly wool'. However, in any case the Permian word has to be an Indo-Iranian borrowing, as Permian *\*r* could not reflect PIE *\*l*.

Komi (Letka dialect) *varneš* 'Schaf' (Joki 1973: 336; Rédei 1986c: 79) has to be borrowed from a later Iranian form from the same root (here the vocalism shows that the borrowing cannot be ancient, contrary to *vurun*).

Finnic *villa* 'wool' (< *\*wilna*) is a well-known borrowing from the Baltic cognate (*\*vilna*) of the Indo-Iranian word (SSA s.v. *villa*).

An Indo-Iranian etymology

**Fi *varsa* 'foal', Ve *varz* 'foal; (young) stallion' (cognates in Ka, Est, Li); Md *E vašo*, M *vaša* id.**

< West Uralic *\*warsa* ~ *\*wīrsa*

← Alanic *\*warsa-*, cf. Oss *wyrs*, *urs* < PI *\*wr̥šan-* (EWAia II: 576–577, s.v. *vr̥šan-*; Cheung 2002: 245)

(Setälä 1928: 306–307; FUV: 139; Joki 1973: 337–338; Harmatta 1977: 171; Katz 1985: 251–252; 2003: 202–203; Lushnikova 1990: 278; Koivulehto 1999a: 226–227 = 2016: 228–229)

This is a clear Iranian etymology, but both the limited distribution and especially the absence of the RUKI rule point to a late borrowing (cf. Ossetic *wyrs*, *urs* 'stallion'), as was observed already by Joki and supported by Koivulehto. Joki (1973) mentions that in earlier research this word has also been considered a very early borrowing (here the absence of RUKI has been interpreted as an archaism), and also Harmatta (1977: 171) supports a similar view, but since the RUKI rule has reflexes in the earliest Pre-II loans (such as *\*mekši*), it is anachronistic to consider this an early loan. Also, the distribution points to the same conclusion. This probably belongs to the small layer of relatively late Old Iranian/Pre-Ossetic borrowings shared by Finnic and Mordvin (as well as Saami), meaning

that the word was acquired into the West Uralic proto-language shortly before the spread of its western dialects into the areas surrounding the Baltic Sea. Although there are only a handful of such borrowings, many of them are related to horse/cattle terminology, showing obvious cultural influence from Iranian pastoralists.

Also KEWA mentions the etymology but does not comment on the age of the borrowing in any way. The Finnic and Mordvin cognates are regular cognates, meaning that the word was probably borrowed into the West Uralic proto-language. However, although the vowel of the word has been usually reconstructed as *\*a*, according to modern research of Aikio (2015b) both Finnic *a–a* and Mordvin *a–a*-stems could also reflect a historical *\*i–a*-stem. It is known that even in the earliest borrowings Indo-Iranian *\*a* was often substituted by Uralic *\*i*, and in this word the high central vowel would be a fitting substitution for the Alanic vowel that later gave Ossetic *y*, *u*. This substitution would be even more logical than *\*a*, as the Ossetic vowels have developed from PI *\*r*, and there is no reason to postulate a Pre-Ossetic form with *\*a*.

Katz (1987: footnote 9) has suggested a Germanic origin for the Finnic word, deriving it from PG *\*farza-* ‘young ox’ (OE *fearr*). As the Finnic word has a cognate in Mordvin, the idea that *\*warsa* would be a Germanic borrowing is not very convincing. Katz’s etymology is refuted by LÄGLOS (s.v. *varsa*), where it is noted that the traditional Iranian etymology is convincing, and the idea that Proto-Germanic *\*f-* could be reflected by Finnic *v-* is not probable, as there are no convincing parallels. Katz also notes that Rédei’s Indo-Iranian etymology is problematic due to absence of the RUKI rule. An alternative Indo-Iranian etymology is suggested by Katz (1985, 2003), who assumes that the Uralic word is borrowed from *\*watsatará-* (> OI *vatsatará-* ‘das entwöhnte Junge’), which is in turn derived from PII *\*watsa-* ‘calf’. Here the Uralic word would result from syncope and metathesis *\*wasara* > *\*wasra* > *\*warsa*. This explanation is very complicated, and it is much more probable that West Uralic *\*warsa* is a loan from an early form of Alanic.

Johnny Cheung has derived the Permian words for ‘stallion’, Ko, Ud *už*, from this same Iranian word, but this etymology is phonologically problematic: it is unexpected that the cluster *\*rs* (or *\*rš*) would be simplified in Permian to *ž*. There is a competing and more probable Iranian etymology for the Permian word (see the entry on *už* < *\*ačwa*).

An Indo-Iranian etymology

### **Finnic *vasa* ‘elk calf’ (der. *vasikka* ‘calf’); Md *vaz***

< West Uralic *\*wasa* ~ *\*wiša*

← PI *\*wasá-* (< PII *\*watsá-*) > Oss *wæs* ‘calf’, Chwarezmian *wsyk* /*wasik*/ id., Khotanese *basaka* (EWAia II: 495, s.v. *vatsá-*)

(FUV: 139; Joki 1973: 338–339; Mayrhofer 1984: 251; Teubner 1974: 304; Katz 1985: 244; 2003: 197; Lushnikova 1990: 281–282; Koivulehto 1999a: 225, 227 = 2016: 227, 229; Parpola 2017: 265–271)

This is a convincing Iranian etymology, and the correspondences between the Finnic and Mordvin words are regular, pointing to inheritance from a common proto-language (or a

very early parallel borrowing, which is impossible to prove). The word can be borrowed from Proto-Iranian or from a later Iranian idiom also ("Pre-Ossetic", as Koivulehto argues), mainly because of its distribution. \*s clearly point to an Iranism. The word could perhaps be already from PI but probably belongs to the same, relatively late layer of Iranian borrowings limited to West Uralic like \*warsa (see above). Many of the loanwords in West Uralic deal with horses or cattle, and \*wasa would fit into this well. The Iranian form goes back to PII \*watsa-, which is derived from an IE s-stem \*wét-os- 'year' (from the root \*wét- 'Jahr; jährlich'), thus originally denoting a one-year-old calf (Stüber 2002: 188–189).

As with the etymology of warsa (see above), here it is impossible to decide whether the Finnic and Mordvin words reflect earlier \*wasa or \*wiša, as in -a-stems the vowels \*a and \*i merged in both branches.

It has also been argued by Joki (1973) and Katz that the word has been borrowed from PII \*watsa- (cf. OI vatsa-) and the -s- is the substitute for the PII cluster -ts-; this remains a possibility, but due to the lack of other elements it is difficult to prove. Since there are other clearly Iranian borrowings with similar western distribution, I would rather consider this a borrowing from Proto-Iranian or a later Iranian language. Kümmel (2018a) notes that \*ts was still possible in Proto-Iranian, and this word can thus be considered a later borrowing. This might belong to the same layer as \*wasa above.

An Indo-Iranian etymology

**Ng bada'ama (tr.aor.sg.1.obj) etc., En barábo (intr.aor.sg.1), NT βᾶδᾶ (tr.) etc., NW βᾶτᾶς (tr.), Slk keŕiŕo (tr.), Km budəl'am (prs.sg.1.), Mt бадьямъ (prs.sg.1.) < PS \*wātā- (tr.) 'erziehen, ernähren', \*wātā- (intr.) 'to grow' (SW: 172)**

< Pre-S \*waksa- or \*wakša-

← PII \*Hwakš-a- or PI \*Hwaxš-a-, root \*Hwakš- 'to grow', OI vakš- 'to grow, become big', perf. vavákša, caus. vakšay 'to let grow big'; OAv vaxš-, uxšieitī 'grows (ints.)', vaxšt 'grows (trans.)', Khot. hušš- 'to grow', Sogd. xwš- 'to grow' < PIE \*h<sub>2</sub>eug-s- (EWAia II: 485–487, s.v. VAKŠ; Werba 1997: 452–453; Cheung 2007 s.v. uaxš)

(Aikio 2002: 52; J. Häkkinen 2009: 25)

Aikio's etymology is significant because it proves that there are independent Indo-Iranian loans in Samoyed that were clearly acquired *before* the Proto-Samoyed sound changes took place. This word has to be quite old then, phonologically either Proto-Indo-Iranian or Proto-Iranian would come in to question. It is unclear whether PI \*xš would actually result in \*kš or whether the spirant \*x was too "weak" to be substituted, but a form like Pre-Samoyed \*waša would also result in \*wātā. From the Iranian point of view the word does not have to be a very early borrowing, as in later Iranian languages such Middle Persian (/waxš/) suitable forms are found, but because of the sibilant change the word has to have been borrowed into Pre-Samoyed relatively early.

Semantically the etymology is also plausible. It is uncertain whether the word was borrowed as a transitive or intransitive verb into Samoyed.

An Indo-Iranian etymology

**Mansi East *wōt*, North *wōt*, *wot*-, South *wāt*, *wat*- (< PM *\*wātə* ‘wind’; *\*wāt*- ‘wehen’); Khanty North *wat*, East *wat*, South *wot* (< PKh *\*wāt*) ‘wind’ < POug *\*wāta* (Zhivlov 2006: 139–140; Honti 1982: 198, No. 724)**

← **Iranian *\*(H)wa(H)ata-* ‘wind’, cf. OI *vāta-* ‘Wind, Windgott’, OAv *vāta-* ‘Wind’, Sogdian, Ossetic < PII *\*HwaHata-* root *\*HwaH-*; < PIE *\*h<sub>2</sub>ueh<sub>1</sub>-ŋ-to-*, root *\*h<sub>2</sub>ueh<sub>1</sub>-* (EWAia II: 542, s.v. *vāta-*; Werba s.v. *vā*)**

(Korenchy 1972: 84, No. 44; Lushnikova 1990: 283 No. 188; Helimski 1991: 222; Helimski 1997b: 121; Sipos 2002: 57)

This is a good and well-established Indo-Iranian etymology. However, the exact age of the loanword is extremely difficult to date, though it could be relatively old as well. The Avestan and Vedic reflexes of the PII *\*HwaHata-* are famous for preserving the “laryngeal hiatus” (see Fortson 2010; Kümmel 2004); the Ob-Ugric words do not show any reflex of the II laryngeal, but this does not mean that it could not have been still present in the donor form, as the reflexes of PII/PI laryngeals in loanwords are few and uncertain.

It is interesting that the Khanty and Mansi reflexes of this word clearly point to a back vowel (reconstructed as long *\*ā* by Zhivlov), as in many other, probably simultaneous, borrowings in Ob-Ugric the Iranian vowel *\*a* is reflected by *\*ä*. While this may be a chronological difference (the most ancient loans show *\*a* or *\*o* as the substitute for *\*a*, not *\*ä*), the different substitutes in Ob-Ugric might also point to synchronic dialectal differences in the Iranian donor languages. The most probable reason for the different substitutions would be that in different environments Iranian *\*a*, *\*ā* were substituted differently, but there is not enough evidence to establish these environments. On the other hand, it is possible that long *\*ā* was substituted differently. Also Ob-Ugric *\*ñātV* possibly reflects Iranian long *\*ā*.

Regarding the vowel developments, the Ob-Ugric *\*ā* : *a* words in Zhivlov’s system seem to correlate with some old *\*a* words. At least the cognates of Hungarian *vág* (reconstructed as *\*waŋi* in the UEW) and *\*kajV-* ‘touch, hit’ (UEW No. 226) belong to the same type.

Helimski (1997b: 121) notes that this belongs to the loanwords that can be either Iranian or Indo-Aryan (> Andronovo Aryan), and that this loan cannot be considered specifically Alanic. The latter argument is not very strong, as the Iranian word is attested in Ossetic, and an Alanic origin could easily yield Ob-Ugric *\*wāta*.

An Indo-Iranian etymology

**Fi *vene*, dial. *venhe*, Ka *veneh* ‘boat’ (< PFi *\*veneh*, cognates in all Finnic languages); SaN *vanas*, *fanas* id. (< PSa *\*vəneš*); Mo *ventš*, *venəš***

< **West Uralic/PU *\*weneš* ‘boat’**

← Pre-II *\*wen-*, *\*wen-e/o-* > OI *ván-* ‘wood’ (root noun), thematized *vána-* (neutr) ‘wood, tree, timber; wooden vessel’; Av *vanā-* (feminine) ‘Baum’ (EWAia II: 500, s.v. *ván-*; Schindler 1972: 43)

(Katz 1985: 318–319; Koivulehto 2001: 250–1)

Koivulehto’s etymology was ridiculed by Helimski (2001), who heavily criticized its semantic development and the need to borrow words for such concepts as ‘boat’. Koivulehto (2001: 250–1) answered Helimski’s criticism and pointed out parallels for borrowed ‘boat’ words, such as Finnic *laiva* (← PG *\*flauja-*) and *karvas* ‘a little boat’ (← PSc *\*karbā*).

Although Koivulehto is right in principle, the etymology cannot be considered absolutely certain in the light of the semantic difficulties. It is possible to derive a word for boat from ‘wood’, but it is difficult to securely prove it. Koivulehto correctly states that the thematic PII noun probably reflects an earlier form with *\*e*, and the Pre-Indo-Iranian etymology is phonologically plausible, but this is again something that cannot be unambiguously proven. Koivulehto does not note that the most archaic form of the word is the OI root noun, and the thematic forms in Indic and Avestan are probably later. It is problematic that the Indo-Iranian word does not have a clear Indo-European background: possible cognates from Germanic (Gothic *winja* ‘Futter, Weide’) have been presented, but EWAia does not consider these as certain. Schindler (1972: 43) does not present any cognates for the Indo-Iranian word.

Theoretically, the Indo-Iranian word could be a later borrowing from some unknown source, but the word had to be present in Proto-Indo-Iranian and after that in the Iranian branch, and thus there is no obstacle to considering the Uralic word an Indo-Iranian borrowing. Also the root noun as a category points to a very early origin of the Indo-Iranian word, as root nouns were not productive in later Indo-Iranian. However, we cannot be absolutely sure that we are dealing with a reflex of Pre-II/PIE *\*e* here. Interestingly, this word does not appear in the list of Indo-Iranian words of unknown origin compiled by Lubotsky (2001b).

The Uralic cognates are regular, so if the etymology is correct, it is an early loanword into a common proto-language of Saami, Finnic and Mordvin; the etymology has to be a very early borrowing because of *\*e*, so it is probable to assume that the word was borrowed already into Proto-Uralic (and was then lost in several branches). All the three branches point to a suffix *-eš*, which is probably a simple Uralic suffix and does not reflect any Indo-Iranian ending.

A parallel (and more convincing) borrowing from a later reflex of the same Indo-Iranian source is Ob-Ugric *\*wǎnV* ‘forest’.

? Indo-Iranian etymology

## Md E *vérges*, *vérgis*, *vérgiz*, M *vərgás*, *vərgás* ‘wolf’ < Pre-Md ?

← PI or PII \**wr̥ka-* ‘wolf’ > OI *vřka-*, Av *vahrka-* id.; < PIE \**włkwō-* (EWAia II: 570–571, s.v. *vřka-*)

(Setälä 1928: 306; FUV: 140; Joki 1973: 342–343; Katz 1985: 133; 2003: 120; Rédei 1986c: 62; Helimski 2000: 193)

This Mordvin word is an obvious Indo-Iranian borrowing, but the etymology involves various problems. On the one hand, the borrowing seems rather old, in spite of the lack of cognates in other related languages, as the ending *-as* in loans often reflects PII \**-as*. On the other hand the Mordvin vowels do not clearly point to any Pre-Mordvin/Proto-Uralic vowel combination. The Mordvin vowel and resonant reflect either a substitute of Proto-Indo-Iranian or Indo-Iranian \**r̥* or some later vocalized Iranian form, but it is difficult to determine what kind of sound substitution has played a role here because of the uncertain vowel development in the Mordvin forms. Katz notes that the non-initial-syllable vocalism in the Mordvin forms is analogical, but this does not explain the development of the vocalism in the initial syllables. Joki considers the word a Middle Iranian borrowing but this can hardly be correct: Mordvin *-Vs* probably reflects the Indo-Iranian masculine ending \**-as* (which, as is well known, was not retained in absolute word-final position even by Old Iranian times, let alone Middle Iranian), meaning that the borrowing has to be significantly older. It is difficult to determine the exact time of borrowing, as both a Proto-Indo-Iranian and Proto-Iranian loan, and perhaps an even later “Old Iranian” loan could give a similar result in the Uralic side.

Although Joki discusses only the Mordvin words, Rédei (1986c: 62) attempts to connect Komi *varges* ‘disobedient, lazy; shrewd, treacherous, shy’ and Old Permic *varkeš* ‘devil’ to the Mordvin word and derives both words from Proto-Finno-Permic. This is a very unfortunate choice, as the Mordvin and Permic words obviously cannot be cognate. Rédei’s ambiguous Proto-Finno-Permic reconstruction \**warkas(e)* ~ \**werkas(e)* also clearly points to the conclusion that the Mordvin and Permic words are not regular cognates. Permic *a* cannot reflect either Uralic \**a* or \**e*, and as the Mordvin word can hardly be reconstructed for Proto-Mordvin, it is obvious that it is not a cognate of the Permic words.

The Permic words might be parallel loans from Iranian, but Saarikivi (2018: 316–317) has considered them loans from Finnic \**varkas* ‘thief’ (> Fi *varas*, Veps *vargas*)<sup>51</sup>. Saarikivi also mentions Udmurt *vargas-* attested in some compounds, a fact which is not mentioned by Rédei. If the Permic words were borrowed from Indo-Iranian, they would have to be very late loans because of the \**a* vocalism and the retained second syllable. Saarikivi’s idea that the Proto-Permic word is borrowed from Finnic is thus more plausible.

Returning to the more precise origin of the Mordvin words, it is possible that the word is a late borrowing, and this explains the ambiguous vowel correspondences within Mordvin. However, in this case Alanic would be the most obvious source language, but the ‘wolf’ word that is attested in Ossetic *biræǰ*, *beræǰ* does not fit the Mordvin word very well. It is uncertain whether the Ossetic word reflects the Proto-Indo-Iranian word, as

<sup>51</sup> Finnish \**varkas* is a well-known borrowing from Germanic \**wargaz*, which is reflected by, among others, OS *warg* m. ‘Verbrecher’, and OHG *warg* m. ‘Feind, Teufel’, as well as ON *vargr* ‘wolf’ (LÁGLOS III s.v.). Saarikivi (2018: 317) further argues that the Germanic word is borrowed from Indo-Iranian \**wr̥ka-*.

according to Cheung (2002: 173), the origin of the word is uncertain. Cheung refers to Abaev's suggestion that Ossetic word is borrowed from Khotanese Saka *birgga*, which is a regular reflect of PII *\*wṛka-*,<sup>52</sup> and he also mentions the resemblance of the Turkic forms like Chagatay *bōri* and the Ossetic forms. Cheung remains unsure of the relationship of these words to Ossetic *biræǰ*, *beræǰ*: the suffix *-æǰ* is difficult to explain if the word is borrowed from Turkic, but the Khotanese geminate *-gg-* would probably not have produced Ossetic *-ǰ*. On the other hand, Cheung assumes that also the Turkic words could be borrowed from Khotanese. Cheung also mentions the Mordvin words, as well as Mari *pire* 'wolf' as possible loans from Alanic (Pre-Ossetic).

The Mari form *pirägy* that is cited by Cheung is probably a compound or a derivation which includes Mari *pire* 'wolf', which according to TschWB is a loan from Tatar. The Mordvin word could rather be borrowed from some Iranian language that retained *\*w-*, akin to Proto-Iranian *\*wṛka-*. The word cannot be very late because of *-s*, as mentioned. It is also possible that the Mordvin word is originally an older loan, and the ambiguous correspondence of the Erzya and Moksha vocalism is due to later interdialectal borrowing. Joki (1973: 342) mentions Ossetic *wærx-æǰ* (*Uærxæg*) 'name of a Nart-hero', which according to Abaev is a regular reflex of the Proto-Iranian wolf word (this is considered as possible also by Thordarson 2009: 12). This word is not mentioned by Cheung (2002). If this word indeed is a relict of an Alanic 'wolf' word, borrowing from this form could explain Mordvin *v-* but hardly the ending.

Despite these difficulties, it is obvious that the Mordvin word originates in Indo-Iranian, as the resemblance of the forms could hardly be coincidental, even though the precise time and place of borrowing cannot be determined. Because of the ambiguous vocalism, the Mordvin word offers no valuable evidence for the substitution of *\*r* in loanwords.

The Samoyedic word *\*wârkhâ* 'bear' is a parallel borrowing from the same Indo-Iranian word (see the entry for more information on the Indo-European background of the Indo-Iranian word). The Mordvin and Samoyed words cannot be cognates.

An Indo-Iranian etymology

**Fi *vermen* 'Oberhaut, Epidermis' (only in Lönnrot's dictionary), *vermes* '(dünnere) Stoff; Halstuch', *verme* 'dünnere, schlechter Stoff, Kleider, Rüstung, Ausrüstung; dünne Eisschicht'; SaN *fierbmi*, SaL *vier'mē*, SaSk *vierem*, SaT *vīrme* 'Fischnetz', < PSa *\*vierme* (Álgu s.v. *fierbmi*)**

< Pre-Fi ? *\*wermV*

← Pre-II *\*wermŋ-* > OI *várman-* 'Panzer, Schutzzung' (EWAia II: 512–513,, s.v. VAR<sup>2</sup>)

<sup>52</sup> Cheung (2002: 173) notes that Ossetic *w-* is the regular reflex of PII *\*w-*. There are some forms that show *b-*, such as *bar* 'will, right' (cf. Av *vāra-*), which is a possible Khotanese borrowing like *biræǰ*, *beræǰ*, and *byjyn* 'to wind', which is possibly borrowed from Slavic (see Cheung 2002: 170, 174). See also the discussion of *b-* forms by Thordarson (2009: 11–12).

(Setälä 1928: 298–300; Mayrhofer 1952: 13–14; Joki 1973: 343; Katz 1985: 345; Rédei 1986c; Lushnikova 1990: 288; Hahmo 1994: 68–70; Helimski 2000: 193; Holopainen 2018b: 150–151)

The Finnic word has been discussed in detail by Hahmo (1994: 68–69), who is critical of the Indo-European/Indo-Iranian origin of the Finnic word because of its scant attestation only in the poetic language. Hahmo sees it possible that *vermen* is simply derived from *verme*, *vermeet* ‘clothes’, and its use in the meaning of ‘shield’ is due to alliteration in the poetic language. According to Hahmo, it is possible that the “existence” of the word has been upheld in later research because it fits the Pre-Indo-Iranian donor form so well. Hahmo further remarks that even if the Finnic word is non-existent/constructed and thus not an early Indo-Iranian loan, the Saami word may well be such a loan.

The Saami and Finnic words are not regular cognates, and according to Lehtiranta the Saami word is borrowed from Finnic. The Saami word is semantically not very close to the Indo-Iranian word, and it is highly doubtful that these two words have anything to do with one another.

The Iranian word is borrowed into Tocharian A as *warm* ‘shuilplaats, hol’ (Isebaert 1980: 89–90, 119).

Not an Indo-Iranian etymology

**Mansi South *wērəw*, West *wēra* (< PMs \**wīrəγ* (-ə)); Khanty East *wär*, South *wár*, North *war* (< PKh \**wār*) ‘Fischzaun’**

< POUG \**wērä* (Zhivlov 2006: 155)

← ? Iranian \**vary-*, NP *vary* ‘Wasserwehr’; OI *vārga-* ‘one who excludes or removes or averts a separate division, class, set, multitude of similar things (...), group’ (EWAia II: 517–518, s.v. *VARJ*)

(Munkácsi 1901: 500; Korenchy 1972: 81–82 No. 41; Lushnikova 1990: 275–276; Helimski 1997b: 121; Sipos 2002: 56; Katz 2003: 243)

The etymology is in principle possible, but the uncertain background and scarce attestation of the Indo-Iranian original makes the connection uncertain. The assumed vowel substitution is shared by some other loans in Ob-Ugric (see above). Semantically the etymology would be possible.

POUG \**wār(γ)V* is Korenchy’s reconstruction, Zhivlov reconstructs long \**ē* for Proto-Ob-Ugric. The same vowel substitution is manifested by some other loans. As this vowel can reflect earlier \**ā*, it is probable that these loans also show \**ā* as the substitution of Iranian *a*. Korenchy reconstructs the Middle Iranian source as \**va/āry-*.

The Mansi word suits the original better. If the word had been borrowed into Proto-Ob-Ugric, it would be difficult to explain why *γ* would have been lost in Khanty, as this would be irregular.

-*γ* in Mansi is a Proto-Ob-Ugric derivational suffix according to Riese (2001: 69). It is possible that the suffix in the loanword has been reanalyzed as an Ob-Ugric (Mansi) suffix, but this does not explain the absence of -*γ* in Khanty.

Katz considers it as possible that the Khanty word could be borrowed from a predecessor of Avestan *vāra-* ‘Deckung, Wehr’ (AiWb: 1411, s.v. *vāra-*).

? Indo-Iranian etymology

**Fi *verso* ‘sprout’, *versoa* ‘to sprout’<sup>53</sup>, Est *võrse* (*\*wersek*) ‘sprout’, Est *võrsu* ‘to sprout’**

< PFI *\*wersV*? ~ *\*wersV* or *\*weréV* ~ *\*weréV*

← Pre-II *\*verčšo-* > PII *\*varčša-* > OI *válšá*-<sup>54</sup> ‘Schössling, Trieb, Zweig’, Av *varəsa-* ‘Haar, Kopfhaar’; or

← PII *\*wrčša-* > OI *vrkšá-* ‘tree’, *varəsa-* ‘a plant’ (EWAia II: 526–527, s.v. *vrkšá-*; Lubotsky 2001b: 313; Beekes 2010: 839–840; Lipp 2009b: 232, footnote 53)

(Parpola 1999: 201; SSA s.v. *verso*; Koivulehto 2001a: 362, footnote 3 = 2016: 260, footnote 3; Aikio 2013b: 167; Holopainen, Junntila & Kuokkala 2017: 119)

This relatively new etymology has not received much critical commentary in the literature of the last couple of decades. The Indo-Iranian etymology is mentioned as a possibility by both SSA and EES. The semantic connection of the Indo-Iranian and Finnic words can hardly be a coincidence, and on the first sight the etymology looks convincing. However, some phonological problems are involved. The Finnic vowel poses a problem to the early Indo-Iranian origin of the word: the correspondence Fi *e* ~ Est *õ* is usually not attested in inherited, Pre-Finnic words, so this word can hardly be a borrowing from an era older than the Proto-Finnic period. But it is fair to say that at the time when Parpola suggested this etymology, the origin of Estonian *õ* was more poorly known than now, and it was not generally accepted that in words like this, Proto-Finnic had *\*õ* and not *\*e*. According to a recent view of Kallio (2014: 160–161), the Proto-Finnic vowel has to be reconstructed as *\*e* (= *õ*) here.

The stem-vowel *o* in the Finnic word is probably secondary. The new Finnic vowel *\*e* first appeared in *\*a*-stems, the Proto-Finnic word was probably *\*vers-a*. Here the *-o* cannot be derived from an earlier combination of *a* and a derivational suffix *j* or *w*, although this is difficult to say as the word is not attested in the relevant languages/dialects, which would show clear reflexes of these earlier glides. The stem-final vowel in Estonian *võrse* points to a different derivative with Proto-Finnic *\*-ek* or *\*-eš*. In any case it looks plausible to assume that this is not a Pre-Finnic *o*-stem – it is usually assumed that the labial vowels become possible in Proto-Finnic or early Proto-Finnic, (Sammallahti 1999: 72–73; Salminen 2002), but Aikio (2015b: 37–39) has argued that *\*o*-stems could have been possible in Proto-Uralic already.

If the word was a real Pre-II borrowing, the non-initial syllable *\*o* ← *\*o* would be a logical substitution, but it is highly disputed whether *-o* existed before the Proto-Finnic

<sup>53</sup> The Karelian verb *versoa* is possibly borrowed from Finnish (SSA s.v.).

<sup>54</sup> For Indic *l*, see the Introduction.

period. It is also dubious to assume that the word was borrowed from a very early form of Indo-Iranian, both because of its limited distribution and its vocalism which rather points to an origin that cannot be earlier than Proto-Finnic. Furthermore, the Estonian word *võrse* shows no trace of the *-o*. It is probable that the *-o* here is secondary, even if it cannot be convincingly explained.

Regarding the two possible Indo-Iranian donor forms mentioned by Parpola, either is possible phonologically. If we assume that the Finnic vowel here goes back to earlier *e*, this vowel can be the substitute for either the *\*e* in *\*werća-* or the syllabic *\*r* in *\*wrčša*. Both are mentioned by SSA, whereas Koivulehto and Aikio, as well as EES, mention only *\*werća-*.

It seems most probable that the Finnic word was borrowed from Indo-Iranian *\*wrčša* or its later Iranian reflex, with Finnic *\*er* or *\*er* substituting for the syllabic resonant *\*r̥* (or a later Iranian reflex with an epenthetic *\*ə*). There are no convincing examples of a Pre-Indo-Iranian loanword that has reflexes only in Finnic, but this is of course still possible, as one can assume that the cognates in other branches have been lost. If the word is borrowed from *\*wrčša-*, we do not have to assume that this is an extremely old borrowing. But in this case one problem appears: the absence of the RUKI rule. In all early loans, Indo-Iranian *\*š* is substituted by Uralic *\*š*, and in apparent exceptions, such as West Uralic *\*warsa* ‘foal’, it can be assumed that the word was borrowed from a later Alanic form where the sibilant *\*š* had merged with *s* in certain positions. There is no reflex of *\*wrčša-* in Ossetic, making it very uncertain to assume that this word could be borrowed from an Alanic variety. Here one could perhaps assume that the cluster *\*-čš-* was simply substituted by Pre-Finnic *\*-ć-* which then yielded regularly *-ś-*.

Indo-Iranian *\*wrčša* is a possible substrate word according to Lubotsky (2001b: 313). However, this has little significance from the point of view of the Uralic etymology, as many other Indo-Iranian loans such as *\*worači* and *\*warkä* are probably also substrate words in Indo-Iranian. The presence of these words in Uralic simply shows that the words were borrowed into Indo-Iranian rather early, before Indo-Iranian speakers came into contact with Uralic speakers.

Due to the various problems mentioned above, the Indo-Iranian origin of the Finnic word remains uncertain but promising.

? Indo-Iranian etymology

**Komi *veres* ‘husband’ (*ver* ‘man’) (? Old Permic *ver* ‘slave, servant’), Udmurt *vor-goron* ‘(old) man’ (< PP *\*ver*; Zhivlov 2010: 171)**

< Pre-Permic *\*wira* or *\*wirä*

← PII/PI *\*wiHrá-* > OI *vīrá-* ‘man’, Av *vira-* id.; from PIE *\*wiHr-o-* (EWAia II: 569–570, s.v. *vīrá-*)

(Munkácsi 1901; Katz 1983b: 175; 2003: 174; Milanova, Holopainen & Bradley [forthcoming])

The etymological suggestion is early but it is missing from many major works because in many etymological sources (such as UEW s.v. *\*orja*; SSA s.v. *orja*; Csúcs 2005: 112) the Komi word is connected with Finnish *orja* ‘slave’, Udmurt *var* id. and their cognates in Uralic (UEW s.v. *\*orja*; another Indo-Iranian borrowing, see the entry). This cannot be correct because the Komi word cannot reflect an Uralic *\*o–a*-stem. Even the Komi *veres* and Udmurt *var* cannot be regular cognates. The new idea of Zhivlov (2010) that Udmurt *vor-* in the obscure compound *vorgoron* is the cognate of the Komi word is, on the contrary, totally regular and plausible.

Although this word is present only in Permic, the borrowing has to be relatively early because it predates the Permic vowel shifts (if the word was borrowed from some later Iranian language to Proto-Permic, the Komi and Udmurt word could not show the regular vowel correspondences typical of inherited Pre-Permic words).

It remains unclear whether Old Permic *ver* ‘slave, servant’ is etymologically connected with the modern Komi word. Semantically the connection is troubling, and it is possible that the modern Komi word *veres* is simply unrelated to the Old Permic word, or that a semantic change from ‘husband’ to ‘servant’ happened in the Komi dialect which was the basis for the Old Permic literary language. But neither can this Old Permic word be derived regularly from PU *\*orja*.

An interesting idea is that Katz (1983b: 175; 1985: 214; 2003: 174) connects North Saami *varris* (*vâres* in the Nielsen orthography) to the Komi word. Theoretically this North Saami variant could reflect Proto-Uralic *\*wirä* or *\*wira*, but *varris* is usually connected to PSa *\*orēs* ‘male’ (Álgu s.v. *varris*; Lehtiranta 2001: 90–91, No. 811), even though the North Saami initial-syllable vowel is irregular. The other Saami languages, such as Inari Saami *ores* and Pite Saami *âries* point regularly to PSa *\*orēs* and could not be cognate with the Komi and Udmurt words discussed here (for the details of the etymology of this Proto-Saami word, see the entry *úr* and Holopainen 2016). It seems best to consider *varris* an irregular variant of the Proto-Saami word and not a reflex of earlier *\*wirä/a*.

An Indo-Iranian etymology

**Fi *viha* ‘hate; venom’; Md E *ožo* ‘green’; Ma (der.) *užar* ‘green’ (< PMa *\*užar*); Ko *vež* ‘envy’, Ud *vož* ‘hate, anger’ < PP *\*vež***

< PU *\*wiša* ‘venom’ (UEW s.v. *\*wiša*; Sammallahti 1988: 554)

← PII/PI *\*wiš-* > OI (Vedic) *viśá-* ‘venom’ (later thematicization), YAv *vīša-*, *viš-* MP *wiš* id. < PIE *\*wis-* (EWAia II: 563–564. s.v. *viśá*)

(Munkácsi 1896: 6; Setälä 1908–1910: 79; Jacobsohn 1922: 27–29; FUV: 141; Katz 1985: 101; 2003: 89; UEW s.v. *\*wiša*; Koivulehto 1999d: 334; 2001e: 369, footnote 9 = 372; 2016: 267, footnote 9, 270; SSA s.v. *viha*; Bereczki 2013 s.v. *užar*)

(see also *\*wiša* ‘hate’ below)

This is an old and widely accepted etymology. The Indo-Iranian etymology for the word ‘venom’ is convincing, but the matter is complicated by the idea, proposed by Parpola

(1999), of separating the words for ‘venom’ and ‘hate’ and deriving these from two different Indo-Iranian originals (see below).

In any case, Fi *viha*, Mo *ožo* and Permic *vež*, *vož* can be regularly derived from PU *\*wiša*, and the old idea to derive the Uralic word from PII *\*wiša-* is totally plausible. The Indo-Iranian word has a plausible Indo-European etymology, it reflects PIE *\*wis-o-s* (from which also Latin *virus* ‘venom’).

In Hungarian, the words for ‘poison’ and ‘hate’ (*méreg* and *mérges*) are clearly etymologically connected (Bárczi 1941: s.v. *méreg*; MSzFE: s.v. *méreg*). Therefore, it seems plausible to suppose that the two Finnic words reflect the same PU stem *\*wiša*.

From Uralic *\*wiša*, words denoting ‘green’ have been derived in several branches: Finnic *viheriä*, *vihreä*, Mari *užar* and Permic *vež*, *vož*. The Permic and Mordvin words are synonymous with the word denoting ‘hate’, whereas both the Finnic and the Mari words contain a suffix. It remains unclear whether Mari *-r* has something to do with Finnic *-rV*. In Mordvin and Mari, the word is attested only in the meaning of ‘green’, and no trace of ‘hate’ can be found. Although the precise development of the words for ‘green’ remains unclear, it does not make the Indo-Iranian etymology less convincing.

An Indo-Iranian etymology

### Fi *viha* ‘hate’, der. *vihata* ‘to hate’ (cognates in all Finnic languages)

#### < Pre-Fi *\*wiša*

← PII *\*dwaiš-* ‘to hate’ > OI *dviṣ-* ‘hassen, anfeinden, verabscheuen’, root-noun *dviṣ-* ‘Hass, Feindschaft’, konkret ‘Feind’; Av *duiṣ-/tbiš-*, root-noun *\*dwiš-* in OAv *daibišuat*, *daibiš-*, YAv *tbiš-*, *duiṣ-* ‘anfeinden, kränken’; < PIE *\*dweis-* (EWAia: 770–71, s.v. *DVEŠ*; Werba 1997: 199 No. 102; Schindler 1972: 26–27)

(Katz 1985: 181–182; 2003: 148–149; Parpola 1999: 201–202; SSA s.v. *viha*)

Both Katz and Parpola have independently suggested that in Finnic there could be two homonymous *viha* words: the other is discussed above. Fi *viha* in the meaning of ‘hate’ would likewise go back to Pre-Finnic *\*wiša*, which in turn would be borrowed from a representative of the Indo-Iranian root *\*dwiš-* ‘to hate’.

It remains very difficult to determine whether the two *\*wiša* words actually represent parallel borrowings or the same word. Uralic *\*w-* can equally well reflect PII *\*w-* or *\*dw-*. It is well known that word-initial consonant-clusters were not permitted in Uralic, and *\*w-* is a plausible substitution for the cluster *\*dw-*. However, there are no parallels for this substitution, and one might suppose that theoretically the word-initial cluster *\*dw-* could also have been accepted into Uralic by substituting *\*t-* or *\*d-* for it instead.

Semantically the etymology is convincing. For the possible connections of the meanings ‘venom’ and ‘hate’, see the entry *\*wiša*<sup>1</sup> above.

? Indo-Iranian etymology (see *\*wiša* ‘venom’)

**Mansi North *wāsiy* , East *wēsi*, West *wēsiy* ‘Elchkalb’**

< PMs *\*wīsəy*

← Iranian *\*wasá-* ‘calf’, compare Ossetic *wæss* ‘calf’, Khotanese *basaka* (< *\*wasa-ka*); < PII *\*watsá-*, derived from IE *\*wet-es-* ‘year’

(Korenchy 1972: 83–84, No. 43.; Lushnikova 1990: 282; Katz 2003: 197)

See also *\*wasa*

The etymology is transparent. The Mansi sibilant *s* reflecting Iranian *s* points to a relatively late loan (an earlier loan from Indo-Iranian form with retained *\*ts* could hardly produce Mansi *s*). Here the Iranian *s* does not reflect earlier *\*ś*, so it is certain that the word was borrowed after the shift *\*ś > s* had occurred in Mansi. Korenchy reconstructs the “Middle Iranian” donor form as *\*vatsa-ka-/vasa-ka*, as she was unable to decide whether the change *\*ts > \*s* had already happened. However, this change happened already in Proto-Iranian, and as the loan has to be relatively late because of the Mansi sibilant, the Iranian donor surely had *\*s* at this point.

Both Korenchy and Katz note the Mansi *-y*, which here probably reflects the Iranian diminutive suffix *-ka* that is attested in the Khotanese form. This is probably correct, although there is also a Mansi derivational suffix *-y*, that is found in some animal names (see *\*päčäy* for discussion). The Mansi word is probably not an Alanic loan, as no form with the suffix *-ka* is found in Ossetic.

The vowel substitution is interestingly Proto-Mansi *\*ĩ*. The West Uralic *\*wasa* (see above) is a separate borrowing from the same source.

An Indo-Iranian etymology

**Ma *urðe-* ‘halten (grosstiere), auf-, grossziehen (Kinder)’ (< PMA *\*wūrðe-*); Ko *verd-*, PO *ve·rd-* ‘ernähren, speisen, füttern (Leute. Tiere)’, Ud *vord-* ‘aufziehen, erziehen, ernähren; gebären’ (< ? PP *\*vordĩ-*; Csúcs 2005: 397)**

?< PU *\*wirta-* (Sammallahti 1988: 554; UEW s.v. *\*wort3-*)

← PII *\*Hwardh-* or PI *\*Hward-*, *várdhati*, < PIE root *\*Hwerd<sup>h</sup>-* ‘to become big, strong’ (EWAia II: 520–521, s.v. *VARDH*; LIV<sup>2</sup>: 228, s.v. *H<sub>2</sub>eRd<sup>h</sup>*)

(UEW s.v. *\*wort3*; Rédei 1986c: 80; Lushnikova 1990: 286–287; Csúcs 2005: 397; Bereczki 2013 s.v. *urðem*)

UEW accepts this old Indo-Iranian etymology, but it is enigmatically noted that if the word is not a loan, Permian *-d* and Mari *-ð* could reflect an old causative suffix. Bereczki (2013) criticizes this idea, as the Uralic causative *\*-tt-* is usually reflected by *-t-* in Mari. As the Indo-Iranian etymology is unproblematic, the idea of a suffix can be rejected. Also Bereczki (2013) accepts the Indo-Iranian etymology.

Sammallahti (1988: 554) reconstructs the word as *\*wirtV-*. In the UEW the reconstruction is *\*wortV-*, which is an impossible preform for both the Mari and Permic vowels. The UEW argues that the Permic vocalism is due to delabialization caused by *\*w-*. Rédei (1986c), contrary to the UEW, assumes that the word was borrowed from a Middle Iranian language into Permic and from there to Mari. This is certainly a more convincing possibility than the assumption of ad hoc vowel developments. It is also possible that the words are parallel borrowings from Indo-Iranian. Csúcs (2005: 397) only mentions the Permic loan as an Iranian loan and he does not comment on the Mari word at all.

If we derive both Komi and Mari forms from *\*wirtV-*, we have to assume that the Uralic vowel *i* is due to the substitution of Indo-Iranian *\*r̥*, as from *\*ar* we could hardly expect *\*ir*. The zero-grade with *\*r̥* is found, for example, in the participles *\*wr̥d<sup>h</sup>ant-* and *\*wr̥d<sup>h</sup>āna-*. This would be the only example of the substitution *\*ir* ← *\*r̥*, however.

Semantically the etymology is convincing. The word denoting the raising of domestic animals fits well with the semantic sphere of the many loanwords connected to cattle and horse breeding.

An Indo-Iranian etymology

### Khanty South *wōć-* ‘eye [?]’ in a compound *wōć-jəŋk* ‘Träne’ < PKh ? *\*woć*

← PII *\*(H)ákš-* ‘eye’ > OI (dual) *ákšī* ‘eyes’, Avestan (dual) *aši id.* (EWAia I: 42–43, s.v. *ákšī*)

(Blažek 1990: 43)

The scarce attestation of the Khanty word and unexpected sound substitutions make this etymology unconvincing. PII *\*kš* would have probably been substituted otherwise, as is shown by such parallels as *\*mekšī*, and Khanty *w-* cannot be explained on the basis of the Iranian word. It is also suspicious that the Khanty word is attested only in a compound with an obscure meaning, so it seems unlikely that this is an early borrowing. There are also early borrowings that have reflexes in Khanty only, but they are usually better attested in Khanty dialects, and especially because of the phonological difficulties we are forced to consider this is a false etymology.

Not an Indo-Iranian etymology

### SaN *oidnit* ‘to see’ (found in all Saami languages) < PSa *\*oajnē-*

< Pre-Sa *\*wojna-* or *\*wajni-*

← PII (or PI) *\*waina-* ‘to see’ > Old Indic *vénati* ‘he/she watches, observes’, Avestan *vaēnaiti* ‘he/she sees’, Old Persian *vaināmiy* ‘I see’ (EWAia II: 582–83, s.v. *VEN*)

(Koivulehto 1999a: 229, 230 = 2016: 231, 232; Holopainen 2018b: 152–153)

See also *\*wajna-*

Koivulehto assumes that the Saami word is an *\*o*-loan, a parallel loan from the same Indo-Iranian source as Mordvin *vano-* (< *\*wajna-*). The etymology is semantically and phonologically convincing. The *o*-substitution is expected after a word-initial labial consonant, as similar substitutions are manifested by other loans such as *\*woraći*.

It is troubling that no reason for the different vowel substitutions for the Mordvin and Saami words can be determined. It would be possible to derive the Saami word also from *\*wajni-* according to the new vowel reconstruction by Aikio (2015b), but this does not solve the problem, as the two forms are still not regular cognates. As noted by Holopainen (2018b), Ante Aikio, (personal communication) has noted that the Lule, Ume and South Saami forms go back to the Proto-Saami form *\*vuojnē-*, not *\*oajnē-*, and these forms would be regular cognates to the Mordvin word and can be derived from PWU *\*wajna-*. If one supposes that the forms apparently reflecting *\*wojna-* can be explained as secondary, one could argue that a Proto-West Uralic word *\*wajna-* was borrowed from PII or PI *\*wayna-*, and this is reflected by Mordvin *vano-* and a portion of the Saami words.

An Indo-Iranian etymology

**Fi *oras, oraisa, ora(i)nen, oroinen, orois* ‘castrated boar’ (cognates in Ka and Ve); Mordvin M *urāzi* ‘boar’** (Zhivlov 2014: 139; UEW: 720)

< P(W)U *\*woraći* or *\*woraća* ‘pig, boar’

← PII *\*warāžhā-* > OI *varāhā-* ‘boar, swine’, Av *varāza-* ‘Männchen des Schweins, Eber, in Verbindung mit dem Namen für Schwein’ (AiWb 1366; EWAia II: 514–515, s.v. *varāhā-*)

(Joki 1973: 296; Gamkrelidze & Ivanov 1984: 924; Katz 1985: 256; 2003: 206; SSA s.v. *oraisa*; Parpola 2005: 44–45)

This is a well-known and convincing etymology, accepted almost universally, but marked with a question mark in the SSA where also the possibility is mentioned that *oras, oraisa* is derived from *ora* ‘awl’. Semantically it is more obvious to assume that the word was borrowed from the Indo-Iranian word with very similar meaning. Moreover, the word cannot be explained as a derivation from *\*ora* at the West Uralic level through any known rules of derivational morphology. Despite its misleading appearance in Finnic *oras*, the word does not belong to the group of “*\*as*-loans”, because here Finnic *-as* reflects PII *-žh-*, part of an Iranian “animal suffix”, and not the nominative masculine ending *\*-as*. Joki mentions that the Finnic word may have become secondarily associated with the unrelated word *ora* ‘awl’, which could have influenced the appearance of its variants.

The Uralic vowel *\*o* as a substitute of PII *\*a* appears in numerous etymologies. Here it could be argued that the labializing effect of word-initial *\*w-* causes the vowel *\*a* to be substituted by a labial vowel. *\*w-* is usually reconstructed for the Uralic word also, although no direct traces of it are attested in Finnic or Mordvin. Usually word-initial *\*w-* is substituted by Uralic *\*w-*. Joki reconstructs the word as *\*oras* ~ *\*woras*. On the other hand, *\*o* also appears frequently as a substitute for *\*a* in front of *\*r*, and Gamkrelidze &

Ivanov have noted that this is one of the words that manifest this rule. As both environments usually cause the \*o-substitution, here it is very difficult to say which one is the precise reason for the substitution in this loan.

It is unclear whether the Uralic word has to be reconstructed as trisyllabic, as the third-syllable vowel undergoes apocope in both Finnic and Mordvin. Traditionally it has been assumed that the phonotactics of Proto-Uralic did not allow consonant-final word stems, but this is not necessarily correct (see Helimski 2000: 191–195; also Aikio 2018 reconstructs consonantal stems into Proto-Uralic).

The Indo-Iranian word does not have a good Indo-European etymology, and de Vaan (2000: 290–1) and Lubotsky (2001b: 312) consider this as a (substrate?) borrowing in Indo-Iranian. This is naturally no obstacle to derive the Uralic word from Indo-Iranian. The general direction of borrowings favours this direction, and there is no (paleolinguistic) evidence that speakers of early varieties of Uralic would have practised pig herding, making borrowing of this kind of word from Uralic to Indo-Iranian very unlikely. According to de Vaan, the last part of the Indo-Iranian word might have been realized as a suffix by the Proto-Indo-Iranians, giving birth to the “animal suffix” \*-ā́ća- that is found in many Indo-Iranian animal names.

Lubotsky (2001b: 309) assumes that the Uralic word was borrowed from *Iranian* because it is a Central Asian substrate word in Indo-Iranian, and thus could have entered Uralic only later, but this circular reasoning: the presence of the word in Uralic shows that the word can be early in PII, existing already before the arrival of the Indo-Iranians in Central Asia. The palatalized outcome of the II affricate rather speaks of a PII origin. In any case, Lubotsky states that the Indo-Iranian word cannot be derived from PIE because of its shape. This belongs to the interesting group of etymologically opaque Indo-Iranian words which have been borrowed into Uralic. The word must have been borrowed into Indo-Iranian rather early, especially if the West Uralic word really is a borrowing from Proto-Indo-Iranian. There are other clear loanwords from Indo-Iranian to Uralic with a similarly opaque etymological background.

The Indic word *\*waraǰha-* is mentioned several times in the Rigveda, and it is possible that its cultural and mythological significance in Indo-Iranian was one of the reasons the word was borrowed. Even though Parpola (2005: 44) rightly states that the boar was usually associated with demons in the Rigveda (for example, Indra’s killing of the boar Emuša is mentioned in RV 8.077.10), it is also attested in a more positive sense, at least in RV 1.114.05, Rudra is called *divó varāhám* (acc.) ‘the boar of heaven’.<sup>55</sup> In any case, in the Avestan texts the boar is positively associated with the victory god *Vərəθraγna* and Indra himself, as is argued by Parpola (2005: 44–45), which means that the animal could have had a notable significance among the early Iranians of the steppe too.

Another possible borrowing denoting ‘pig’ from Indo-Iranian is *\*poréas*, which is, however, derived from Balto-Slavic by Kallio (ms.); see the entry.

### A convincing Proto-Indo-Iranian etymology

---

<sup>55</sup> In the translation of Jamison and Brereton (2014): “We call down with reverence the boar of heaven, flame-red, with braided hair, turbulent in form. Bearing in his hand desirable healing remedies, he will extend shelter, covering, and protection to us.”

Ng *ɲarka* (? < PS *ǎrkā*), En *bóggo*, NT *wark*, NF *βar̥k*, Slk *qorki* < PS *\*wǎrkā* ‘bear’ (SW: 170)

< ? Pre-S *\*wurka*

← PI or PII *\*wr̥ka-* ‘wolf’ > OI *vṛka-*, Av *vǎhrka-* *id.* M; < PIE *\*wŕkʷo-*, cf. Lith *vilkas*, OCS *vьlkъ*, etc. (EWAia II: 570–571, s.v. *vṛka-*; for the problems with the irregular Greek and Latin reflexes of the IE word, see KEWA s.v. III *vṛkah*, de Vaan [2009, s.v. *lupus*] & [Beekes 2010, s.v. λύκος])

(Janhunen 1983: 121; Carpelan & Parpola 2001: 127)

Janhunen’s etymology for the Samoyedic word is rather convincing, although the divergent form of Nganasan compared to the rest of the Samoyed languages poses some problems for the Proto-Samoyed reconstruction (the Nganasan word reflects a PS form *\*ǎrkā* without the initial *\*w-*); this could perhaps point to parallel contacts with Iranian, especially if Nganasan truly was the first Samoyed language to branch off from Proto-Samoyed, as Janhunen has argued elsewhere (1998). This would point to a rather late contact scenario, and from the Indo-Iranian point of view, the donor language could be anything between PII and some “Old Iranian” type language. The breakup of Proto-Samoyed probably started sometime in the last millennium BC (Janhunen 1998: 457), and it can be assumed that an Iranian contact language at this time would have reflected the “Old Iranian” type, although it is extremely difficult to say anything in more detail of an unattested Iranian language spoken thousand of kilometers from the areas where the early Iranian languages were attested.

If the loss of *\*w-* is secondary in Nganasan, the development has to be very early, as *b-* is the normal reflex of PS *\*w-* in Nganasan. In the material of SW, there appear a couple of other irregular developments of initial *\*w-* in Nganasan: *\*woaǰ* ‘Insel’ > Ng *ũai*, *\*wotʰ* ‘Beere’ > Ng *ũta*, *ɲota*.

No other convincing loan etymology has been offered for the Samoyed word. Katz (1985: 136, footnote 131) has attempted to derive this word from the Indo-Iranian word for ‘bear’, *\*Hr̥kša-* (or *\*h₂r̥kša-* in the notation of Katz) but this is phonologically impossible. Katz assumed that *\*w-* could be a substitute for *\*h₂*, but this is unconvincing and lacks parallels. Katz argues that Janhunen’s etymology is not convincing because of the atypical semantic change, for which no parallels are presented by Janhunen.

Blažek (2017a: 179) has supported Janhunen’s etymology, mentioning it as a better alternative to the one presented by Katz. Blažek considers Janhunen’s idea of taboo in the semantic development as plausible, and also gives semantic parallels: the well-known Old English personal name *Beowulf* is literally a compound ‘bee-wolf’ and denotes ‘bear’, and in a variety of the Romani language spoken in Russia, the inherited Indo-Aryan word for ‘bear’, *ryč* (< *\*Hr̥kša-* or *\*Hr̥tša-*), is used to denote ‘wolf’.

The semantic match between Indo-Iranian and Samoyed words is not perfect, and this could be considered the weakest point of the etymology; this is criticized especially by Kallio (2002: 37, footnote 12). Janhunen’s original argumentation stresses the fact the bear possessed the same kind of central role in the culture and belief system of the northern pastoral peoples as wolf did in those of the Iranian and later Turkish pastoralists of the steppe. Even assuming a simple metonymic change from a name of one big predator to

another would not be a very big obstacle; for instance, the Finnic *lehmä* ‘cow’ and Mordvin *lišme* ‘horse’ clearly originate from the same (probably substrate) source, and while the semantic difference is great, it is still possible to reconstruct a common meaning ‘large domestic animal’ as Aikio does (2015).<sup>56</sup>

It has been suggested that the Indo-European word for ‘wolf’ derives from an adjective *\*ulk<sup>w</sup>ó-* ‘dangerous, wolf-ish’ (see, for example, Pooth [2015: 66], who also states that the original meaning is still visible in the OI word *a-vr̥ká-* ‘not-harming, undangerous etc.’). If this adjective still existed in the Iranian languages of the steppe (which is possible very difficult to prove), and if it was still transparent to the speakers of this Iranian variety that the words ‘wolf’ and ‘dangerous’ are connected, the meaning of the Samoyed word ‘bear’ could be rather easily explained. Interestingly, it has been suggested that the cognate of this Indo-Iranian word in Lydian *walwe-* meant ‘lion’, not ‘wolf’ (Lehrman 1987; Steer 2015: 49, 78). Kloekhorst (950–951) considers the connection of Lydian *walwe-* with PIE *\*ulk<sup>w</sup>ó-* formally impossible (PIE *k<sup>w</sup>* yields Lydian *k<sup>w</sup>*, not *w*). However, the Anatolian languages might have a cognate of the ‘wolf’ words, namely Hittite *ualkuua-* ‘something negative’, but this remains also unclear because of the uncertain semantics of the Hittite word (Kloekhorst op. cit.).

Phonologically it is interesting that the Proto-Samoyedic sequence *\*ə̂r* could reflect PU *\*ur*. This would be a fitting substitution for PII/PI *\*r̥* in an environment after a labial semivowel. Not many such parallels are known among the Indo-Iranian loans, but the substitutions of the syllabic resonant are divergent anyway, and a similar substitution is found in many of the tentative PIE etymologies of Koivulehto (2001) in more westernly Uralic languages. Here one has to take in mind the possible divergent reflexes of PI *\*r̥*, which could have developed into a sequence *\*ər* already by the time the word was borrowed. The word may have also been borrowed during the Proto-Samoyedic stage, where *\*ə̂r* would simply result from the substitution of PI *\*r̥*/*\*ər* and not from earlier Pre-Samoyedic *\*ur*.

Helmski (2000: 193) has suggested an alternative Indo-European source *\*worgo-* ‘cuning, thief’ for the Samoyed word, but this is semantically not more convincing. The Indo-European word is not attested in Indo-Iranian, and it is difficult to see what would be the source language here. It is difficult to assume that a meaning ‘thief’ would have developed into ‘bear’; although in Mansi a development ‘thief’ > ‘wolverine’ is attested, for the sacred animal, bear’, it would be less credible. Also phonologically the etymology is not an improvement to Janhunen’s etymology (*\*wo-?*).

Mordvin *vérges* ‘wolf’ is probably a parallel borrowing from the same Indo-Iranian source.

## An Indo-Iranian etymology

---

<sup>56</sup> For the sake of clarity, it has to be stressed here that Aikio does not state that the Finnic and Mordvin words are inherited from Proto-Uralic but that they are parallel borrowings from an undefined source.

**Mansi West *wəniy*, North *wāni* ‘forest’, East *woāni* ‘eine von einer Flußbiegung umgebene Landzunge’, South *wānəw* (< PMs *\*wānəy*); Khanty East *wantəŋ*, South *wántəŋ* ‘Landzunge, Landspitze’ (< PKh *\*wāntəŋ*)**

< POug *\*wānV* (Zhivlov 2006: 145–146)

← Iranian *\*van-*, compare YAv *vanā-* ‘tree’,<sup>57</sup> MP *wan*, Sogdian *wn-* ‘tree’, Ossetic *byn*, *bun* ‘forest’; OI (Vedic) *ván-* ‘Baum, Holz’, also thematized *vána-* ‘Baum, Holz, Wald’ < PII *\*van-* (EWAia II: 500, s.v. *ván-*)

(Setälä 1900: 32; Munkácsi 1901: 150; Korenchy 1972: 79 No. 38; Katz 1985: 129–130; 2003: 109; Lushnikova 1990: 272–273; Sípós 2002: 55)

See also *\*weniš*

According to Zhivlov (2006: 145–146), the Ob-Ugric words are regular cognates, which means that this might be a considerably old loan. The substitution (I)I *\*a* > Ob-Ugric *\*ä* seems to occur in many Indo-Iranian loanwords. It is unclear whether they should be labelled “Middle Iranian” loanwords as Korenchy does, but they could reflect a development of PII *\*a* towards a front vowel. On the other hand, because loans of varying age in Ob-Ugric manifest the same substitution, it is probable that the donor language had *\*a* which was simply substituted by *\*ä*. Zhivlov (2013) and Helimski (1997b) have argued that the words that show this substitution are borrowed from the unattested Andronovo Aryan language.

Korenchy notes that the Khanty and Mansi words are derivatives, supposedly going back to an underived *\*wānV*.

Semantically the etymology is convincing. Korenchy convincingly argues that the meanings of the Khanty and Mansi words can be derived from ‘forest’.

An Indo-Iranian etymology

**Mansi South *wəŋgər*, West *wəxəp*, North *wānkrip* ‘Haken’, East *wəxɾ-*, West *wəxrit-* ‘mit einem Haken ziehen’ (? < PMs *\*wāŋkV*); Khanty North *wəŋrep* ‘Haken, Bootshaken, hölzerner Haken’ (? < PKh *\*wāŋkV*)**

< ? POug *\*wāŋkV*, *\*wāŋkV(rV)* (Honti 1982: 195)

← ? Iranian *\*wank-*; OAv *vašīietē* ‘strömt heraus, sprudelt heraus, wogt’, YAv *vaš-* ‘sich spontan äußern, hervorsprudeln’; OI *vakrá-* ‘geboren, krumm’, *vānkri-* ‘Rippe’, *vāncati* ‘geht krumm, wankt’ (EWAia II: 492–493, s.v. *VAÑC*)

(Munkácsi 1901: 118–119; Jacobsohn 1922: 119, footnote 208; Korenchy 1972: 79–80 No. 39; Katz 1985: 111; Sípós 2002: 55–56)

<sup>57</sup> Korenchy also mentions Avestan *van-*, but according to EWAia there is no trace in Avestan of this root noun that would correspond to Vedic *ván-*. Korenchy does not mention the thematic noun *vanā-*.

The etymology is phonologically possible, but semantically not very obvious, and it is possible that the resemblance of the Khanty, Mansi and Indo-Iranian words is simply accidental. Korenchy argues that the original meaning of the Iranian words could have been ‘krumm, gebogen’. The Ob-Ugric words are not mentioned by Zhivlov. Honti (1982: 195) reconstructs Proto-Ob-Ugric *\*wāŋk*. Korenchy notes that words with a derivational suffix in *-r* are attested only in Indo-Aryan, not in Iranian. If the etymology is correct, the Ob-Ugric words should have been borrowed from Proto-Indo-Iranian already. Korenchy mentions Khanty (South) *wan̄kl̄s* as a possible example of a form without *-r*, but it is uncertain whether the word really belongs here.

? Iranian etymology

**Mansi North *wār* ‘machen’, East *woar*, South *wār-* ~ *\*wār-* (< PM *\*wārä*, *\*wār-*); Khanty North *wer-t̄š*, East *wer-tä*, South *wer- id.* (< PKh *\*wār-*)**

< POug *\*wārV-* ‘machen’ (Zhivlov 2006: 132–133)

← ? Iranian *\*varz*, YAv *vərəz* ‘wirken, handeln’, 1sg *vərəziāmi* < PII *\*warǵ* < PIE *\*wergʷ* (Werba 1997: 438 No. 579; Cheung 2007: 425–427, s.v. *\*u<sub>2</sub>arz*)

(Korenchy 1972: 80–81 No. 40; Katz 1985: 228; 2003: 185; Lushnikova 1990: 279–280; Helimski 1997b: 121)

If this etymology is correct, we have to assume a simplification of the *\*rz* (? *\*rdz*) cluster, possibly because of phonotactic reasons. This makes the etymology slightly uncertain. The root *\*wergʷ* ‘to do, make’ is attested widely in Indo-European, but it is most probable that the Ob-Ugric words are borrowed from Iranian and not some other branch. It is difficult to date the borrowing precisely, but *\*ä* (/ *\*ā*) as a substitution of earlier PII *\*a* seems to occur in a stratum that is shared only by Khanty and Mansi (same substitution occurs also in Permic, but these are probably independent loans). The Ob-Ugric cognates are regular, at least according to Zhivlov, meaning that the word could be an early borrowing to a common proto-language of Khanty and Mansi. Helimski (1997b) considers the etymology as problematic.

Katz also mentions Permic forms: Ko *veřni*, Ud *vörni* ‘sich bewegen etc.’ (< PP *\*vöri-* Csúcs 2005: 396) which he considers parallel borrowings or reflections of a different Uralic “ablaut grade”. These words are certainly not cognates of the Ob-Ugric words. The possible Iranian etymology of the Permic words suffers from the same problem with the *\*rz* cluster.

Honti (2017) argues that Hungarian *ver* ‘to plait, to lay the rope; to hit’ could be a cognate of the Ob-Ugric words, and thus the word could be reconstructed for Proto-Ugric. A Turkic etymology for *ver* has been suggested: Turkic *ör-* ‘to plait’ < Proto-Turkic *\*hōr-* (WOT s.v. *ver*), but Honti refutes this. I see no reason to support the connection of the Hungarian word with the Ob-Ugric words because of the semantic differences.

? Indo-Iranian etymology

**Khanty North *wārəs* ‘Roßhaar’ East *wārəs* ‘Schwanzhaar des Pferdes’; North *uwīrəs* ‘Birkenwald’, East *wīrəs* ‘Dickicht, Gestrüpp von Birken und Espen’**

< PKh *\*wār(ə)s*

← Iranian *\*wārša-*, compare Avestan *varāsa-* ‘hair of human and animal, mostly hair of the head’, OI *válša-* < PII *\*wárca-* (EWAia II: 545, s.v. *válša-*)

(Munkácsi 1901: 499; Korenchy 1972: 82–83 No. 42; Joki 1973: 336; Katz 1985: 309; 2003: 247; Lushnikova 1990: 278–279; Zhivlov 2013: 11; Aikio 2015a: 167)

See also *\*wersV* (> Fi *verso*)

This old explanation for the Khanty word is a transparent and clearly good Indo-Iranian etymology. Although the word has been considered a late borrowing in Khanty by Joki and recently Zhivlov (2013), the situation might not be this simple. It has been thought that the Khanty word has been borrowed after the change *\*ś > s* from an Iranian form with *s* (from earlier *\*é-*) but theoretically, the Khanty word could have been borrowed from a Proto-Indo-Iranian form *\*wárca-* before the loss of the palatality of *\*ś* in Khanty. This was noted already by Korenchy (1972). On the other hand, Khanty *\*ā* has been considered to be peculiar to later borrowings, but as the dating of the Indo-Iranian loans of the Ob-Ugric languages is not very solid in general, this is not a decisive argument. Also the loans with *\*ä*-vocalism can probably be older than has previously been thought (see *\*šāṅkV*, *\*säptä*). There are not many other borrowings from Iranian to Khanty after the sound change *\*ś > s*, and also in Mansi there are also only a couple of examples. Lushnikova (1990) considers the word as a late loan. For the time being, the dating of the etymology remains open, but the Indo-Iranian source itself is clear and obvious.

Mayrhofer (EWAIA II: 545) states that the Indo-Aryan *l* in the cognates of this word is secondary, which is probably correct and corresponds to the mainstream view that the change *\*l > \*r* was complete in Proto-Indo-Iranian (Mayrhofer 2004; Kümmel 2018a).

An Indo-Iranian etymology

**Ma *wäryə*, *weryə* ‘kidney’ (< PMA *\*weryə*); Ko *ver̥k* id. (< PP *\*vörk* Csúcs 2005: 394)**

< ? PU *\*wärkä* ‘kidney’ (UEW s.v. *wärkä* (*wärrkä*))

← PII *\*wrtka-*,<sup>58</sup> > OI *vr̥kkáu* (du.) ‘Nieren’, Oss *wyrg*, *urg* id., Zoroastrian Middle Persian *gurdad* < *\*wrtka-ka-*, Sogdian *γwrdtq-* < *\*wrtká-* ~ *wrt(k)kǎ-* (Cheung 2002: 30; EWAia II: 571, s.v. *vr̥kká-*; Korn 2005: 183, + footnote 36 with references; Sims-Williams 1983: 49)

<sup>58</sup> Koivulehto (1999b: 159) mentions a PIE pre-form *\*wrtko-* but this has to be an error, as the Indo-Iranian word had to have a cluster of *\*tk*.

(FUV: 140; Joki 1973: 341; Katz 1985: 156; Rédei 1986c: 79; Koivulehto 1999d: 159; Berezcki 2013 s.v. *werye*)

This is a well-known and convincing etymology. However, the etymological background of the Indo-Iranian word is uncertain. Lubotsky (1988: 81; 2001: 313) considers the word a borrowing in Indo-Iranian, and sees the traditionally argued derivation from the PII root *\*wart-* ‘to turn’ (< PIE *\*wert-*, see Werba 237 No. 182) as impossible, because *-ka-* is a denominal suffix in Indo-Iranian. Nevertheless, the Uralic form should still be considered a borrowing from Indo-Iranian, as the simple Uralic cluster *\*-rk-* can easily be explained from PII *\*-r<sub>ṣ</sub>tk-*, whereas the opposite would be very unlikely, so the Indo-Iranian word is certainly not an Uralic borrowing. The Indo-Iranians must have borrowed the word before coming into contact with Uralic speakers, and the same is true of several other Indo-Iranian loanwords in Uralic, whose etymological background in the source language is unclear. Of course, it is theoretically possible that both PU and PII borrowed the word from some common source, but this is not probable in the view of the large stock of II borrowings in Uralic but the lack of common substrate words in both families.

Interestingly, the *-rk* cluster of the Komi word cannot regularly reflect earlier *\*-rk-*, as in a simple *\*-rk-* cluster the *\*k* would have disappeared. Because of this, the UEW reconstructs the form as *“\*wärkä (\*wärrkkä)”*. The Mari word points to simple *\*-rk-* (compare Mari *ur<sub>ṣ</sub>e-*, *ḥr<sub>ṣ</sub>e-* ‘to sow’ < PU *\*worka-*), but there are no Uralic words with *\*-rkk-* which could show how this cluster developed regularly in Mari. One could assume that here the Iranian cluster *\*-tk-* has been substituted by a Permic geminate *-kk-*, but because of the Mari evidence, it looks more likely that the *\*-rk-* cluster in this word shows Permic secondary gemination in a similar way as the *\*-rt-* clusters in loanwords might show (see *\*kertä*, *\*mertä*). It seems, however, that among the loanwords there are no exact parallels to the development *\*-rk-* > *\*-rkk-*, but it seems plausible to assume that a similar development affected both *\*-rt-* and *\*-rk-* clusters in Pre-Permic.

Lubotsky (2001b: 309) supposes that the Uralic word was borrowed specifically from Iranian, because it is a Central Asian substrate word in Indo-Iranian, and thus it could have entered Uralic only later; cf. s.v. *oras* in this work for a similar argumentation. This is possible in the case of this word, but it is difficult to prove it for certain, and it is important to note that regarding some other loans, Lubotsky’s arguments of a specifically Iranian origin are untenable. The limited distribution in Uralic could speak of a late borrowing, perhaps separately into Mari and Permic from a later Iranian source. Rédei (1986c: 79) has considered the Mari word a borrowing from Permic, but no such explanation is found in Rédei (1988a: 662). According to Rédei (1986c: 79) the Permic word is borrowed from “← miran. *\*vr<sub>ṣ</sub>ka-*”, but this argumentation is misleading: in no Middle Iranian language could have a form like *\*vr<sub>ṣ</sub>ka-* have existed, as the syllabic resonant certainly was vocalized by then, and the borrowing into Uralic may well be much earlier than the Middle Iranian stage. Interestingly, the UEW states that the word was borrowed from Proto-Indo-Iranian or Proto-Iranian.

In many Middle Iranian languages the *\*-rk-* cluster develops into *-lk-* or *-lg-* (such as Khotanese *bilga*), but not in Ossetic. The Mari and Komi vowels could reflect PII or PI *r<sub>ṣ</sub>*, but there is no obstacle to assuming that the Uralic words could have been borrowed from an Alanic source that later developed into Ossetic *wyrg*, *urg*.

It remains uncertain whether the Mari and Komi words are parallel loans or inherited from an earlier proto-language. Both could reflect PU *\*wārkä* regularly, so there is no reason to consider the words late parallel loans, although this cannot be ruled out for certain. Koivulehto (1999d: 159; 2016: 200) notes that the etymology of Pre-Permic *\*pärtä* manifests a similar substitution (see the entry).

The Khanty forms like (South) *wirək* ‘kidney’ are borrowed from Komi, and so is probably Mansi *vureχ* which is attested only in the expression *vureχ päŋk pūm* ‘Achillea millefolium’ (Toivonen 1956: 20; UEW).

An Indo-Iranian etymology

**Ko *egir* ‘coal’; Ud *egir* id. < PP *egir*** (Csúcs 2005: 331; Zhivlov 2013: *ägir*)

< Pre-P *\*äjkarä*

← I *\*angāra-*, cf. OI *ángāra-* ‘coal’, ? Parachi *yâr*, *yār* id. (EWAia I: 48, s.v. *ángāra-*)

(Joki 1973: 258; Lushnikova 1990; Csúcs 2005: 331; Zhivlov 2013: 219)

Despite the word’s limited distribution – it is found only in Permic – the word cannot belong to the latest layer of loans in that branch because it has gone through the regular denasalization of the *\*-ŋk-* cluster. According to Rédei and Zhivlov, the *a* vocalism points to earlier *\*ä*. This can be either a sign of fronted articulation of *\*a* in the Iranian donor language, or it can simply reflect a sound substitution. It has often been assumed that the *\*ä* shown in loans into Permic and Ob-Ugric reflects a later Iranian development (Korenchy 1972; Rédei 1986c: 35), but the evidence points to early substitution *\*a* ← *\*ä* at least in Ob-Ugric, and it would not be an impossible idea to assume a similar development in Permic.

Zhivlov (2013) lists this among the Andronovo Aryan loanwords, assuming that the vocalism *\*ä* is typical of exactly this hypothetical branch of Indo-Iranian. As the *\*ä* substitution can be found in Ugric loans which in no way fill the criteria that Helimski (1997b) and Zhivlov present for the Andronovo Aryan loans, it is better to reject this idea. This Permic word is probably simply an Iranian borrowing, as has been usually assumed.

Semantically the etymology is convincing. The Indo-Iranian word does not have certain cognates in any other branch of Indo-European, and it is uncertain whether the Parachi form cited by Zhivlov is related to the Indic word, as it is not mentioned by EWAia. The Indic word appears already in the Rigveda, so the word can well-have been present in Proto-Indo-Iranian already, and it is possible that its later Iranian cognate has disappeared.

An Indo-Iranian etymology

**Mansi North *ārās* ‘Feuerstelle’, West *ārās* ‘fire’ (< PMs \**ārV*) ‘fire, fire-place’**

← **Iranian ?**, cf. **PI \**ātar-* ‘fire’ > Av *ātar-*, *āθr-* ‘fire’, MP *ādur*, Scythian personal names *Ἀρθάμων*, *Ψεθδαρτάκη*, Sogdian *”tr*, *’rd*, Oss *art* ‘fire’ (Cheung 2002: 154)**

(Munkácsi 1901: 244; Korenchy 1982: 53–54, No. 4; Lushnikova 1990: 182; Helimski 1991: 221; Helimski 1997b: 121)

The borrowing is considered uncertain even by Korenchy. She reconstructs the donor form as “Middle Iranian \**art/ärt-*”. Although semantically the etymology is plausible, the problem is that the \**θr/rt* cluster would have been simplified in Mansi; it is uncertain whether the donor form had \**θr* (or \**tVr*) like Avestan does or whether it had already gone through the metathesis which the Scythian and Ossetic forms show, and which must have happened early as it is also shown by a part of the Sogdian material.

The etymology of Korenchy has been considered as very doubtful by Helimski (1991: 221; 1997: 121). The etymology is rejected by Katz (2003: 240), but he assumes that the word is borrowed from the same stem as \**etrā* (see the entry), more precisely from the accusative \**ātram* (> Av *ātrəm*). Korenchy assumes that the word could have originally appeared mostly in the compound \**āret-tāwtV* ‘fire-fire’ in Mansi, and the *-t* would have been dropped because of reanalysis. This explanation is ad hoc. It is also difficult to understand why the Mansi speakers would have used the two words for ‘fire’ in a compound. In general, Mansi has preserved consonant clusters quite well.

Riese (2001: 112) refers to Korenchy’s etymology and considers the Iranian etymology as possible, but he does not comment the etymology in more detail. Riese notes that the word might include the Mansi denominal suffix \**-ś*, but this is uncertain.

The isolated Mansi word is probably borrowed from somewhere, as there are other, inherited words denoting ‘fire’ in Mansi, such as the inherited word \**tāwt* (cognate to Hu *tűz* etc.), and this word must have appeared from somewhere. As ‘fire’ was important for the mythology of the Iranian peoples, it would be a natural source for borrowing, but for the time being the possible Iranian origin of the Mansi word remains highly uncertain.

? Indo-Iranian etymology

**Mansi North *ēsərma*, East *ēsərmə*, South *īserəm*, West *ēsürəm* ‘Scham, Schade, Schamgefühl’**

< PMs \**āsVrVmV*

← **Alanic \**æfsærm*, > Oss *æfsærm*, *æfsarm*, cf. YAv *fšarəma-* ‘Scham(gefühl)’ < PI \**fšarma-* (Cheung 2002: 157)**

(Munkácsi 1901: 567–568; Korenchy 1972: 55; Joki 1973: 260–261)

The Mansi word is clearly an Alanic loanword, as is shown by the prothetic vowel. The Ossetic *æ*-prothesis is dated to after the text of Tzetzes in the 12th century (Cheung 2002: 36) which means that the loan can be relatively late, although it is possible that prothesis

took place at different times in the different parts of the Alanic speaking-area. This specific Ossetic word is probably attested in the Alanic of Tzetztes as φάρ, which Bieltmeier (1993: 16–17) amends to \*σφαρ or \*φσαρ. If the word had been borrowed from an earlier, Old Iranian-type form, one would expect the word-initial consonant cluster to have been simplified, as there are no examples of a vowel prothesis in the early Iranian loans into Mansi/Ob-Ugric. Korenchy also remarks that the retained sibilant *s* shows that the word has to have been borrowed after the change \*ś > \*s had taken place in Mansi. Korenchy also mentions the \*ä vocalism here as a sign of an Alanic-type borrowing. Here also the sibilant *s* instead of \*š points to an Alanic donor, as \*š would also have been possible in Proto-Mansi (Honti 1999: 125–129). Proto-Iranian \*š is retained in most languages, but not in this environment in Alanic/Ossetic.

Korenchy explains the different vowels through Proto-Mansi ablaut. It is however possible that the loan is so late that the different anlaut vowels reflect different substitutions.

Hu *szemérem* ‘shame’ has been considered as a loan from the same Iranian source (Korenchy mentions that Munkácsi considered the Hungarian and Mansi words cognates, but this is, of course, impossible), but Joki states that this is unlikely.

An Indo-Iranian etymology

### 2.3. NEW ETYMOLOGIES

**Mordvin** E *moro* ‘song’, *mora-* ‘to sing’; **Ma** *muro*, *mârâ* ‘Lied, Vogellied’ < PMA \**mūrə*, *mura-* ‘to sing’; ? **Khanty (East)** *morəy-* ‘Unwesen treiben lärm (Menschen)’

< PU \**mura*

← PII or PI \**mrawH-*, \**mruH-* ‘to speak’ (< PIE \**mlewh<sub>2</sub>-*), > **Avestan** *mrao-*, *mru-*; *mrawaiti* ‘he/she speaks’, *fra-mrū-* ‘auflegend, rezitierend’, **OI** *brav<sup>i</sup>-* ‘sprechen, sagen’ (AiWb: 1192–1196, s.v. *mraw-*; EWAia II: 235–236, s.v. *BRAV<sup>i</sup>*)

It is unclear whether the Khanty word belongs here, but it seems that the Mordvin and Mari words are regular cognates and could reflect an earlier (Proto-Uralic/Pre-Mari/Pre-Mordvin) \**mura*. There is no etymology for the Mari and Mordvin words beside this, but as these two branches share only very little vocabulary that is not found in other languages or not borrowed from some common source (such as Turkic or Russian), it is natural to search for a foreign, that is, substrate or loan etymology for this word.

Keresztes (2017: 140) considers the Mari and Mordvin cognates part of the “Volgaic vocabulary” (consisting of more than 20 words), the origin of which remains unclear; he states that they can be either retentions whose cognates have disappeared elsewhere, or loans from an unknown source. Also TschWb (396) considers *muro* a “Volgaic” word, and Zaicz (2005: 125) considers this a possible “Volgaic” etymology.

The Khanty words are considered uncertain cognates by the UEW, but as the sound correspondence Mo *o* ~ Ma *u*, *â* is regular, and as according to Sammallahti [1988: 500–510] PU \**u* is reflected as Proto-Khanty \**o* in most environments, then the Khanty word could also be derived from the same Uralic source. The UEW is cautious of the semantic

difference between the Khanty word and the rest, but ‘making noise’ and singing are semantically not that far from each other.

The semantic development from Indo-Iranian ‘speak’ to Uralic ‘sing’ can be fairly well explained, as the meanings of ‘singing’ and ‘speaking’ are not very far from each other, and the Mari word seems to express also other kinds of sounds than purely singing: TschWb mentions ‘krähen (Hahn)’, ‘rufen (Kuckuck)’, ‘krächzen (Krähe)’, ‘schlagen (Nachtigall, Wachtel)’, etc. In Mordvin the meanings of the noun *moro* are more restricted to ‘singing’, but the verb *morams* can express other kinds of sounds as well (MWb: 1284). Although the Avestan root mainly denotes speaking, it can also mean recitation, and it frequently appears in hymns in which gods and divine beings are addressed. However, if the Khanty word is related to Mari and Mordvin words, the original meaning of the Uralic word might have indeed been not ‘singing’ but some kind of ‘noise’ or ‘sound’.

A competing loan etymology has been suggested, but it has not received general acceptance. Katz (2003: 279–80) has derived the Volgaic word(s) from “frühurarisches” *\*g<sup>w</sup>əra-* (< Proto-Indo-European *\*g<sup>w</sup>ṛH-*), a root which has *inter alia* the meaning ‘to sing’. Although semantically this etymology is possible, phonologically it cannot be correct. PIE *\*g<sup>w</sup>* would be reflected in Uralic as a stop, as there are parallel examples to this, but no credible parallels to a substitution by *\*m* (Katz has some such examples, but none of them have been generally accepted by other researchers). And what is more important from the point of view of this study, the labiovelars were not retained in Indo-Iranian, and if the etymology of Katz was correct, it could only be a Proto-Indo-European loanword, not an Indo-Iranian one.

**Mansi South *ńāt-*, West *ńōt-*, North *ńōt-* ‘to help’ (< PM *\*ńāt-*); Khanty North *ńot-*, South *ńot-* id. (< PKh *\*ńāt-*)**

< POUg *\*ńātV-* ‘to help’ (Zhivlov 2006: 138)

← PII/PI *\*HnātHa-*, OI *nāthá-* m. ‘refuge, help’ < PIE *\*Hnēt-h<sub>2</sub>-o-*; ? root *\*HnaH* < *\*h<sub>2</sub>neh<sub>3</sub>* (Kroonen 2013: 388, s.v. *\*nēpō-*; EWAia II: 33–34, s.v. *NATH*; Werba 1997, s.v. *\*nā*)

See the entry *\*ńāt-*, where this new etymology is suggested instead of the earlier, less likely explanation of Korenchy (1972).

**Md M *śarəndə-*, *śarənd’ə-* ‘to shit’; Ma E *šora-*, W *šara-* id. (< PMa *\*sāra-*); Hu *szar* ‘shit’, *szarik* ‘to shit’**

< PU *\*śara-* (Aikio 2015b: 56; UEW s.v. *śar<sub>3</sub>* (*ćar<sub>3</sub>*))

← PII *\*ćar-* > Av *sairiia-* ‘Mist, Dünger’, < PIE *\*k<sup>er</sup>H-* (LIV<sup>2</sup>: 326, s.v. *\*k<sup>er</sup>H-*; Steer 2015: 229, 229, footnote 719)

Katz (2003: 132–133) had already derived the Khanty word mentioned above, as well as Hungarian *szar*, from Indo-Iranian *\*sákr̥t-* (> OI *\*sákr̥t-*) as parallel loans, but he assumes that the rest of the forms reflect a PU noun *\*śara*. It is argued here that the PU word can be

explained as an Indo-Iranian loan, and that the Khanty word probably does not belong here.<sup>59</sup>

The Uralic origin of Hungarian *szar* is well-known and can be found in all recent etymological dictionaries of Uralic or Hungarian (UEW; TESz; Zaicz 2006 s.v. *szar*). However, the cognate set that these sources present is partly outdated. The Khanty word (Kaz., Ni.) *śor* ‘Kot’ with ś- cannot be related to the Hungarian word because of the obvious irregularity. It is possible that the word is a parallel loan, although also this is unlikely because of the sibilant. The UEW attempts to explain the Khanty sibilant with “Reihenübergang \*ś > \*ć” or “durch einen Wechsel \*ś ~ \*ć in der Grundsprache”. The latter explanation is impossible, and also the attempt to derive Khanty ś from secondary (?) PU \*ć would be irregular, as there is very little evidence for the varying outcome of \*ś and \*ć in Khanty. Aikio (2015) presents a set of regular cognates from Hungarian, Mari and Mordvin. In the UEW the Proto-Uralic stem vowel is left unspecified, but Aikio reconstructs the word as an \*a-stem. For the initial syllable \*a is more likely than \*i because of Mari *o ~ a* (< PMA \*ā), which usually reflects PU \*a.

TESz mentions that a Slavic origin for the Hungarian word has been suggested in earlier research, but TESz considers it wrong. The exact Slavic source is left unmentioned, but this Slavic source is probably Russian *сепу, срать* ‘to shit’; the Slavic etymology of the Hungarian word is obviously wrong because of the wide distribution of the word within Uralic, but as the Slavic word has an Indo-European etymology, it is not at all impossible that the Uralic and Slavic words are in some way related. LIV (326) derives the Slavic verb from the PIE root \**ker(H)-*, which is also reflected by Latin *mūs-cerda* ‘mouse shit’ and Avestan *sairiia-* ‘Mist, Düngé’. LIV further states that the root might be identical to the set root \**kerh<sub>2</sub>-* which is widely reflected in various branches of Indo-European. However, here the crucial thing is that the root denoting ‘shit’ in Proto-Indo-European has a cognate in the Indo-Iranian branch. This cognate would be a perfect fit for the origin of the Uralic word: Avestan *sairiia-* reflects Proto-Indo-Iranian \**car(H)ya-*, and although only the noun is attested, the root is originally verbal anyway, so postulating an Indo-Iranian root \**car(H)-* is unproblematic. This root, if borrowed, would result in Uralic \**śara-* (~ \**ćara-*).

Theoretically also a Proto-Balto-Slavic source would be possible, but the Uralic vowel would be much more difficult to explain from PBSl \**śer(H)-*.

Steer (2015: 229) reconstructs the PIE noun as heteroclitic \**sók-r/sék-n* (based on Hittite *šakkar : šakn-* ‘excrement, dung, faeces’); and mentions that the Avestan and Russian words, as well as the ON noun *skarn* ‘Düngen, Mist’ (PG \**skarna-*) continue a formation \**skor-*. Although this is radically different from the reconstruction of LIV in PIE terms, this does not radically influence the etymology of the Uralic word, as the PII predecessor of the Avestan word would have in this case been something like \**ścar-*, which would have easily produced PU \**śara-*. De Vaan (2008: 109) reconstructs the predecessor of the Latin noun as \*(*s*)*ker-d-* and assumes a relation to the root \*(*s*)*ker-* ‘to cut’. The views of Kroonen (2013: 440) are largely the same as the ones expressed by Steer. Kroonen does not mention the Avestan word, but he does mention Latvian *sārni* m.pl. ‘excrement’.

---

<sup>59</sup> This etymology is also included in the unpublished manuscript of Koivulehto, but as it has not been published, and as I came up with the etymology before I had the possibility to consult Koivulehto’s manuscript, it is presented as a new etymology here. Furthermore, many phonological details presented here were not mentioned by Koivulehto (ms.).

In Mari, also a noun *šur* (< PMA \*šūr) ‘shit’ is found, in addition to the verb *šora*, *šara* (< PMA \*sāra-). The noun is not mentioned by Aikio (2015b) but is listed among the cognates by UEW. The Mari noun and verb cannot be connected through any regular derivational process, and their relationship remains obscure. It is possible that the noun \*šūr reflects a later, parallel borrowing into Mari.

In summary, the resemblance of the Indo-European and Uralic words is hardly accidental, and the Indo-Iranian origin is the most suitable for the source for Uralic \*šara-, even if there are several possibilities for the ultimate Indo-European etymology of the Iranian word.

### **PU \*tora- ‘fight’**

← PII/PI \*tarH-, \*tr̥H- ‘to overcome’

See above in the entry \*tora-, where this etymology is discussed instead of the earlier suggestion by Aikio (2015b).

## 3. RESULTS

In this last chapter, the results of the etymological analysis are summed up. The convincing, uncertain and unconvincing etymologies are listed, and the sound substitutions and the distribution of the loanwords are presented and discussed. As the substitutions were presented in the Introduction already, here only the more problematic and relevant cases are discussed in further detail.

### 3.1. ANALYSIS OF THE SOUND SUBSTITUTIONS

Here the sound substitutions occurring in the convincing etymologies are discussed.

### 3.2. VOWELS

#### 3.2.1. PII/PI \*a and \*ā

It is clear that PII \*a and \*ā are substituted by several different Uralic sounds: \*a, \*o and \*i. It has been pointed out many times that this most probably due to the realization of \*a in Indo-Iranian. Because of the very simple vowel system, the Indo-Iranian phoneme \*a was probably realized in widely different ways. However, it is not that clear how much the different substitutions in loanwords are due to the different sound environments and the resulting allophony of \*a.

Uralic \*a seems to be the most frequent substitution for PII and PI \*a, although the merger of \*a and \*i in many branches makes it impossible to determine the exact substitution in a large number of cases.

Unambiguous cases of \*i ← \*a are few. Most of these are either Proto-Indo-Iranian loanwords or loans that cannot be securely classified as either Proto-Indo-Iranian or Proto-Iranian. This means that we have solid evidence for this substitution only among the Proto-Indo-Iranian loans, as well as among a couple of later Iranian loanwords. The Proto-Iranian loans often have a distribution restricted to West Uralic, where the difference between \*i and \*a cannot be seen in \*a-stems. None of the clearly Proto-Iranian etymologies can be certainly considered as \*i loans. However, as the “Old Iranian”-type loan \*sīrńa in Ugric does show this substitution, it is possible that some of the \*a-vowels in western Uralic can reflect earlier \*i. Interestingly, a couple of Proto-Mansi words of later Iranian origin show similar substitution, the reflex of PI \*a is substituted by Mansi \*j in \*wīsiy and possibly \*pńtV.

The \*o-loans present a peculiar case, as some of the o-loans can reflect Pre-II \*o (these are discussed separately below), but most of the loanwords clearly show \*o as the substitution for \*a. Of the rules suggested in earlier research, the view that PII and PI \*a was substituted by \*o before \*r is supported by a number of examples, even though also several counterexamples can be found. The vowel \*o also prevails after labial consonants, although there are several counterexamples to this rule as well.

Even though there is reason to assume that PII *\*a* and *\*ā* were phonetically rather different, at least based on their later reflexes in Sanskrit, where *\*a* was phonetically probably [ə] or [ʌ] and *\*ā* [e:] (Lubotsky 2018: 1875), in loanwords this difference cannot be observed, but both phonemes behave similarly. There are no substitutions that are unique for short *\*a* but not for its long counterpart, and vice versa.

Here it has to be noted that the number of loanwords which were borrowed from forms with *\*ā* is rather low, and even in these cases it often cannot be determined for certain that the donor form indeed had long *\*ā*. Many of the long vowels were caused by contraction of earlier *\*VH*, and there is no way to prove whether the laryngeals had already caused the lengthening of the vowels by the time the words were borrowed into Uralic. In some cases the long vowel is caused by the lengthened grade, namely in *\*ārya-* (if this is indeed a vrddhi-formation of *\*ari-* ‘stranger’) and *\*dāywa-*, *\*dāiwya-* (from which Finnic *toivo-* ‘to wish’). Both of these words are *o*-loans, but *\*orja* ← *\*ārya-* can be explained by the substitution rule that *\*a* (and *\*ā*) is reflected by Uralic *\*o* in front of *\*r*.

It is also notable that there seems to be no difference in the substitution of PII *\*a*, *\*ā* and later Iranian *\*a*, *\*ā*, and that loanwords with western (West Uralic or “Finno-Permic”) and eastern (Ugric) distribution show similar substitutions, with the exception of the early *\*ā*-loans in Ugric (see below).

It has been shown in the analysis of the etymologies that the Indo-Iranian accentology (the role of stressed and unstressed syllables) does not offer enough evidence for the different substitution of Indo-Iranian vowels, contrary to the hypothesis of J. Häkkinen (2009: 21–24).

Although modern Ossetic shows various conditional developments of PII *\*a*, in early Alanic loans *\*a*-substitution prevails. However, the vowel substitution in the Alanic loanwords manifest a murky picture. In Alanic-type loans of West Uralic, *\*a* is reflected by Uralic *\*a* (? *ɨ*) in Pre-Fi *\*arwa* and *\*wasa*. The Alanic reflex of PI *\*r* likewise yields West Uralic *\*a* (? *ɨ*) in *\*warsa*. In the more eastern loans, Hu *ár* probably reflects earlier *\*a* as well, although from a Uralic point of view *\*o* would be equally possible. Mansi *\*šār*kVśV ‘eagle’, which likewise is a probable Alanic loan, shows *\*ä*.

Due to the scarcity of the material, it is difficult to see to which extent the substitution patterns of the Alanic-type loans differ from the earlier ones. Probably here too several different layers of Alanic loanwords are present, and in any case the loans have been acquired separately into the western and eastern Uralic languages. However, it seems clear that the *\*a* : *\*ä* vocalism of the loans does not correspond to the later differences in the Ossetic vocalism.

Uralic *\*ä* does appear as the substitution for Indo-Iranian *\*a* in a number of well-known loanwords in Ugric and Permic, which probably belong to the later layers of loanwords. This has been seen as a sign of later Iranian vowel developments (Korenchy 1972; Rédei 1986). However, it has been shown here that *\*ä* appears also in earlier loanwords such as *\*säptä* and *\*päčäy*, and there is little reason to assume that *\*a* had become fronted by that time. It is more likely that the *\*ä* loans simply reflect a sound substitution of PII/PI *\*a*, for the same reasons that were discussed above.

The early *\*ä*-loans include *\*pätäri-* ‘to fall’ (the only word with a wide Uralic distribution, with reflexes in Saami and Mansi), *\*päčäy* ‘reindeer calf’ (from Proto-Iranian). PUg *\*säptä* ‘seven’, Pre-P *\*änkärä*. It is more difficult to state how old loan PMA *\*pärcä* is.

### 3.2.1. PRE-II \*e

Although there are versatile substitutions of PII \*a, the \*e-vocalism in Uralic loans clearly points to Pre-Indo-Iranian times, as it always reflects PIE \*e (the case is thus different from the \*o-loans, which can also reflect \*a from earlier \*e), that is, there are no possible loan etymologies where the Uralic form had \*e but the Indo-Iranian form reflected Indo-European \*o or \*a.

(\*ertä ‘side’), \*kečrā ‘spindle’, \*mekši ‘bee’, (\*meti ‘honey’), peijas ‘feast’, \*pejmā ‘milk’, \*peji- ‘to milk’, \*pertä ‘wing’, \*repä ‘fox’, \*rečmä ‘rope’

### 3.2.2. Pre-II \*o

Regarding the possible Pre-II o-loans, it has been suggested here that in a couple of cases we have a reason to assume that the Uralic form reflects retained \*o-vocalism:

\*ońca ‘part’, \*orpa ‘orphan’ (can also be from other IE)

### 3.2.3. \*r

It is rather difficult to distinguish between true reflexes of syllabic resonants and possible later Iranian reflexes where an epenthetic vowel ə might have developed before \*r, as all the loanwords show various combinations of a vowel and \*r in Uralic, due to obvious phonotactic reasons. In some cases it is also difficult to say whether the word has been borrowed from an Indo-Iranian form with \*r or \*ar.

### 3.2.4. \*r<sub>h</sub>

Evaluating the substitution of PII \*r<sub>h</sub> and its Iranian reflexes is challenging, as there was no phonetically close equivalent in this sequence of either of its phonemes. It is thus rather difficult to determine whether the substitutions like \*ar really mirror the Iranian development \*r<sub>h</sub> > \*ar, or whether they are substitutions of the PII cluster. One could maybe claim that it is the most economical solution to consider the instances of \*ar clear proof of the Iranian origin of these words, but since there seem to be otherwise few traces of laryngeals in the loanwords, the substitution of \*r<sub>h</sub> can be seen as an equal process compared to substitution of the simple \*r, and in these cases there are a number of different vowel + \*r clusters in the different words in Uralic; but the consistency in substituting this with specifically \*a and no other vowel could indeed point to Iranian.

However, it is clear that there are no cases of Uralic \*ir in place of PII \*r<sub>h</sub>, so this cluster is never reflected in a way that would point to Indo-Aryan donor language. Helimski’s and Zhivlov’s Andronovo Aryan has \*ar in this place, as reconstructed by Zhivlov (2013), but this probably a mirage created by sound substitution. According to Cantera (2001) and de Vaan (2003), PII \*r<sub>h</sub> became PI \*ər, and there is no reason not to suppose a similar phonetic realization to PII already. It is unclear whether the phonetic realization of \*r<sub>h</sub> in Proto-Indo-Iranian was different, and whether there are any

differences of PII *\*r̥H* and its later Iranian reflex. In any case, there is no reason to consider the substitution of *\*r̥H* by *\*ar* to be a feature which points particularly to Iranian, or to Andronovo Aryan.

### 3.2.5. The substitutions of PII and Pre-II vowels and the vowel reconstruction by Tálós (1984; 1987) and Abondolo (1996; 1998)

As noted in the Introduction, Kümmel (2017, 2018a, 2019) has attempted to derive the vocalism of the Indo-Iranian loanwords from the system advocated by Tálós (1987) and Abondolo (1996; 1998), with two tongue heights for vowels. Kümmel argues that if the traditional reconstructions of Uralic *\*e* and *\*o* should rather be reconstructed as *\*ä* and *\*ā*, respectively, one can simply assume that the cases traditionally reconstructed with Uralic *\*e* or *\*o* simply reflect different front-and back-vocalic substitutions of Proto-Indo-Iranian and Proto-Iranian *\*a*. These languages did not, of course, distinguish between front and back vowels, so this idea is plausible.

While the general plausibility of this alternative vowel reconstruction cannot be decided here, based on the results of Chapter 2 it is clear that all the substitutions which occur in the convincing Indo-Iranian etymologies can be explained from the “traditional” reconstruction of Proto-Uralic vowels. The idea of reconstructing *\*e* rather than *\*ä* for the Uralic words that are Pre-II loans fits very well the Indo-European vocalism reconstructed for these words (= there are no cases where Uralic shows *\*e* and we are forced to assume that PII *\*a* reflects some other PIE vowel), and also the possible cases of retained Pre-II *\*o* fit the traditional system.

Some other arguments against the plausibility of Tálós’s Proto-Uralic vowel reconstruction is that the assumed vowel length in Uralic does not correlate with the vowel length in Indo-Iranian. This is necessarily not an obstacle, though, as one can suppose that because the short and long *\*a* in Indo-Iranian were in any case not distinguished only by length but by quality too (see the Introduction), the difference in length would not have been necessarily crucial for the speakers of Tálós-type Proto-Uralic. However, there are almost no examples of the substitution of Indo-Iranian *\*a* by short *\*ä* in the earliest loans, although substitutions by long *\*ā* prevail.

All this does not mean that the reconstruction with two tongue heights would certainly be wrong, or that the traditional reconstruction would not have its problems, simply that resorting to the alternative reconstruction of Uralic vocalism does not present a more coherent and regular picture of vowel substitutions than the commonly accepted view. Future research on Uralic vocalism will show whether there are further arguments to support the vowel reconstruction of Tálós.

### 3.2.6. Non-initial syllables

On the basis of the etymological analysis, it is clear that the overwhelming majority of the Indo-Iranian loans were borrowed as *\*a*-stems. The number of *\*i*-stems is surprisingly low compared to the roughly equal occurrence of *\*a*- and *\*i*-stems in the Proto-Uralic vocabulary. Among Proto-Indo-Iranian or Proto-Iranian loans, only *\*šarwi* and *\*jini* were

certainly borrowed as *i*-stems, but there exist a number of words which theoretically could be either *\*a-i* or *\*o-a*-stems, as the reflexes of these two stem types merged in Mordvin and Saami, and reflexes of *\*a* and *\*o* are in general not easily distinguished in Mari and Hungarian.

Many Pre-Indo-Iranian loans were borrowed as *ä*-stems. It seems that Pre-Indo-Iranian *o*-stems were acquired as *a*-stems, while *i*-stems and *u*-stems (*\*mekši*, *\*medhu-*) were borrowed as Uralic *i*-stems.

The idea of possible *\*o*-stems in Proto-Uralic is a disputed idea, and among the loanwords, PU *\*asora* ← PII *\*asura-* is the only possible example. Other loanwords such as *\*počaw* ‘reindeer’ show that the roundedness of a vowel in the donor language could have been reflected by a sequence of a vowel and glide in Uralic, which casts doubt on the idea that *\*asora* should be reconstructed with *-o*.

For the loans with the ending *-as*, no regular substitution rule can be established. All the convincing cases reflect Indo-Iranian masculine nouns, but there are a number of counterexamples where *-as* is not reflected, but rather the word was borrowed as a simple *a*-stem. It has been repeatedly noted in earlier research that the ending *-as* was probably acquired via loanwords (Helimski 2000: 194; Nikkilä 1994: 104–; Hahmo 1997: 95–96), and this idea is convincing. Probably the *\*as*-stems became productive in the western Uralic languages at some point after the divergence of Proto-Uralic.

It is also notable that those later Iranian loans that are preserved only in Permic or Ugric tell us little about the stem vowels, because in these branches the PU stem vowels either underwent total apocope or at least merged and reduced to such a condition that it is almost impossible to determine the original stem type, although the initial-syllable vowel correspondences often tell us something of the original stem type of the word.

Koivulehto (2009) has stated that the choice of an *\*i-* or *\*a-*stem in loanwords was influenced by the weight of the syllable, with *\*i*-stems appearing in words which have a heavy first syllable. However, this is clearly an over-simplified statement, as there exist more stem types than these two among the borrowed Indo-Iranian words, and in the light of modern research the “heaviness” of the stem cannot necessarily be reconstructed for Proto-Uralic in the way that Koivulehto assumed.

### 3.3. Consonants

#### 3.3.1. Laryngeals

##### 3.3.3.1. VHV

Parpola has argued that the spirant *\*y* represents a substitute for the laryngeal in the etymology of *\*sayi-* ‘to get, obtain’ (← *\*saH*). This is, in principle, convincing, but because it is not backed by parallel examples the evidence is not decisive.

### 3.3.3.2. #HV

There are only very few possible examples of a consonantal substitution of the word-initial laryngeal. It seems probable that the word-initial laryngeal, if it was retained, was not substituted in any way in Uralic.

\**karšV* (> Fi *karhu*), an uncertain etymology, is the only possible example.

### 3.3.3.3. CH

In a few examples Koivulehto has argued that the laryngeals were substituted by Uralic \**k*. Uralic \**ćolki* is probably not an Indo-Iranian borrowing, but \**siŋka* ‘old’ almost certainly is. As argued in the entry for \**soŋka*, it is possible that the \**k* in the Indo-Iranian original \**sanaka-* reflects a “hardened” laryngeal, but this development is disputed by Indo-Europeanists. Even if \**k* was a result of laryngeal hardening, the development would probably be earlier than Proto-Indo-Iranian, meaning that by the time the word was borrowed, the Indo-Iranian word simply had the stop \**k* that was regularly substituted by Uralic \**k*.

## 3.3.4. Stops

### 3.3.4.1. Unvoiced stops

There is nothing special in the substitution of the plain unvoiced stops of Proto-Indo-Iranian and Proto-Iranian, as they had counterparts in Uralic and were thus reflected rather faithfully as corresponding unvoiced stops in all positions. This can be said of all unvoiced stops, except in consonant clusters; in the word-initial cluster \**st* the *t* element was dropped, as happened in early Germanic and Baltic loans as well. There are not many examples of this development in Indo-Iranian loans (\**sampas* ‘pillar’ being perhaps the only one), but this suits the general substitution pattern of the early loanwords into the Uralic languages.

### 3.3.4.2. Voiced and aspirated stops

As there were no voiced stops in Uralic, the Indo-Iranian voiced stops were reflected by their unvoiced counterparts in the loanwords. The situation is similar to the other early Indo-European loans.

The same can be said of the aspirated stops, which were likewise substituted by Uralic \**p*, \**t*, \**k*. Because of this, it is impossible to distinguish the substitutions of Proto-Indo-Iranian aspirated stops from their later, deaspirated reflexes in Iranian.

### 3.3.4.3. Iranian \*x, \*θ, \*p

The Iranian fricatives that developed from unvoiced stops in various clusters were substituted by Uralic stops (for example, \*šukta ← \*tsuxta-), and in most cases it is impossible to determine whether the Uralic word reflects PII stops or later fricatives.

### 3.3.5. Sibilants

#### 3.3.5.1. PII, PI \*s

The substitution of PII \*s is uniformly Uralic \*s. Because of the large sibilant inventory of both Uralic and Indo-Iranian, it was probably never the case that \*s was substituted by other Uralic sibilants, unlike in the layer of earliest Germanic loanwords in Finnic, where Proto-Germanic \*s is often substituted by Early Proto-Finnic\*š.

As noted above, in the \*st- cluster in word-initial position, only \*s is reflected in Uralic. This fits well with the evidence from other Indo-European (Baltic and Germanic loans), but is represented by only one Indo-Iranian example.

The etymology of the Uralic word \*poši ‘penis’ is a notable example to this substitution. Here it would be unlikely that the Indo-Iranian and Uralic words are only coincidentally similar, but it is impossible to explain the Uralic \*š.

There is no evidence for Iranian \*s > \*h in early loanwords. On the other hand, there is also no evidence that would force us to consider \*s > \*h a late change, as words with retained \*s do not show other obvious Iranisms, such as the affricates \*ts or \*dz.

#### 3.3.5.2. PII, PI \*š

The reflexes of the sibilant \*š (originally the RUKI allophone of \*s) are clearly distinguishable in the loanwords. The substitution in PII and PI loans is unambiguously \*š, and the loans point to the evidence that the “RUKI change” was very early, predating even the Indo-Iranian merger of non-high vowels. The loans which do not show the RUKI reflex, such as \*warsa, are probably later loans from a language in which \*s and \*š merged (as certainly happened in Alanic). The apparent exceptions to \*š ← \*š are not convincing, so one can state that Uralic \*š was the regular substitution of Proto-Indo-Iranian \*š.

In some loans that are found only in Finnic, the cluster *rs* could theoretically reflect also earlier \*rś, and it is impossible to know whether the substitution was \*rs or \*rś. If \*rš was not allowed in Proto-Uralic, it is also possible to assume that the apparent exceptions can be explained by the substitution \*rś ← \*rš.

There are a couple of loans in Permic which show \*ś from \*š. In the case of *buriś*, it is possible that the folk etymology ‘good hair’ has played a role, and \*ś can be explained as secondary. In the etymology of \*śumis this cannot be the option, and it is possible that the substitution was \*ś ← \*š. Here we have to take into account that the Pre-Permic sibilant was probably \*š rather than \*ś. Mari \*pārća is also an exception to the rule that Uralic \*š is the substitute of PII \*š.

Beekes (1997: 7) reconstructs the RUKI sibilant as a retroflex \*ʂ. This is known to be the reflex of the sound in Old Indo-Aryan, but the retroflexion of \*ʂ can be attributed to a more general phenomenon of retroflex consonants in Indo-Aryan (Burrow 1955: 96–99). There is no direct evidence for retroflexion in Old Iranian, and also Beekes assumes \*ʂ as the Proto-Iranian reflex of this sibilant. It is of course doubtful how well we can reconstruct the exact realization of this sibilant in Proto-Indo-Iranian anyway. There is also the terminological question of what is meant by “retroflexion” exactly. The Uralic loanwords do not help in this, as Proto-Uralic \*ʂ was in any event the most fitting substitution for either \*ʂ or \*ʂ.

### 3.3.6. Affricates

#### 3.3.6.1. PII \*ć, \*ǰ, \*ǰʰ

The “palatal” character of \*ć and \*ǰ is always reflected in the early loans, as they are substituted by Uralic \*ś (\*ć?). This is one of the most decisive and often the only feature to distinguish later Iranian loans from earlier Proto-Indo-Iranian borrowings. In most cases it is impossible to decide whether the Uralic substitute was the sibilant \*ś or the affricate \*ć, supporting the claim by Zhivlov (2014: 114, footnote 3) that these were one consonant in reality. If Pre-Ob-Ugric \*śarapa ~ ćarapa ‘elk’ (which I deem uncertain) and Proto-Mansi \*āćtər are possible counter-examples, as these words do not give the usual reflex of \*ś > s in Mansi. There seems to be no difference in the substitution of \*ć and \*ǰ, ǰʰ. For the latter, there are only a few examples, and interestingly they have a West Uralic distribution (\*śasra ~ \*śisra, \*waćara, \*woraći).

#### 3.3.6.2. PI \*ts, \*dz

PI \*ts is substituted by \*ʂ in word-initial position. In word-internal position it appears as \*č as Koivulehto initially argued, and it seems that this substitution prevailed in both the area of West and Central Uralic, as well as in more eastern areas as is shown by POUG \*pächäy. The reasons for word-initial \*ʂ instead of \*č remain unclear, as \*č would have been possible in word-initial position in Proto-West Uralic (and Proto-Uralic).

The possible case of \*ʂ ← \*ts in word-initial position, Uralic (?) \*(j)iša ‘skin’, is a false etymology.

Koivulehto’s idea of the substitution \*ks ← \*ts can be rejected, as it is based on one etymology only, \*teksä ‘ten’, which has been shown to be unconvincing. The substitution \*ks ← \*dz is supported only by two possible examples. It would be crucial to have more evidence. At this point the substitution rule remains hypothetical. The cases where \*dz is substituted by \*č are not convincing.

The cases that possibly show \*s from \*dz are too uncertain. It is possible that the convincing etymologies that apparently show this substitution can be explained from later \*z.

### 3.3.6.3. PII & PI \*č, \*ǰ

There are very few examples of these sounds, making it difficult to determine anything with certainty. The substitution of these sounds seems to be practically similar to those of PI \*ts and \*dz. Pre-Permic \*poči- shows \*č from Indo-Iranian \*č. Among the uncertain etymologies there are examples of possible substitutions of \*ǰ with Uralic \*š or \*č, but because these etymologies are not convincing, there is no point in discussing this further.

### 3.3.6.4. Later Iranian \*s, \*z (from \*ts, \*dz)

These are reflected by Uralic \*s in all loans. Koivulehto argued that Uralic \*s- in word-initial position could substitute for \*dz. If this was correct, there would be no way to distinguish \*z and \*dz in the word-initial position in the loanwords, but there is no positive evidence for this (see above).

### 3.3.6.5. Glides

The glide \*y was substituted with the corresponding Uralic glide \*j in all positions. The same can be said of \*w.

### 3.3.6.6. \*l, \*r

The loanwords support the view that PIE \*l became \*r already in Proto-Indo-Iranian, as there are no plausible loan etymologies which show Uralic \*l as the substitute of l that could be reconstructed based on Sanskrit forms that have allegedly retained PIE \*l. All the convincing loanwords point to PII or PI \*r, and the few tentative words with \*l (\*śali, \*śolki, \*talas) have been shown to be false or at least improbable etymologies (some being borrowings from some other branch of Indo-European). The substitutions also show that the shift to r took place very early, as it is reflected even in the Pre-Indo-Iranian borrowings like \*kekrä and maybe \*kečrā, if this word goes back to a PIE form with \*-tlo- and not \*-tro-.

### 3.3.6.7. Evidence for Andronovo Aryan and Indo-Aryan loanwords?

None of the loanwords have to be considered as Andronovo Aryan or Proto-Indo-Aryan based on the criteria that were presented in the Introduction. The Uralic palatal affricate \*č or sibilant \*ś can in all cases be explained from Proto-Indo-Iranian \*č, and there is no need to assume that it should reflect Andronovo Aryan \*č or PIA \*ś. In the etymological material of this study, no further positive evidence was found for the distinction of PU \*ś and \*č as substitutions of the Proto-Indo-Iranian affricates. This means that at least in word-initial position there probably was no difference between \*č and \*ś, and even though we do not

know what this sound was phonetically, it is safe to assume that Uralic words showing \*ś reflect a sound substitution of Indo-Iranian \*ć and \*ź.

Regarding the other phonological criteria, it was shown already in the Introduction that the substitutions of \*r<sub>̣</sub>H clusters are so diverse that one cannot determine what their development was in Andronovo Aryan.

The various substitutions of Indo-Iranian \*a cannot be used as evidence for either Indo-Aryan or Andronovo Aryan loans. The possible Andronovo development \*a > \*ä is also shown by some loans that are clearly from Proto-Iranian, which means that Uralic \*ä reflects a simple sound substitution, not a development \*a > æ on the Iranian side.

Regarding the distribution of the etymologies within Indo-Iranian, all the loanwords which cannot be from Iranian because of the lack of attested Iranian cognates have a more or less secure Proto-Indo-Iranian etymology, and nothing prevents us from assuming that these words reflect Proto-Indo-Iranian borrowings. It is also possible that some words with solid Proto-Indo-Iranian etymologies were present in Iranian but were lost before the first Old Iranian texts were composed.

### 3.4. LIST OF CONVINCING, UNCERTAIN AND UNCONVINCING INDO-IRANIAN ETYMOLOGIES

Here the etymologies are listed and stratified into different loanword layers. The loanwords that can belong to several layers are mentioned in brackets. The words that can be either from Proto-Indo-Iranian or Proto-Iranian are listed as one group. Here the phonological criteria is used as the main criterion and those loanwords that do not show Proto-Indo-Iranian, Proto-Iranian or later Iranian features are listed under the ambiguous loanwords. In the case where the distribution within Indo-Iranian or semantics give definite traits of Iranisms, the words are listed as Iranian, but otherwise they are listed as ambiguous, even if they have a limited distribution within Uralic.

#### 3.4.3. Pre-Indo-Iranian

\*ertä ‘side’, \*kekrä ‘wheel’, \*kečrä ‘spindle’, \*mekši ‘bee’, (\*meti ‘honey’), \*ońća ‘part’, (\*orpa ‘orphan’), \*peijas ‘feast’, \*pejmä ‘milk’, Pre-P \*pertä ‘wing’, \*repä ‘fox’, \*rečmä ‘rope’, \*sejti ‘bridge’

#### 3.4.4. Proto-Indo-Iranian

\*áčtara ‘whip’, \*anti/onta, \*ora ‘awl’, \*orja ‘slave; south’, (\*orpa ‘orphan’), \*poši ‘penis’, \*šańka ‘handle’, Pre-Md \*šaya ‘goat’, \*šarwi ‘horn’, \*šađa- ‘to rain’, šara- ‘shit’, \*šita ‘hundred’, Pre-P \*šVta ‘hundred’, \*šasra ‘thousand’, \*šišta ‘wax’, \*šoma- ‘sad’, \*wačara ‘hammer’, \*worači ‘boar’

#### 3.4.2.1. Uncertain cases

\*čańka- ‘to sting’, \*čara- ‘brown; ? to dawn’, \*čarapa ‘elk’, \*čero ‘hill-top’,? \*čerti ‘group’, \*porčas, PUg \*šeńkV ‘nail’

### 3.4.3. Proto-Iranian

(\**ačwa* ‘horse’) \**mača* ‘insect’, \**očra* ‘barley’, \**počaw* ‘reindeer’, \**päčäy* ‘reindeer’, \**serä* ‘old’, \**šoji-* ‘to care’, \**šukta* ‘burnt patch’

#### 3.4.4.1. Uncertain cases

\**maksa-* ‘to pay’, \**paksu* ‘thick’

### 3.4.5. Old Iranian and ambiguous early Iranian (no PI criteria, no Osseticisms)

\**asma* ‘voracious’, PP \**domV* ‘to tame’, Pre-P \**dura /tura* ‘long’, Pre-Fi \**isa-* ‘appetite’, Pre-P \**kaða-* ‘to steal’, Pre-P \**karka* ‘chicken’, Pre-P \**parša ~ \*barša* ‘mane’, Pre-P \**parta* ‘knife’, PFi \**pərna* ‘spleen’, PMs \**rāsnā* ‘rope’, \**sernā* ‘metal’, \**siṛnā* ‘metal’, PFi \**sooja* ‘shelter’, PP \**šud* ‘luck’, PP \**šumīs* ‘band’

### 3.4.6. Alanic/“Pre-Ossetic”

\**arwa* (> Fi *arvo* ‘price’), \**arwV* (> Hu *ár* ‘price’), POUg \**ērāy* ‘song’, POUg \**eträ* ‘clear sky’, PKh \**layər* ‘Panzerhemd’, PMs \**širyV* ‘sword’, \**šārkVšV* ‘eagle’, \**warsa* ‘calf’, \**wasa* ‘foal’, PMs \**wīsəy* ‘calf’, PMs \**āsVrVmV* ‘shame’

These words probably represent several different chronological layers, but due to the small number of securely “Alanic” loans, far-reaching conclusions cannot be made. Some words, like \**šārkVšV* or *wasa* ‘foal’, PMs \**wīsəy* ‘calf’ could also have been borrowed from some other later Iranian language, as similar phonological developments are attested in Chwarezmian. \**layər* could equally well be borrowed from a pre-form of the Pashto word *zyar*, but due to geographical reasons Alanic is a more probable origin than other forms of Eastern Iranian.

### 3.4.7. Ambiguous early loans (can be either from PII or PI)

\**ajša* ‘shaft’, \**asVra* ‘lord’, \**iha* ‘yearning, passion’, \**ihta* ‘lust’, \**jama* ‘twin’, \**jawi/jowa* (> Mo *juv*) ‘awn’, \**jawi* (> PS \**jāš*) ‘flour’, \**jini* ‘way, path’, \**juma* ‘god’, \**kana-* ‘to dig’, \**kara-* ‘to dig’, \**kata-* ‘to graze’, \**kertä-* ‘to bind’, \**kintaw* ‘tree stump’, \**kürtiV* ‘iron’, PKh \**kärtV* ‘iron’, \**kärtä* ‘iron’, \**martas* ‘dead’, \**nātV-* ‘to help’, \**pakas* ‘god’, \**para* ‘good’, Kh *pənt* ‘way’, PMs \**pəntV* ‘brother-in-law’, \**pora* ‘old’, \**poči-* ‘to boil’, Pre-P \**porta* ‘vessel’, \**puntaksi* ‘bottom’, Pre-Ma \**pänti-* ‘to bind’, PMa \**pärca* ‘ear of corn’, \**pätäri-* ‘to flee’, \**sayi-* ‘to get, obtain’, \**sampas* ‘pillar’, \**sanjka* ‘old’, \**sara* ‘lake’, \**sasara* ‘sister’, \**säptä* ‘seven’, \**tajwas* ‘sky’, \**takra* ‘piece of flesh’, \**tarna* ‘grass’, \**tojuV* ‘wish’, \**toraksi* ‘through’, \**tora-* ‘to fight’, \**täjV* ‘milk’, \**täjinV* ‘cow’, \**täši*, \**uška* ‘bull’, \**wakša-* (> PS \**wātā-*) ‘to grow’, \**wajna-* ‘to see’, \**wojna-* ‘to see’, \**wiša* ‘venom’, \**wirna* ‘wool’, \**wärkä* ‘kidney’, PS \**wärkâ* ‘wolf’, \**wirtV-* ‘to hold, raise’, \**änkärä* ‘coal’

#### 3.4.7.1. Uncertain cases

PFi \**aiwa* (← Germanic ?), Ma \**arša* ‘mane’, \**aštira* ‘barren earth’, \**itä-* ‘to appear’, Pre-Fi \**karšV* ‘bear’, \**kota* ‘chum’, Pre-Sa \**kupa* ‘pit’, PFi \**kärsä* ‘snout’, PFi \**mana-*, PUG ? \**mañci*, Ma *marij* ‘Mari; man; husband’, POUg \**meja* ‘wedding’, \**mykkä* ‘dumb’, \**orpV* ‘relative’, \**pijka* ‘psychedelic mushroom’, POUg \**pVrV* ‘phratry’, Pre-Sa \**poti* ‘against’, \**sentü-* ‘to be born’, Pre-Sa \**soma/sami* ‘some’, PP \**sur* ‘beer’, PFi \**süte-* ‘to hit’ (< ? \**sewči-*), PUG \**taja-* ‘secret’, Pre-Fi \**terni* ‘young’, \**terwV* ‘healthy’, ? \**towkV* ‘spring’,

PWU *\*utarV* ‘udder’ (← Germanic ?; Mari *\*waðar* ← II), *\*waŋka* ‘hook’, Md E *vérges*, M *vərgás* ‘wolf’

The ambiguity means that many of the loans with limited distribution in Finno-Permic and Ugric can be considered Iranian loanwords. If we take the wide distribution of *\*tora-* into account, it is more probable that this word is also from Proto-Indo-Iranian. The same can be said of other widespread loans, like *\*asVra*.

### 3.4.8. Other Middle Iranian/later Iranian

POug ? *\*mōŋki* ‘forest-spirit’, PS *\*pulā* ~ *\*pīlā* ‘bridge’, PP *\*sarid’z* ‘sea’

### 3.4.9. List of uncertain Indo-Iranian etymologies

PFi *\*aiwa* (← Germanic ?), Ma *\*arša* ‘mane’, PMs *\*ārV* ‘fire’, *\*aštira* ‘barren earth’, POug *\*ćakV* ‘hammer’, *\*ćara-* ‘brown; ? to dawn’, *\*ćero* ‘hill-top’, *\*ćerti* ‘group’, *\*itā-* ‘to appear’, Pre-Fi *\*karšV* ‘bear’, PMs *\*kīrV* ‘iron’, *\*kota* ‘chum’, Pre-Sa *\*kupa* ‘pit’, PFi *\*kārsā* ‘snout’, *\*maks-* ‘to pay’, PFi *\*mana-*, PUg ? *\*mańci*, Ma *marij* ‘Mari; man; husband’, *\*meja* ‘wedding’, *\*mykkā* ‘dumb’, PP *\*oč* ‘corn’, *\*orpV* ‘relative’, PFi *\*paksu* ‘thick’, *\*peji-* ‘to milk’, *\*pijka* ‘psychedelic mushroom’ POug *\*porV* ‘phratry’, Pre-Sa *\*poti* ‘against’, Pre-Fi *\*šatas* ‘germ’, *\*sentü-* ‘to be born’, *\*šerā-* ‘to wake up’, Ms *šVšwəŋ* ‘hare’, PUg *\*šeŋkV* ‘nail’, Pre-Sa *\*soma/sami* ‘some’, PP *\*sur* ‘beer’, PFi *\*süte-* ‘to hit’ (< ? *\*sewči-*), Hu *szekér* ‘wagon’, Kh *aiķər* ‘Narte’ PUg *\*taja-* ‘secret’, Pre-Fi *\*terni* ‘young’, *\*terwV* ‘healthy’, ? *\*towkV* ‘spring’, PWU *\*utarV* ‘udder’ (← Germanic ?; Mari *\*waðar* ← II), *\*waŋka* ‘hook’, Mo E *vérges*, M *vərgás* ‘wolf’

### 3.4.10. Etymologies that were probably borrowed from another Indo-European source (PIE, PBSI, Germanic, Baltic)

*\*aisa* ‘shaft’ ← Balto-Slavic, PFi *\*aiwa* (← Germanic ?), *\*apV* ‘help’ ← Germanic, *\*jewā* ‘grain’ ← Balto-Slavic, Ma *karaš* etc. ‘honeycomb’ ← Baltic, (*\*meti* ‘honey’ ← ? PIE,) Fi *\*ojas* ‘shaft’ ← Slavic, *\*ola* ← Baltic, *\*oŋki* ← Germanic, *\*porćas* ← Balto-Slavic, Pre-Sa *\*porta* ‘vessel’ ← Germanic, *\*salV* ‘salt’ (cannot be reconstructed for PU, various later parallel loans), *\*šilkaw* ← Balto-Slavic, *\*sammu-* ← Germanic, *\*šuka* ← Balto-Slavic, Mari *\*šūžar* ← Baltic/Balto-Slavic or Slavic, *\*tejniš* ‘pregnant animal’ ← Baltic/Balto-Slavic, PWU *\*utarV* ‘udder’ (? ← Germanic)

### 3.4.11. Unconvincing etymologies

PUg *\*arV* ‘kinship-term’, PSa *\*čārēs* ‘coarse’, *\*čača-* ‘to grow’, *\*eštā-* ‘to be able’, *\*idi* ‘year’, ? *\*icā* ‘father’, *\*icāntā* ‘master’, *\*išana* ‘wonderful’, *\*iði* ‘year’ *\*jiša* [?] ‘skin’, *\*juyi-* ‘to drink’, *\*jäyi-* ‘to leave’, *\*miyi* ‘earth’, PMs *\*mul’V-* ‘to win’, PFi *\*mürkkü* ‘venom’, Fi *naaras* ‘female’, *\*niska* ‘neck’, POug ? *\*nud* ‘river’, POug *\*nāk* ‘nail’, PUg ? *\*omtVrV*, Est *oide* ‘grass’, *\*pošja* ‘bottom’, Ma *\*pörš* ‘Reif’, Mo *pušto*, Ma *pušta* ‘Haferbrei’, PKh ? *\*rawV* ‘river’, Mansi *rēk<sup>w</sup>*, Khanty *rōwa* ‘Ufer’, *\*šali* ‘guts’, *\*sarajas* ‘sea’, PS *\*ser* ‘white’, PU ? *\*šertā* ~ *\*šerā* ‘Art’, *\*sini* ‘blue’, *\*šiwā* ‘good’, *\*šolki* ‘fibula’, Ma *šör*, *šəžer* ‘milk’, *\*tala* ‘vessel’, *\*talV* ‘hut’, *\*tarwas* ‘sickle’, Mns *tas* (*xum*) ‘Fremder’, *\*teksä* ‘ten’, Md *ukso*, Ma *oškā*, PSa *\*wāčəm*, *\*wari* ‘forest; hill’, *\*waji* ‘butter’, *\*waćća* ‘stomach’

### 3.5. EARLY LOANS INTO ALREADY DIFFERENTIATED BRANCHES

It is often difficult to distinguish the loanwords which are acquired into individual branches from those which were borrowed already into the proto-languages. Here only the convincing cases are listed.

Only in Finnic:

*\*ačnas* ‘voracious’, *\*iha* ‘wish’, *\*ih̥ta* ‘lust’, PFI *\*isV* ‘appetite’, *\*martas* ‘dead’, *\*očra* ‘barley’, *\*peijas* ‘feast’, *\*pejmä* ‘milk’, *\*peřna* ‘spleen’, *\*sampas* ‘pillar’, *\*sooja* ‘shelter’, *\*tajwas* ‘sky’, *\*takra* ‘piece of flesh’, *\*terwV* ‘healthy’, *\*tojuwV* ‘wish’

All of these words, with the exception of *\*sooja* ‘shelter’, were clearly borrowed into Early Proto-Finnic (Pre-Finnic) at the latest. Formally most of the loans could be from PII or PI: *\*peijas* and *\*pejmä* are probably from Pre-II, *\*sooja* and *\*isa* from later Iranian. *\*peřna* has to be formally from Iranian (PI or later).

Only in Saami:

*\*kata-* ‘to graze’, *\*kertä-* ‘to bind’, *\*pora* ‘old’, *\*wojna-* ‘to see’

All of the loans were acquired before the Saami vowel changes. Formally all could be either from Proto-Indo-Iranian or Proto-Iranian.

Only in Finnic and Saami:

*\*asma* ‘voracious’, *\*jama* ‘twin’, *\*kekrä* ‘wheel’, *\*mača* ‘insect’

Of these, *\*mača* from Proto-Iranian and *\*jama* is ambiguous. As the *-sm-* in *asma* does not point to Proto-Indo-Iranian *\*ć*, this is probably an Iranian loan too. It is possible that these words were borrowed into Proto-West Uralic, as there is no general support for a Finno-Saamic proto-language today. As the cognates within Finnic and Saami are regular, there is no need to assume parallel borrowings. *\*kekrä* has to be from Proto-Indo-Iranian.

Only in Mordvin:

*\*čaya* ‘goat’, *\*jawi/jowa* (> Mo *juw*), *\*kürtnV* ‘iron’, *\*pakas* ‘god’, *\*serä* ‘old’, *\*sernä* ‘metal’, *\*wajna-* ‘to see’

Most of these loanwords are ambiguous (they could have been borrowed from Proto-Indo-Iranian, Proto-Iranian or even from some later form of Iranian). *\*serä* and *\*sernä* have to be from Iranian because of the sibilant *\*s*, and it is unlikely that *\*kürtnV* ‘iron’ is from Proto-Indo-Iranian, for obvious historical reasons. *\*čaya* is the only loan that appears to be earlier than Proto-Iranian, but due to the unclear development of PII *\*ś*, it cannot be ruled out that this word is borrowed from Proto-Iranian too.

Only in Mordvin and/or Finnic and/or Saami (can point to a borrowing into Proto-West Uralic):

\**jini* ‘way’, \**kečrā* ‘spindle’, \**rečmä* ‘rope’, \**šaŋka*, \**wačara* ‘hammer’, \**warsa* ‘foal’, \**wasa* ‘calf’, \**worači* ‘pig’

Based on phonological criteria, these loans do not form a chronologically coherent layer, but probably their modern distribution is accidental (their original distribution can have been wider). \**kečrā* ‘spindle’ and \**rečmä* ‘rope’ are from Pre-II, \**šaŋka*, \**wačara* and \**worači* from PII, \**warsa* and \**wasa* from later Iranian (Alanic). \**jini* is ambiguous. Also the loans confined to Finnic and Saami mentioned above probably were borrowed into Proto-West Uralic, as it is a more convincing taxonomic entity than Proto-Finno-Saamic.

Only in Mordvin, Finnic and/or Saami and Mari

\**juma* ‘good’

This loan can be either from PII or PI. As it is obvious that these four branches do not form any taxonomical entity (Salminen 2002; J. Häkkinen 2009), it is only logical that there are no other loanwords with a “Finno-Volgaic” distribution.

Only in Mari:

\**kVrtnV* ‘metal’ (← PII, PI or later), Pre-Ma \**pānti-* ‘to bind’, PMa \**pārca* ‘ear of corn’, \**sijña* ‘gold’ (← Old Iranian)

Only very few early Indo-Iranian loans can be found in Mari and in no other Uralic language. It is unclear what the reason for this is. It is, of course, possible that some uncertain loanwords like *marij* ‘man; Mari’ turn out to be correct after all, but even that does not make the number of loans in Mari very high. The situation has to be explained either with loss of vocabulary and replacement by later loans (from Turkic, and also perhaps from Permic) or with Mari’s location on the periphery at the time of the later contacts with the Iranian languages. Agyagási (2019: 254–258)<sup>60</sup> argues that the current area where Mari is spoken was formed only relatively late, after the Mongol invasion in the High Middle Ages. If this is indeed correct, and Mari was spoken in more northern areas before that, it can be assumed that Pre-Mari had only sporadic contacts with the Iranian languages after it split off from Proto-Uralic.

Only in Permic (early loans; for later loans confined to Permic, see Appendix I):

\**a(č)wa* ‘stallion’, PP \**ju* ‘awn’, \**kertä* ‘house’, \**kärtä* ‘metal’, \**kada-* ~ \**gada-* ‘to steal’, \**karka* ‘chicken’, \**parša* ~ \**barša* ‘mane’, \**parta* ‘knife’, \**pertä* ‘wing’, \**poči-* ‘to

---

<sup>60</sup> Agyagási’s arguments are based on a combination of evidence from archaeology, mediaeval written sources and alleged substrate vocabulary, and her views require a detailed analysis and commentary before they can be accepted as such. However, from the point of view of Indo-Iranian loanwords the idea that the area of Mari speakers has changed radically during mediaeval times is possible.

boil', \**porta* 'vessel', \**dura* 'long', \**domV* 'to tame', PP \**šumīs* 'band', PP \**šud* 'luck', \**uška* 'bull', \**wirna* 'wool', \**wirä* 'man, husband', \**āṅkärä* 'coal'

The number of loanwords in Permic is relatively high, and many of these can be considered to be Iranian loanwords. Technically many loans are ambiguous, but as some of the words were borrowed late due to historical reasons ('iron'), and some were borrowed into a Pre-Permic which already had a phonological system that was different from Proto-Uralic (\**šud-* has *d* which cannot reflect PU \**š*).

It is probable that the Permic languages were in continuous contact with the Indo-Iranian languages from the time they split from Proto-Uralic until the early mediaeval era. Although many later loans listed in Appendix I show Ossetic phonological features, among the earlier loanwords there are not many traits that would force us to consider the words to be borrowings from specifically Alanic, but many could theoretically be from Proto-Iranian or a more generic "Old Iranian". This makes the more precise dating and stratigraphy of the loanwords in Permic difficult.

Only in Permic and Mari:

\**wärkä* 'kidney', \**wirtV-* 'to hold, raise'

Both loans are ambiguous (they could be from PII or PI). The modern distribution is probably accidental (loss from related branches). It is also possible that these words are parallel loanwords.

Only in Khanty and Mansi (regular cases):

POUg \**ēräy* 'song', POUg \**eträ* 'clear sky', POug \**mɔ̄ŋki* 'forest-spirit', \**nātV-* 'to help', \**päčäy* 'reindeer'

The number of these etymologies is so low that it is very difficult to determine whether these words were borrowed into Proto-Ob-Ugric or some earlier proto-language, such as Proto-Ugric. These words show regular phonological relations, but this can also be an illusion: Zhivlov (2006) reconstructs rather many stem types and correspondence types for Proto-Ob-Ugric, and some of these include only a few etymologies, which can point to the conclusion that the words showing these rare correspondences are not real cognates after all.

\**päčäy* has to be a Proto-Iranian loanword because of \**č* (← PI \**ts*), \**mɔ̄ŋki* has to be a late, Middle-Iranian-type loan, and \**nātV-* is ambiguous. The rest of the loans can be from Alanic.

Only in Khanty and/or Mansi and/or Hungarian (regular cases):

\**säptä* 'seven' (Khanty + Hungarian regular), \**sara* 'lake'

There are so few convincing loanwords with a "Ugric" distribution that they provide very little evidence. Either of these loans could be from Proto-Indo-Iranian or Proto-Iranian, if we assume that \**s* > \**h* was a common Iranian sound change. Both loans were acquired

before \*s > θ, \*ś > s, meaning that they were borrowed into a form of Uralic that was not yet Proto-Ugric.

Only in Mansi:

\**aćtirV* ‘whip’, PMs \**pēhtV* ‘brother-in-law’, PMs \**rāsnā* ‘rope’, \**širyV* ‘sword’, \**šārkVśV* ‘eagle’, PMs \**sāt* ‘seven’ (borrowed independently from Pre-Ug \**sāptā*), \**wēšəy* ‘calf’, \**āsVrVmV* ‘shame’

These words consist of different layers, with \**širyV* ‘sword’, \**šārkVśV* ‘eagle’ and \**āsVrVmV* ‘shame’ probably from Alanic. PMs \**rāsnā* ‘rope’ and \**wēšiy* have to be late loans, and \**aćtirV* has to be very early because of the palatal \*ć. The rest are ambiguous.

Only in Khanty:

PKh \**kārtV* ‘iron’, PKh \**layər* ‘Panzerhemd’, PKh \**pint* ‘way’, PKh \**wārəs* ‘horse’s mane’

PKh \**layər* is a probable Alanic loan which was clearly borrowed before \*s > \*<sub>l</sub>, but the rest are more or less ambiguous. Because of the small number of loanwords, nothing very far-reaching can be said about the contacts between Khanty and the Indo-Iranian languages.

Only in Hungarian (early loans):

\**arwV* (> Hu *ár* ‘price’), *ostor* ‘whip’ (< ? \**oćtVrV*), \**surV* ‘lord’, \**tājV* ‘milk’, \**tājinV* ‘cow’

As is the case with the other Ugric languages, the number of the early loanwords in Hungarian is quite low, which again prevents us from making far-reaching conclusions. Many of the loanwords are ambiguous, but *ár* most probably reflects an Iranian form \**ar̥a-*, which means that the word could have been borrowed from Alanic (or another form of early Eastern Iranian). If the *s* in *ostor* reflects earlier \*ć, like in Mansi \**aćtar*, the word has to be a very early loan. The examples of \*ć > s [š] in Hungarian are few, and because of this, the development remains hypothetical. On the other hand, it could also be borrowed from some much later form of Iranian with \**-št-*, although in this case the word should be quite late, borrowed at a time when *s* was already possible in Hungarian. Hungarian *s* can also reflect earlier \*č, and it could also be assumed that \**-čt-* was the substitution for Indo-Iranian \**-ćt-* or \**-št-*.

Only in Samoyed:

\**jāwi* (> PS \**jāē*), PS \**pulā* ~ \**pilā* ‘bridge’, \**tājki* ‘spear’, PS \**wār̥kā* ‘wolf’, Pre-S \**tāši* (> PS \**tāt*), \**wakša-* (> PS \**wātā*) ‘to grow’

Of these, only \**wātā-* has to be a very early loan because of \*s > \**t*. \**jāwi* (> PS \**jāē*) and PS \**wār̥kā* were possibly acquired before the Proto-Samoyed vowel developments, making them probably early loanwords too. Formally all of them could be either from PII or PI. \**pulā* ~ \**pilā* could have been borrowed into Proto-Samoyed (with Iranian \**u*

corresponding to Samoyed *\*u*), and because of the *\*l* the word is probably from a relatively late, Middle Iranian language.

The following loanwords have a distribution with a cognate in both Samoyed and some other branch:

*\*śaḍa-* ‘to rain’, *\*tora-* ‘to fight’ (also *\*itä-*, which is more uncertain, belongs here)

The following loanwords have a distribution with regular cognates with at least one Ugric branch and some other branch, which points to early borrowing. Although formally *\*kana-* and *\*kara-* are ambiguous, they are probably from Proto-Indo-Iranian because of their distribution. The rest of the loans are from Pre-II or PII.

*\*kana-* ‘to dig’, *\*kara-* ‘to dig’, *\*meti* ‘honey’, *\*mekši* ‘bee’, *\*orpV* ‘orphan’, *\*ora* ‘awl’, *\*peji-* ‘to milk’, *\*pätäri-* ‘to flee’, *\*śara-* ‘shit’, *\*śoma-* ‘sad’

The following loanwords are found in at least two non-adjacent branches of Uralic (the ones listed in the above categories are not counted). As there are no widely accepted criteria for a word to be considered “Uralic”, all of these could be considered loanwords into Proto-Uralic, in this case probably from Proto-Indo-Iranian or Pre-Indo-Iranian.

*\*ajša* ‘shaft’, *\*anti/onta* ‘grass’, *\*ertä* ‘side’, *\*kiñtaw* ‘tree stump’, *\*mertä* ‘human’, *\*orja* ‘slave’, *\*para* ‘good’, *\*počaw* ‘reindeer’, *\*puntaksi* ‘bottom’, *\*sayi-* ‘to get, obtain’, *\*repä* ‘fox’, *\*siŋka* ‘old’, *\*sasara* ‘sister’, *\*sejti* ‘bridge’, *\*šišta* ‘wax’, *\*tarna* ‘grass’, *\*toraksi* ‘through’, *\*wiša* ‘venom’

### 3.6. DISCUSSION ABOUT THE DISTRIBUTION AND ITS IMPACT ON URALIC TAXONOMY

Although much weight has been given to Indo-Iranian loanwords in the study of Uralic taxonomy, since Koivulehto’s works it has been obvious that the issue is very complicated and that the loanwords offer very little evidence for this. Hopefully this study has enforced this view. It is now obvious that the mere distribution of the loanwords say little of the age of the borrowings, as already Koivulehto (1999a: 208–209) noted.

However, some observations can be made. It thus seems that there are Proto-Iranian loanwords which were borrowed simultaneously into several early branches of Uralic, making it likely that Uralic had split into several branches by the time of these contacts. Also the fact that many of the Proto-Indo-Iranian loanwords either show a restricted distribution (such as West Uralic *\*wačara*, *\*worači*) or irregular correspondences (*\*asVra*, *\*śasra*, *\*šiŋta*) can point to the conclusion that Proto-Uralic was fragmenting by the time when contacts with Proto-Indo-Iranian took place.

The earlier, Pre-Indo-Iranian loanwords usually show a wider distribution and regular sound correspondences. Although the number of these earliest loans is quite small, based on their distribution and regular correspondences it can be assumed that the Pre-Indo-Iranian stage (after RUKI, *\*l* > *\*r* and the merger of velars and labiovelars but before the

merger of non-high vowels) was concurrent with Proto-Uralic, with the changes leading to Proto-Indo-Iranian happening after the dispersal of Proto-Uralic.

The distribution of loanwords reinforces the old idea that Samoyed is a lexical outlier, as only few convincing Indo-Iranian etymologies for Proto-Uralic words (*\*saδa-* ‘to rain’, *\*tora-* ‘to fight’) have a convincing reflex in Samoyed. However, the fact that such etymologies exist means rather that the situation is due to lexical loss in Samoyed, and that the earliest contact occurred before Samoyed split off from Proto-Uralic.

There are very few loanwords that have a Ugric distribution (being found in at least one Ob-Ugric branch and Hungarian), and likewise rather few in Ob-Ugric. The few loans that have a distribution confined to Ugric were borrowed before the change *\*s > \*θ* took place. This means that the Ugric distribution does not mean much from the point of view of chronology or taxonomy, as the words were borrowed into a language that was still identical to Proto-Uralic. Even some loans borrowed into Khanty and Mansi have to be so early. On the other hand, as the change *\*s > \*θ* probably was areal rather than a true innovation, and we do not know precisely when this change occurred, the “Pre-Ugric” loans do not have to be very early.

### 3.7. IMPACTS ON DATING AND THE LOCATION OF THE CONTACT ZONES

The area where the contacts took place was probably not far from the area where Proto-Uralic and its early dialects were spoken, that is, the Volga-Kama area west of the Urals. However, it is impossible to define the contact zones in more detail. Both the loans with “Finno-Permic” and “Ugric” distribution can be quite easily explained through contacts that occurred west of the Urals, although at this point the early Uralic dialects were probably spoken in a relatively wide region. The small number of the loanwords attested in Samoyed is certainly partly due to the greater loss of lexicon in Samoyed, but also the early migration of the Pre-Samoyed speakers to east of the Urals may be part of the reason.

Because of the very limited number of convincing etymologies found only in Finnic or Saami, it is probable that there were not (extensive) contacts with Pre-Finnic or Pre-Saami after the split of Proto-West Uralic.

The great number of loanwords of varying ages in Permic inevitably points to the conclusion that the pre-form of the Permic branch had been constantly spoken in an area that was adjacent to the Iranian languages. The different layers of loanwords in Permic clearly point to chronological differences in the donor languages, but it also seems that Permic was in contact with various forms of Iranian and not with different diachronic stages of the same language.

It is an interesting issue that several of the loanwords do not have a good Indo-European etymology, and they are probably loanwords (substrate words) in Indo-Iranian. Although already Lubotsky (2001b) observed that some of the Indo-Iranian substrate words have been borrowed into Uralic, he assumed that the words have to be from Iranian, as in his view these words entered Proto-Indo-Iranian in the Bactria-Margiana Archaeological Complex in Central/South Asia, which is too far to the south and east from the Uralic contact areas. However, this idea has to be reviewed on the basis of these

etymologies: at least some of these Uralic words have to be borrowed from Proto-Indo-Iranian because of phonological archaisms, and it is thus possible that these words were borrowed into Proto-Indo-Iranian already earlier and in more northern areas.

### **3.8. LEXICAL EVIDENCE OF CULTURAL INFLUENCE**

Although it is not the main purpose of this study, it is interesting to give a brief summary of the areas of culture reflected by the loanwords. This is central in that the different loanword layers reflect different areas of cultural vocabulary, and in this respect the issue is related to the topic of this dissertation, the stratigraphy of the Indo-Iranian loanwords. In some cases the information on cultural aspects could provide additional information in support of phonological or distributional arguments.

However, the limited number of loans does not give much possibilities to consider the cultural influence that speakers of Proto-Indo-Iranian and later forms of Iranian had over the speakers of Uralic. As the reconstruction of the vocabulary of proto-languages and the culture of their speakers is in general fraught with challenges due to vocabulary loss and semantic change (Mallory 1991: 111–113; Campbell 1999: 350; Ringe 2004: 1023–1024), the whole issue remains speculative, and the remarks made here have to be understood in that sense.

In general, the words that have been borrowed are typical cultural words, and the contacts between Indo-Iranian and Uralic seems to have been a typical contact situation in which a culturally less-advanced language group borrows various cultural terms from a more “advanced” group. The words in various loanword layers related to horse and cattle breeding show obvious cultural influence in the field of domesticated animals, and the borrowing of some names of grains points to agricultural influence from the Indo-Iranians on the speakers of Uralic.

As some very basic terms, such as words for ‘kidney’ and some kinship terms were borrowed too, it is probable that the contacts were quite intensive. The limited number of loans, which is probably partly due to vocabulary loss in Uralic during the subsequent millennia, makes a more detailed assessment of the intensity impossible.

### **3.9. DIRECTIONS FOR FURTHER STUDIES**

The results of this dissertation will certainly be challenged by future works, but this is only natural and desirable. As Joki (1973: 353) and Szemerényi (1977a: 132) noted, interim reports on loanwords studies are crucial, but they only remain relevant for so long. The progress being made in the study of both Uralic and Indo-Iranian historical phonology at the moment is immense, and it is certain that from the present material new results can be achieved in the future. Equally likely is that new etymologies will be presented.

The later loanwords from Indo-Iranian have been dealt with only cursorily (the clearly later loans into Permic are shown in Appendix I), and many details concerning the relative (and absolute?) chronology of these loans, as well as the characteristics of the donor

languages, remain open. In the same way, the later loanwords into Hungarian remain a topic where new results can be achieved.

In addition to the Indo-Iranian loanword layers in the Uralic languages, also many ideas concerning the other early loanwords (Northwest Indo-European, Balto-Slavic, Proto-Indo-European) are now in a period of transition. As was shown in this work, some of the assumed Indo-Iranian loanwords must have been borrowed instead from some other branch, and these words await a largescale investigation with the same methods used here, and with the help of the results of modern Uralic linguistics as well as Indo-European studies. Hopefully the results of this work will make the study of later Iranian loanwords and the other early Indo-European loans easier.

An almost completely new topic of research would be the possible lexical influence from the Uralic languages on Indo-Iranian. Although some of the suggested etymologies were discussed here (see the listing in Appendix II), certainly more could be said on this topic in the future. At least all the suggested etymologies could be discussed in more detail.

As stated in the Introduction, this study had the sole objective of examining the lexical influence from Indo-Iranian on Uralic. The location of the contact zones, the absolute chronologies and correlation with the results of archaeology were deliberately left out. If this work has achieved its purpose in presenting a convincing model of loanword stratigraphy, then archaeologists and researchers of other related fields can base their new research on these results. Also the matter of possible cultural influence (such as the alleged Indo-Iranian elements in the mythology of Uralic peoples) can, hopefully, be better approached in the future.

## BIBLIOGRAPHY

- Abaev, Vaso I. 1958–1989: *Историко-этимологический словарь осетинского языка*. Vols. I–V. Moscow & Leningrad: Akademiya Nauk SSSR.
- Abondolo, Daniel 1996: *Vowel rotation in Uralic: Obug[r]ocentric evidence*. SSEES Occasional papers no. 31.; London: School of Slavonic and East European Studies, University of London.
- 1998: Introduction. – Abondolo (ed.), 1 – 42.
- Abondolo, Daniel (ed.) 1998: *The Uralic languages*. Routledge Language family descriptions; London and New York: Routledge.
- Agyagási, Klára 2019: *Chuvash historical phonetics. An Areal Linguistic Study. With an Appendix on the Role of Proto-Mari in the History of Chuvash Vocalism*. Turcologica 117; Wiesbaden: Harrassowitz.
- Ahlqvist, Anders 1857: Venäläisiä sanoja suomen kielessä. – *Suomi. Tidskrift i fosterländska ämnen*. Helsingfors: Finska Litteratur-Sällskapets tryckeri.
- Aikio, Ante & Aslak Aikio 2001: Heimovaelluksista jatkuvuuteen. Suomalaisen väestöhistorian tutkimuksen pirstoutuminen. – *Muinaistutkija* 4/2001: 2–21.
- Aikio, Ante & Petri Kallio 2005: Zu Hartmut Katz: Studien zu den alteren indoiranischen Lehnwörtern in den uralischen Sprachen. Aus dem Nachlas herausgegeben von Paul Widmer, Anna Widmer und Gerson Klumpp. – *Die Sprache: Zeitschrift für Sprachwissenschaft* 45: 212–222.
- Aikio, Ante 2002: New and old Samoyed etymologies. – *Finnisch-Ugrische Forschungen* 57: 9–57.
- 2004: An Essay on Substrate Studies and the origin of the Saami. – Irma Hyvärinen, Petri Kallio und Jarmo Korhonen (Hrsg.), *Etymologie, Entlehnungen und Entwicklungen*. Mémoires de la Société Neophilologique de Helsinki 63; Helsinki: Société Neophilologique.
- 2006a: New and old Samoyed etymologies (Part II). – *Finnisch-Ugrische Forschungen* 59: 9–34.
- 2006b: On Germanic-Saami contacts and Saami prehistory. – *Journal de la Société Finno-Ougrienne* 91: 9–55.
- 2007: Etymological Nativization of Loanwords: a Case Study of Saami and Finnish. – Ida Toivonen & Diane Nelson (eds.), *Saami linguistics*. Amsterdam & Philadelphia: John Benjamins. 17–52.
- 2009: *Saami loanwords in Finnish and Karelian*. PhD dissertation. University of Oulu.
- 2012a: On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic \*x. – Tiina Hyytiäinen, Lotta Jalava, Janne Saarikivi & Erika Sandman (eds.), *Per Urales ad Orientem. Iter polyphonicum multilingue. Festschrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012*. Mémoires de la Société Finno-Ougrienne 264; Helsinki: Société Finno-Ougrienne. 227–250.
- 2012b: An essay on Saami linguistic prehistory. – Riho Grünthal & Petri Kallio (eds.), *A Linguistic map of Prehistoric Northern Europe*. Mémoires de la Société Finno-Ougrienne 266; Helsinki: Société Finno-Ougrienne.
- 2013a: Uralilaisen kantakielen vokaalistosta. Etymologia ja kielihistoria: Erkki Itkosen ja Aulis J. Joen 100-vuotisjuhlaseminaari 19.4.2013 (handout). Helsinki.
- 2013b: Studies in Uralic etymology I: Saami etymologies. – *Linguistica Uralica* XLIX: 161–174.
- 2014b: Studies in Uralic etymology III: Mari etymologies. – *Linguistica Uralica* L: 81–94.

- 2014c: On the reconstruction of Proto-Mari vocalism. – *Journal of language relationship* 11: 125–157.
- 2014d: The Uralic-Yukaghiric lexical correspondences: genetic inheritance, contact of chance resemblance? – *Finnisch-Ugrische Forschungen* 62: 7–76.
- 2015a: Studies in Uralic etymology IV: Ob-Ugric etymologies. – *Linguistica Uralica* LI: 1–20.
- 2015b: The Finnic ‘secondary *e*-stems’ and Proto-Uralic vocalism. – *Journal de la Société Finno-Ougrienne* 95: 26–66.
- 2018: Uralic Etymological Dictionary. (Näyte tulevan sanakirjan käsikirjoituksesta; Société Finno-Ougrienne 16.2.2018)
- (manuscript). *Uralic Etymological Dictionary*. Unpublished manuscript.
- AiWb = Bartholomae, Christian 1961: *Altiranisches Wörterbuch*. 2., unveränderte Auflage. Berlin: De Gruyter.
- Alatalo 2004 = *Sölkupisches Wörterbuch* aus Aufzeichnungen von Kai Donner, U. T. Sirelius und Jarmo Alatalo. Zusammengestellt und herausgegeben von Jarmo Alatalo. Lexica Societatis Finno-Ugricae XXVII; Helsinki: Société Finno-Ougrienne.
- Aleman, Agustí 2000: *Sources on the Alans: a critical compilation*. Handbook of Oriental Studies. Section 8 Uralic & Central Asia Studies; Leiden: Brill.
- AlEW = Wolfgang Hock, R. Fecht; A.H. Feulner; E. Hill; D. S. Wodtko 2015: *Altlitauisches Etymologisches Wörterbuch*. Studien zur historisch-vergleichenden Sprachwissenschaft 7; Hamburg: Baar.
- Algeo, John 1990a: Semantic change. – Edgar C. Polomé (ed.), *Research guide on language change*. Berlin – New York: Mouton de Gruyter. 399–408.
- 1990b: Borrowing. – Edgar C. Polomé (ed.), *Research guide on language change*. Berlin – New York: Mouton de Gruyter. 409–414.
- Álgu – sámegielaid etymologalaš diehtovuodđu. The Etymological database of the Saami languages. Institute for the Languages of Finland, November 2006 –. Constantly updated. <http://kaino.kotus.fi/algu/> [last accessed in November 2019]
- Anttila, Raimo 1989: *Historical and comparative linguistics*. Current Issues in Linguistic Theory 6; Amsterdam/Philadelphia: John Benjamins.
- 2000: *Greek and Indo-European Etymology in Action: Proto-Indo-European \*agʷ-*. Current Issues in Linguistic Theory 200; Amsterdam – Philadelphia: John Benjamins.
- 2006: Review of Katz (2003). – *Journal of Indo-European Studies* 34: 231–6.
- Bailey, Harold 1979: *Dictionary of Khotan Saka*. Cambridge: Cambridge University Press.
- Bakró-Nagy, Marianne 2013: Mit tudunk az ugor fonológiáról? – Agyagási Klára, Hegedűs Attila és É. Kiss Katalin (szerk.), *Nyelvelmélet és kontaktológia II*. Szerk. Piliscsaba: PPKE BTK.
- Bárcki, Géza 1941: *Magyar szófejtő szótár*. Budapest: Királyi Magyar Egyetemi Nyomda.
- Bartens, Raija 1999: *Mordvalaiskielten rakenne ja kehitys*. Mémoires de la Société Finno-Ougrienne 232; Helsinki: Société Finno-Ougrienne.
- Beekes, Robert Stephen Paul 1988: Laryngeal developments: a survey. – Alfred Bammesberger (hrsg.), *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*. Heidelberg: Winter. 59–105.
- 1997: Historical phonology of Iranian. – *Journal of Indo-European Studies* 25: 1–26.
- 2010: *Etymological dictionary of Greek*. I–II. Leiden Indo-European Etymological Dictionary Series, vol. 10; Leiden – Boston: Brill.
- Benveniste, Emil 1949: Noms d'animaux en indo-européen. – *Bulletin de la Société de linguistique de Paris* 45: 74–103.
- Bereczki, Gábor 1988: Geschichte der wolgafinnischen Sprachen. – Denis Sinor (ed.), *The Uralic languages: description, history and foreign influences*. Handbuch der Orientalistik 8: Handbook of Uralic studies 1; Leiden: E. J. Brill. 314–350.

- 1992: *Grundzüge des Tscheremissischen Sprachgeschichte II*. Studia Uralo-Altaica 34; Szeged: József Attila Tudományegyetem.
- 1994: *Grundzüge des Tscheremissischen Sprachgeschichte I*. Studia Uralo-Altaica 35; Szeged: József Attila Tudományegyetem.
- 2002: *Cseremisztörténeti alaktan*. Debrecen: Debrecen University Press.
- 2013: *Etymologisches Wörterbuch des Tscheremissischen (Mari)*. *Der einheimische Wortschatz*. Nach dem Tode des Verfassers herausgegeben von Klára Agyagási und Eberhart Winkler. Veröffentlichungen des Societas Uralo-Altaica 86; Wiesbaden: Harrassowitz Verlag.
- Bergsland, Knut 1963: Suomalais-skandinaavinen \*rebäs 'kettu'. – *Virittäjä* 67: 150–154.
- Bielmeier, Ronald 1977: *Historische Untersuchungen zum Erb- und Wortschatzanteil im ossetischen Grundwortschatz*. Europäische Hochschulschriften, Reihe 27, Asiatische und Afrikanische Studien; Peter Lang; Frankfurt am Main.
- 1989: Sarmatisch, Alanisch, Jassisch und Altossetisch. – Schmitt (ed.), 236–245.
- 1993: Das Alanische bei Tzetzes. – Wojciech Skalmowski and Alois van Tongerloo (eds.), *Medioiranica. Proceedings of the International Colloquium organized by the Katholieke Universiteit Leuven from the 21st tot he 23rd of May 1990*. Leuven: Peeters & Departement Orientalistiek. 1–28.
- 2001: Zwei Beiträge zum Jungavestischen 2001. – Heiner Eichner, Peter-Arnold Mumm, Oswald Panagl, Eberhard Winkler (hrsg.) unter Mitarbeit von Roland Hemmauer, Susanne Knopp und Velizar Sadovski, *Fremd und eigen: Untersuchungen zu Grammatik und Wortschatz des Uralischen und des Indogermanischen ; in memoriam Hartmut Katz*. Wien: Praesens.
- Bjorvand, Harald & Fredrik Otto Lindeman 2000: *Våre arveord. Etymologisk ordbok*. Oslo: Institutt for sammenlignende kulturforskning.
- Blažek, Václav & Irene Hegedűs 2012: On the position of Nuristani within Indo-European. – Roman Sukač & Ondřej Šefčík (eds.), *The Sound of Indo-European 2: papers on Indo-European phonetics, phonemics and morphophonemics*. Lincom Studies in Indo-European Linguistics 41; Munich: Lincom. 40–66.
- Blažek, Václav 1990: New Fenno-Ugric-Indo-Iranian lexical parallels. – В.В. Иванов, Т.М. Судник, Е.А. Хелимский, (ред.). *Uralica-Indogermanica II*. Институт славяноведения и балканистики. Москва. 40–45.
- 2003: Toward the Fenno-Ugric cultural lexicon of Indo-Iranian origin. – *Indogermanische Forschungen* 108: 92–99.
- 2005: Indo-Iranian elements in Fenno-Ugric mythological lexicon. – *Indogermanische Forschungen* 110: 162–185.
- 2016: On Indo-European 'barley'. – Benedicte Nielsen Whitehead, Thomas Olander, Birgit Anette Olsen, Bjarne Simmelkjaer Sandgaard Hansen (eds.), *Etymology and the European lexicon*. Proceedings of the 14th Fachtagung der Indogermanischen Gesellschaft, 17–22 September 2012, Copenhagen. Wiesbaden: Reichelt. 53–67.
- 2017a: Indo-European "bear". – *Historische Sprachforschung* 130: 148–192.
- 2017b: Indo-European "Gold" in Time and Space. – *Journal of Indo-European studies* 45: 267–311.
- 2018: Indo-European Dendronyms in the Perspective of External Comparison. – *Journal of Indo-European studies* 46: 1–45.
- Bozzone, Chiara 2016: The Origin of the Caland System and the Typology of Adjectives. – *Indo-European Linguistics* 4: 15–52.
- Burrow, Thomas 1955: *The Sanskrit language*. Delhi: Faber and Faber.
- 1976: Review of Joki (1973). – *Kratylos* 21: 59–63.
- Caland, Willem 1892: Beiträge zur Kenntnis des Avesta, no. 19. – *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 31: 266–268.

- 1893. Beiträge zur Kenntnis des Avesta, no. 26. – *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 32: 592.
- Campbell, Lyle & William J. Poser 2008: *Language Classification: History and Method*. Cambridge: Cambridge University Press.
- Campbell, Lyle 1998: Nostratic: a personal assessment. – Joseph C. Salmons and Brian D. Joseph (eds.), *Nostratic: sifting the evidence*. Current Issues in Linguistic Theory 142; Amsterdam & Philadelphia: John Benjamins.
- 1999: *Historical linguistics: an introduction*. First MIT Press edition. Cambridge, Massachusetts: The MIT Press.
- 2006: Long-range comparison: methodological disputes. – Keith Brown (ed.), *Encyclopedia of Language and Linguistics* (2nd edition). Oxford: Elsevier. 7.324–331.
- Cantera, Albert 2001: Die Behandlung der Lautfolge (C)rHC- im Iranischen. *Münchener Studien zur Sprachwissenschaft* 61:7–27.
- 2017: The phonology of Iranian. – Jared Klein & Brian D. Joseph with Matthias Fritz (eds.), *Handbook of comparative and historical Indo-European linguistics: an international handbook*. Berlin: De Gruyter Mouton. 481–503.
- Chantraine, Pierre 1933: *La Formation des noms en grec ancien*. Collection Linguistique publiée par la Société de Linguistique de Paris, XXXVIII; Paris: Champion.
- 2009: *Dictionnaire étymologique de la langue grecque: histoire de mots. Nouvelle édition*. Paris: Klincksieck.
- Cheung, Johnny 2002: *Studies in the historical development of the Ossetic vocalism*. Beiträge zur Iranistik, Band 21; Wiesbaden: Dr. Ludwig Reichert Verlag.
- 2007: *Etymological dictionary of the Iranian verb*. Leiden Indo-European Etymological Dictionary Series, vol. 2; Leiden – Boston: Brill.
- Cincius, Vera = В. И. Цинциус. 1975–1977: *Сравнительный словарь тунгусо-маньчжурских языков / Отв. ред. В. И. Цинциус. Т. I–II. Л.*
- Clackson, James 1994: *The Linguistic relationship between Armenian and Greek*. Oxford: Blackwell.
- 2007: *Indo-European linguistics. An introduction*. Cambridge Textbooks in Linguistics; Cambridge: Cambridge University Press.
- Čop, Bojan 1974: Indouralica II. – *Ural-Altische Jahrbücher* 44: 162–178.
- Cowgill, Warren & Manfred Mayrhofer 1986: *Indogermanische Grammatik*. Band I. Heidelberg: Winter.
- Cowgill, Warren 1965: Evidence in Greek. *Evidence for Laryngeals*. Edited by Werner Winter. London: Mouton. 142–180.
- Csúcs, Sándor & Zoltán Molnár 2009: Alternatív rekonstrukció az UEW-ben. – *Nyelvtudományi közlemények* 106: 7–71.
- Csúcs, Sándor 2005: *Die Rekonstruktion der permischen Grundsprache*. Budapest: Akadémiai Kiadó.
- Czentnár, András 2015: Egy finnugor etimológia török kapcsolatai. – *Keletkutatás 2015. tavasz*: 9–19.
- Darms, Georges 1978: *Schwäher und Schwager, Hahn und Huhn. Die Vrddhi-Ableitung im Germanischen*. München: Kitzinger
- DatSemShift = DatSemShift 2.0. Database of Semantic Shifts in languages of the world. <http://datsemshift.ru/> [last accessed 26 May 2019]
- Delamarre, Xavier 1992: Finnois *karhu*, aryen *hr̥kšas*, indoeuropéen *\*h<sub>2</sub>rtk<sup>h</sup>os* ‚ours‘. – *Historische Sprachforschung*. 151–154.
- Derksen, Rick 2008: *Etymological dictionary of the Slavic inherited lexicon*. Leiden Indo-European Etymological Dictionary Series, vol. 4; Leiden – Boston: Brill.
- 2015: *Etymological dictionary of the Baltic inherited lexicon*. Leiden Indo-European Etymological Dictionary Series, vol. 13; Leiden – Boston: Brill.

- DEWOS = Wolfgang Steinitz 1966–1993: *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. Berlin: Akademie Verlag.
- Diefenbach, Lorenz 1851: *Vergleichendes Wörterbuch der Gothischen Sprache II*. Frankfurt am Main: Verlag von Joseph Baer.
- Dolgopolsky, Aharon 1989: Cultural contacts of Proto-Indo-European and Proto-Indo-Iranian with neighbouring languages. – *Folia linguistica historica VIII/1–2*. 3–36.
- Dunkel, George 1995: More Mycenaean survivals in later Greek. – H. Hettrich, W. Hock, P.A. Mumm & N. Oettinger (eds.), *Verba et structurae. Festschrift für Klaus Strunk zum 65. Geburtstag*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. 1–21.
- 2014: *Lexikon der indogermanischen Partikeln und Pronominalstämme 1–2*. Heidelberg: Winter.
- Durkin-Meisterernst, Desmond 2004: *Dictionary of Manichean Middle Persian and Parthian*. Turnhout: Brepols.
- Durkin, Philip 2009: *The Oxford Guide to Etymology*. Oxford: Oxford University Press.
- Edel'man, D.I. = Д.И. Эдельман 1986: *Сравнительная грамматика восточноиранских языков. Фонология*. Москва: Наука.
- 1999: *Языки мира: Дардские и нуристанские языки*. Москва: Индрик.
- EES = *Eesti etimoloogiasõnaraamat*. 2012. Koostanud ja toimitanud Iris Metsmägi, Meeli Sedrik, Sven-Erik Soosaar. Tallinn: Eesti Keele Sihtasutus.
- EMS = *Eesti murtete sõnaraamat*. <https://www.eki.ee/dict/ems/> [Last accessed 24 May 2019]
- EIEC = *Encyclopedia of Indo-European culture*. Edited by J. P. Mallory and Douglas Q. Adams. 1997. London and Chigaco: Fitzroy Dearborn Publishers.
- Emmerick, Ronald 1992a: Old Indian. – Jadranka Gozdanovic (ed.), *Indo-European numerals*. Berlin: De Gruyter Mouton. 163–198.
- 1992b: Iranian. – Jadranka Gozdanovic (ed.), *Indo-European numerals*. Berlin: De Gruyter Mouton. 289–346.
- Erdosy, George 1995: Language, material culture and ethnicity: Theoretical perspectives. – George Erdosy (ed.), *The Indo-Aryans of Ancient South Asia: Language, Material Culture and Ethnicity*. Indian Philology and South Asian Studies 1; Berlin: de Gruyter. 1–31.
- EWAia = Mayrhofer, Manfred 1986–2001: *Etymologisches Wörterbuch des Altindiarischen I–III*. Heidelberg: Universitätsverlag C. Winter.
- EWUng = *Etymologisches Wörterbuch des Ungarischen Sprache*. 1992–1997. Hrsg. von Loránd Benkő. Budapest: Akadémiai Kiadó.
- Fortson, Benjamin W. 2010: *Indo-European language and culture: an introduction*. Second edition. Oxford: Blackwell publishing.
- Fox, Anthony 1995: *Linguistic reconstruction: An introduction to theory and method*. Oxford – New York: Oxford University Press.
- Fraenkel, Ernst 1962–65: *Litauisches etymologisches Wörterbuch*. Heidelberg: Winter.
- Frisk, Hjalmar 1954–1972: *Griechisches Etymologisches Wörterbuch*. Heidelberg: Winter.
- FUV = Björn Collinder 1955: *Fenno-Ugric vocabulary*. Stockholm: Almqvist & Wiksell.
- Gamkrelidze & Ivanov = Гамкрелидзе Т.В. & В.В. Иванов 1984: *Индоевропейский язык и индоевропейцы*. Тбилиси: Издательство Тбилисская университета
- Garnier, Romain, Laurent Sagart & Benoît Sagot 2017: Milk and the Indo-Europeans. – Martine Robbeets, Alexander Savalyev (eds.), *Language Dispersal Beyond Farming*. John Benjamins. 291–311.
- Garnier, Romain 2004–2005: Śánkate, kóghkos et védique śánkú-. – *Bulletin d'Études Indiennes* 22–23: 189–210.
- 2015: Secondary roots in PIE. Handout, 15. Fachtagung der Indogermanisches Gesellschaft. Vienna, September 2016.

- Gharib, B. 1995: *Sogdian Dictionary: Sogdian-Persian-English*. Tehran: Farhangan Publications.
- Gippert, Jost 1985: Rez. Symposium Saeculare Societatis Finno-Ugricae. – *Kratylos* 30: 157–62.
- 1997: Laryngeals and Vedic metre. – Alexander Lubotsky (ed.), *Sound Law and Analogy: Papers in Honor of Robert S.P. Beekes on the Occasion of His 60th Birthday*. Leiden Studies in Indo-European, Vol. 9; Amsterdam/ Atlanta, GA: Rodopi. 63–79.
- Gombocz, Zoltán 1920: A magyar mélyhangú *i* kérdéséhez. – *Magyar nyelv* XVI. 2–9.
- 1940: *Összegyűjtött művei* II. Budapest.
- Gordeev = Ф. И. Гордеев 1967: Балтийские и иранские заимствования в марийском языке. – Г. А. Архипов, Д. Е. Казанцев (ред.), *Происхождение марийского народа*. Материалы научной сессии, проведенной Марийск. НИ яз., литературы и истории. 23-25 дек. 1966 г. 25–29.
- Gosselain, Olivier P 2016: To hell with ethnoarchaeology! *Archaeological Dialogues* 23(02): 215–228.
- Gotō, Toshufumi 2013: *Old Indo-Aryan Morphology and its Indo-Iranian Background*. Veröffentlichungen zur Iranistik Nr. 60; Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Grünthal, Riho 2002: Prehistoric contacts between Mordvinic and Indo-European. – Rogier Blokland & Cornelius Hasselblatt (eds.), *Finno-Ugrians and Indo-Europeans: Linguistic and Literary Contacts*. Studia Fenno-Ugrica Groningana 2; Maastricht: Shaker Publishing. 84–91.
- 2012: Baltic loanwords in Mordvin. – Riho Grünthal & Petri Kallio (eds.), *A Linguistic map of Prehistoric Northern Europe*. Mémoires de la Société Finno-Ougrienne 266; Helsinki: Société Finno-Ougrienne. 297–343.
- Gulya, János 1977: Megjegyzések az ugor őshaza és az ugor nyelvek szétválása kérdéséről. – Antal Bartha, Károly Czeglédy & András Róna-Tas (szerk.), *Magyar őstörténeti tanulmányok*. Budapest: Akadémiai Kiadó. 115–121.
- 1992: Eine Bemerkung zur indogermanisch-finnougrischen Problematik. – *Ural-Altäische Jahrbücher*. Neue Folge. 11: 41–43.
- 2003: Széjjegyzetek az indoeurópai-finnugor kutatásokhoz. – Zoltán Molnár & Gábor Zaicz (eds.), *Permistica et Uralica. Ünnepi könyv Csúcs Sándor tiszteletére*. Fenno-Ugrica Pázmányienseia I. Piliscsaba: Pázmány Péter Katolikus Egyetem. 79–82.
- Hahmo, Sirkka-Liisa 1988: *Avoliitto 'avoin liitto'?* – Sirkka-Liisa Hahmo, Tette Hofstra, A.D. Kylstra, Osmo Nikkilä. *Omaa vai lainattua: itämerensuomen germaanisiin lainasanoihin liittyviä kirjoitelmia*. Helsinki: Suomalaisen Kirjallisuuden Seura. 67–83.
- 1994: *Grundlexem oder Ableitung? die finnischen Nomina der Typen kämmen und pähkinä und ihre Geschichte*. Studia Fennica 5; Helsinki: Suomalaisen Kirjallisuuden Seura.
- 1997: Finnische nominale Ableitungssuffixe fremder Herkunft. – Sirkka-Liisa Hahmo, Tette Hofstra, László Honti, Paul Van Linde & Osmo Nikkilä (hrsg.), *Finnisch-ugrische Sprachen in Kontakt. Vorträge des Symposiums aus Anlaß des 30-jährigen Bestehens der Finnougristik an der Rijksuniversiteit Groningen 21.-23. November 1996*. Maastricht: Shaker.
- Häkkinen, Jaakko 2007: *Kantauralin murteutuminen vokaalivastaavuuksien valossa*. Unpublished MA thesis, University of Helsinki. <https://helda.helsinki.fi/bitstream/handle/10138/19684/kantaura.pdf?sequence=2&isAllowed=y> [last accessed 24 May 2019]
- 2009: Kantauralin ajoitus ja paikannus: perustelut puntarissa. *Journal de la Société Finno-Ougrienne* 92: 9–56.
- 2012: After the protolanguage: Invisible convergence, false divergence and boundary shift. – *Finnisch-Ugrische Forschungen* 61: 7–28.

- 2019: Kantasuomen keskivokaaleista: paluu. – Santeri Junntila & Juha Kuokkala (toim.), *Petri Kallio rocks*. Helsinki. 24–40.
- Häkkinen, Kaisa & Terttu Lempiäinen 1996: Die ältesten Getreidepflanzen der Finnen und ihre Namen. – *Finnisch-Ugrische Forschungen* 53: 115–182.
- Häkkinen, Kaisa 1983: *Suomen kielen vanhimmasta sanastosta ja sen tutkimisesta*. Väitöskirja. Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja 17; Turku: Turun yliopisto.
- 1984: Wäre es schon an der Zeit, den Stammbaum zu fällen? – *Ural-Altäische Jahrbücher, Neue Folge* 4; 1–24.
- 1987: Vanhat lainat uusissa kehyksissä. [Arvostelu teoksesta Károly Rédei: Zu den indogermanisch-uralischen Sprachkontakten]. – *Sananjalka* 29. 241–247.
- 2004: *Nykysuomen etymologinen sanakirja*. Juva: WSOY.
- Harmatta, János 1955: Hozzászólások. – *A Magyar Tudományos Akadémia Nyelv- és Irodalomtudományi Osztályának Közleményei VII*: 249–254.
- 1977: Irániak és finnugorok, irániak és magyarok. – Antal Bartha, Károly Czeglédy & András Róna-Tas (szerk.), *Magyar őstörténeti tanulmányok*. Budapest: Akadémiai Kiadó. 167–182.
- 1992: The Emergence of the Indo-Iranians: the Indo-Iranian languages. – A. H. Dani & V. M. Masson (eds.), *History of Civilizations of Central Asia. Vol 1*. Paris: Unesco Publishing. 346–370.
- Haspelmeth, Martin & Uri Tadmor 2008 (eds.): *Loanwords in the world's languages: a comparative handbook*. Berlin: De Gruyter Mouton.
- Haugen, Einar 1950: The analysis of linguistic borrowing. – *Language* 26: 210–231.
- Hayyim, Sulayman 1934–1936: *New Persian-English dictionary*. Teheran: Librairie-imprimerie Béroukhim. <https://dsal.uchicago.edu/dictionaries/hayyim/> [last accessed 24 May 2019]
- Heggarty, Paul & Colin Renfrew 2014: South and Island Southeast Asia: Languages. – Colin Renfrew & Paul Bahn (eds.), *The Cambridge World Prehistory*. Cambridge: Cambridge University Press. 534–558.
- Heidermanns, Frank 2001: Lautgesetz versus Lautersatz. Zur Mehrfachsubstitution bei lexikalischer Interferenz. – Heiner Eichner, Peter-Arnold Mumm, Oswald Panagl, Eberhard Winkler (hrsg.) unter Mitarbeit von Roland Hemmauer, Susanne Knopp und Velizar Sadovski, *Fremd und eigen: Untersuchungen zu Grammatik und Wortschatz des Uralischen und des Indogermanischen; in memoriam Hartmut Katz*. Wien: Praesens. 39–53.
- Heikkilä, Mikko 2014: *Bidrag till Fennoskandiens språkliga förhistoria i tid och rum*. PhD dissertation. Helsinki.
- Helimski, Eugene 1982 = Хелимский, Е. А. 1982. *Древнейшие венгерско-самодийские языковые параллели*. Москва: Наука.
- 1984: Permi szófejtések. – *Nyelvtudományi közlemények* 86: 196–199.
- 1985: [Рецензия на] Honti, 1982. – *Советское финно-угроведение XXI*: 63–73.
- 1991: [Рецензия на] Lushnikova 1991. – *Linguistica Uralica XXVII*: 218–222.
- 1997a: *Die matorische Sprache*. *Studia Uralo-Altica* 41; Szeged: JATE Finnugor Tanszék.
- 1997b: The southern neighbours of Finno-Ugrians: Iranians or an extinct branch of Aryans („Andronovo Aryans“)? – Sirkka-Liisa Hahmo, Tette Hofstra, László Honti, Paul Van Linde & Osmo Nikkilä (hrsg.), *Finnisch-ugrische Sprachen in Kontakt. Vorträge des Symposiums aus Anlaß des 30-jährigen Bestehens der Finnougristik an der Rijksuniversiteit Groningen 21.–23. November 1996*. Maastricht: Shaker. 117–25.
- 1998: Nganasan. – Abondolo (ed.), 480–515.

- 1999: Basic Vocabulary in PU and PFU – Remarks on Etymology and Reconstruction. A handout delivered at the Budapest Uralic Workshop 2: Etymology, September 1999, organized by the Nyelvtudományi Intézet, Budapest.
- 2000: *Компаративистика, уралистика: Лекции и статьи*. Studia Philologica. Москва: Языки русской культуры.
- 2001: Early Indo-Uralic linguistic relationships: real kinship and imagined contacts. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: linguistic and archaeological considerations*. Mémoires de la Société Finno-Ougrienne 242; Helsinki: Société Finno-Ougrienne. 187–206.
- 2003: Areal groupings (Sprachbunde) within and across the borders of the Uralic language family: a survey. – *Nyelvtudományi Közlemények 100*: 156–167.
- Hintze, Almut 1994: *Der Zamyād-Yašt: Edition, Übersetzung, Kommentar*. Beiträge zur Iranistik 15; Wiesbaden: Dr. Ludwig Reichert Verlag.
- 1998: The Migrations of the Indo-Iranians and the Iranian Sound-Change s > h. – Wolfgang Meid (hrsg.), *Sprache und Kultur der Indogermanen*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. 139–156.
- 2000: "Lohn" im indoiranischen. Eine semantische Studie des Rigveda und Avesta. Beiträge zur Iranistik 20; Wiesbaden: Reichert.
- Hock, Hans Henrich & Brian Joseph 2009: *Language History, Language Change, and Language Relationship*. 2<sup>nd</sup> edition. Berlin – New York: Mouton de Gruyter.
- Hock, Hans Henrich 2016: The languages, their histories and genetic classification. – Hans Henrich Hock & Elena Bashir (eds.), *Languages and linguistics of South Asia*. De Gruyter Mouton. 9–240.
- Höfler, Stefan 2017: Hittite *hišša*- c. thill, shaft (of a cart)'and the feminine gender in Proto-Indo-European. Conference Handout: The Split – Reconstructing Early Indo-European Language and Culture. University of Copenhagen, September 13–15, 2017. [https://www.academia.edu/34569864/Hittite\\_hišša-c\\_thill\\_shaft\\_of\\_a\\_cart\\_and\\_the\\_feminine\\_gender\\_in\\_Prot Indo-European](https://www.academia.edu/34569864/Hittite_hišša-c_thill_shaft_of_a_cart_and_the_feminine_gender_in_Prot Indo-European) [accessed May 2019]
- 2018: The Caprice of O. On a Proto-Indo-European substantivization type and its excesses in Ancient Greek. Presentation, ECIEC 37, June 15–17, 2018, University of Michigan, Ann Arbor.
- Hofstra, Tette 1985: *Ostseefinnisch und Germanisch. Frühe Lehnbeziehungen im nördlichen Ostseeraum im Lichte der Forschung seit 1961*. Groningen.
- 1992: Germanisch-ostseefinnische Lehnbeziehungen: Notizen zu Adjektiven. – László Honti et al. (eds.), *Finnisch-ugrische Sprachen zwischen dem germanischen und dem slavischen Sprachraum*. Amsterdam – Atlanta, GA: Rodopi. 55–63.
- Holopainen, Sampsa & Santeri Junntila & Juha Kuokkala 2017: Indoeurópai jövevényszavak és a második szótagi labiális magánhangzók fejlődése az uráli nyelvekben. – Forgács, T., Németh, M. & Sinkovics, B (szerj.), *Nyelvtudományi kutatások újabb eredményei*. Szeged: SZTE. 109–136.
- Holopainen, Sampsa 2016: On the etymology of Hungarian *úr* 'lord, gentleman'. – *Folia Uralica Debreceniensia* 23: 57–68.
- 2018a: Substitutions of Indo-European palatal stops in Uralic? Handout of a presentation held at the 4th Indo-European research colloquium, 5.–6. April 2018, Zürich. [https://www.academia.edu/37664846/Substitutions\\_of\\_Indo-European\\_palatovelars\\_in\\_Uralic](https://www.academia.edu/37664846/Substitutions_of_Indo-European_palatovelars_in_Uralic) [accessed 26 May, 2019]
- 2018b: Indo-Iranian loans confined to Saami? – Sampsa Holopainen & Janne Saarikivi (eds.), *Peri orthótētos etýmōn. Uusiutuva uralilainen etymologia*. Uralica Helsingiensia 11; Helsinki; Société Finno-Ougrienne. 135–179.
- 2019: *Itää*-verbin mahdollisesta lainaetymologiasta ja etäsukukielistä vastineista. – Santeri Junntila & Juha Kuokkala (toim.), *Petri Kallio rocks*. Helsinki. 73–78.

- Honti, László 1982: *Geschichte des obugrischen Vokalismus der ersten Silbe*. Budapest: Akadémiai Kiado.
- 1997: *Az ugor alapnyelv kérdéséhez*. Budapesti finnugor füzetek 7; Budapest: ELTE; Numi-Tórem Alapítvány.
- 1999: *Az obiugor konszonantizmus története*. *Studia uralo-altaica supplementa* 9; Szeged: Szegedi Tudományegyetem.
- 2017: *A magyar és a nyugati ótörök szókészletli kapcsolatairól*. Budapest: Tinta Könyvkiadó.
- Hyllested, Adam 2009: Internal Reconstruction vs. External Comparison: The Case of the Indo-Uralic Laryngeals. – Jens Elmegård Rasmussen and Thomas Olander (eds.), *Internal reconstruction in Indo-European: Methods, results, and problems*. *Copenhagen studies in Indo-European* 3; Copenhagen: Museum Tusulanum Press. 3–28.
- 2014: *Language contacts at the gates of Europe*. Doctoral dissertation, University of Copenhagen.
- IEW = Pokorny, Julius 1959–1969: *Indogermanisches Etymologisches Wörterbuch*. Bern: Francke.
- Illich-Svitych, Vladislav Markovich 1968: *Sootvetstviya smychnykh v nostraticheskikh jazykakh*. Moskva: Nauka.
- Isebaert, Lambert 1980: *De Indo-Iraanse bestanddelen in de Tocharische woordenschat*. PhD dissertation. Leuven.
- 1991: Quelques considérations sur les emprunts iraniens en tokhariens. – Lambert Isebaert (ed.), *Studia Etymologica Indoeuropaea. Memoriae A.J. van Windekens dicte*. Leuven: Peeters & Departement Oriëntalistiek. 141–150.
- Itkonen, Erkki 1946: Zur Frage nach der Entwicklung des Vokalismus der ersten Silbe in den finnisch-ugrischen Sprachen, insbesondere im Mordwinischen. – *Finnisch-Ugrische Forschungen* 29:
- 1954: Zur Geschichte des Vokalismus der ersten Silbe im Tscheremissischen und in den permischen Sprachen. – *Finnisch-Ugrische Forschungen* 31: 149–345.
- 1956: Etymologisches Wörterbuch der uralischen Sprachen [Review of FUV]. – *Ural-Altäische Jahrbücher* 28: 56–80.
- 1973: Zur Geschichte des Partitivs. – *Finnisch-Ugrische Forschungen* 40: 278–339.
- Itkonen, Terho 1983: Välikatsaus suomen kielen juuriin. – *Virittäjä* 87: 190–229, 349–386.
- 1987: Erään vokaalivyyhden selvittelyä. – *Virittäjä* 91: 164–208.
- Itkonen, T. I. 1928: Lappalais-suomalaisia sanavertailuja. II. – *Juhlakirja Yrjö Wichmannin kuusikymmenvuotispäiväksi*. Mémoires de la Société Finno-Ougrienne 58; Helsinki: Société Finno-Ougrienne. 26–56.
- Jacobsohn, Hermann 1922: *Arier und Ugrofinnen*. Göttingen: Vandenhoeck & Ruprecht.
- 1933: Zu den ältesten arischen Lehnwörtern in den finnisch-ugrischen Sprachen. – *Liber semisaecularis Societatis Finno-Ugricae*. Mémoires de la Société Finno-Ougrienne LXVII; Helsinki: Société Finno-Ougrienne. 136–147.
- Jamison, Stephanie W & Joel B. Brereton 2014: *The Rigveda. The earliest religious poetry of India*. Translated by Stephanie W. Jamison and Joel B. Brereton. South Asia Research. Oxford – New York: Oxford University Press.
- Janda, Michael 2000: *Eleusis: das indogermanische Erbe der Mysterien*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Janhunen, Juha 1981: Uralilaisen kantakielen sanastosta. – *Journal de la Société Finno-Ougrienne* 77: 219–274.
- 1982: On the structure of Proto-Uralic. – *Finnisch-Ugrische Forschungen* 44: 23–42.

- 1983: On early Samoyed Indo-European contacts. – Juha Janhunen, Anneli Peräniitty & Seppo Suhonen (toim.), *Symposium Saeculare Societatis Fenno-Ugricae*. Mémoires de la Société Finno-Ougrienne 185; Helsinki: Société Finno-Ougrienne. 115–128.
- 1998: Samoyedic. – Abondolo (ed.), 457–479.
- 2009: Proto-Uralic: what, where and when? – Jussi Ylikoski (ed.), *The quasiquincentennial of the Finno-Ugrian Society*. Mémoires de la Société Finno-Ougrienne 258; Helsinki: Société Finno-Ougrienne 57–78.
- Janurik, Tamás 2012: Volt-e déli-szamojed (PSS) alapnyelv? – Tiina Hyytiäinen, Lotta Jalava, Janne Saarikivi & Erika Sandman (eds.), *Per Urales ad Orientem. Iter polyphonicum multilingue. Festschrift tillägnad Juha Janhunen på hans sextioårsdag den 12 februari 2012*. Mémoires de la Société Finno-Ougrienne 264; Helsinki: Société Finno-Ougrienne. 145–162.
- Joki, Aulis J. 1952: *Die Lehnwörter des Sajansamojedischen*. Mémoires de la Société Finno-Ougrienne 103; Helsinki: Société Finno-Ougrienne.
- 1959: Paleolingvistiikkamme ongelmia. – Pertti Virtaranta, Terho Itkonen & Paavo Pulkkinen (toim.), *Verba docent. Juhlakirja Lauri Hakulisen 60-vuotispäiväksi*. Helsinki: Suomalaisen Kirjallisuuden Seura. 48–67.
- 1973: *Uralier und Indogermaner*. Mémoires de la Société Finno-Ougrienne 151; Helsinki: Société Finno-Ougrienne.
- 1988: Zur Geschichte der uralischen Sprachgemeinschaft unter besonderer Berücksichtigung des ostseefinnischen. – Sinor (ed.), 575–595.
- Jokl, Norbert 1921: Das Finnisch-Ugrische als Erkenntnisquelle für die ältere indisch-germanische Sprachgeschichte. – *Prace lingwistyczne ofiarowane Janowi Badouinowi de Courtenay dla uczczenia jego działalności naukowej*. Kraków: Uniwersytet Jagielloński. 97–112.
- Junttila, Santeri 2012: The prehistoric context of the oldest contacts between Baltic and Finnic languages. – Riho Grünthal & Petri Kallio (eds.), *A Linguistic map of Prehistoric Northern Europe*. Mémoires de la Société Finno-Ougrienne 266; Helsinki: Société Finno-Ougrienne. 261–296.
- 2015: *Tiedon kumuloituminen ja trendit lainasanatutkimuksessa. Kantasuomen balttilaislainojen tutkimushistoria*. Doctoral dissertation; Helsinki.
- 2016: Die baltisch-slawische Frage im Lichte der alten Baltischen Lehnwörter des ostseefinnischen. – *Baltistica* LI (2): 217–238.
- 2018: Altlettgallische Lehnwörter in den mordwinischen Sprachen? – *Finnisch-Ugrische Forschungen* 64: 72–91.
- Junttila, Santeri (manuscript): Huomioita itämerensuomen herV-sanoista.
- Jürine, Anni, Karl Pajusalu, Renate Pajusalu, Ilona Tragel & Ann Veismann 2016: Estonian *jumal* 'god'. – *Folklore. Electronic Journal of Folklore* 64: 9–32.
- Kalima, Jalo 1933: Fi. *sammás* 'grenzstein'. – *Finnisch-Ugrische Forschungen* 31: 128–137.
- 1936a: *Itämerensuomalaisten kielten balttilaiset lainasanat*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- 1936b: Über die indoiranischen und baltischen Lehnwörter der ostseefinnischen Sprachen. – Helmut Arntz (hrsg.), *Germanen und Indogermanen: Volkstum, Sprache, Heimat, Kultur: Festschrift für Hermann Hirt*. 1–2. Indogermanische Bibliothek, 3 Abt, 15: 1–2 Bd; Heidelberg.: Winter. 199–214.
- 1950: Suomen *taivas* ja sen liettualainen vastine. – *Virittäjä* 54: 328–9.
- Kallio, Petri 2000: Germanic maggot and moth. – Dirk Boutkan & Arend Quak (eds.), *Language Contact: Substratum, Superstratum, Adstratum in Germanic Languages*. Amsterdamer Beiträge zur älteren Germanistik 54; Amsterdam – Atlanta, GA: Rodopi. 117–122.
- 2002: Prehistoric Contacts between Indo-European and Uralic. – Karlene Jones-Bley, Martin E. Huld, Angela Della Volpe & Miriam Robbins Dexter (eds.), *Proceedings of the*

- Thirteenth Annual UCLA Indo-European Conference*. Journal of Indo-European Studies Monograph Series 44; Washington, DC.: Institute for the Study of Man. 29–44 .
- 2004: Tocharian loanwords in Samoyed? – Irma Hyvärinen, Petri Kallio und Jarmo Korhonen (hrsg.), *Etymologie, Entlehnungen und Entwicklungen: Festschrift für Jorma Koivulehto zum 70. Geburtstag*. Mémoires de la Société Neophilologique de Helsinki 63; Helsinki: Société Neophilologique. 129–137.
- 2006: Suomen kantakielten absoluuttista kronologiaa. – *Virittäjä* 110: 2–25.
- 2007: Kantasuomen konsonanttihistoriaa. – Jussi Ylikoski ja Ante Aikio (doiammahan), *Sámit, sánit, sátnehámit. Riepmočála Pekka Sammallahtii miessemánu 21. beaivve 2007*. Mémoires de la Société Finno-Ougrienne 253; Helsinki: Société Finno-Ougrienne. 229–249.
- 2008: On the "Early Baltic" Loanwords in Common Finnic. – Alexander Lubotsky, Jos Schaeken & Jeroen Wiedenhof (eds.), *Evidence and counter-evidence: Essays in honour of Frederik Kortlandt 1, Balto-Slavic and Indo-European linguistics*. Studies in Slavic and General Linguistics 32; Amsterdam – New York: Rodopi. 265–277.
- 2009: Stratigraphy of Indo-European loanwords in Saami. – Tiina Äikäs (ed.), *Mattut – maddagat: The Roots of Saami Ethnicities, Societies and Spaces / Places*. Publications of the Giellagas Institute. Oulu: Giellagas Institute at the University of Oulu. 30–45.
- 2012a: The non-initial-syllable vowel reductions from Proto-Uralic to Proto-Finnic. – Tiina Hyytiäinen, Lotta Jalava, Janne Saarikivi & Erika Sandman (eds.), *Per Urales ad Orientem. Iter polyphonicum multilingue. Festschrift tillägnad Juha Janhunen pa hans sextioarsdag den 12 februari 2012*. Mémoires de la Société Finno-Ougrienne 264; Helsinki: Société Finno-Ougrienne. 163–175.
- 2012b: The prehistoric Germanic loanword strata in Finnic. – Riho Grünthal & Petri Kallio (eds.), *A linguistic map of prehistoric Northern Europe*. Mémoires de la Société Finno-Ougrienne 266; Helsinki: Société Finno-Ougrienne. 225–238.
- 2013: Review of LÄGLOS. – *Kratylos* 58: 214–218.
- 2014: Diversification of Proto-Finnic. – Joonas Ahola, Frog, Clive Tolley (eds.), *Fibula, Fabula, Fact – the Viking Age in Finland*. Studia Fennica Historica 18; Helsinki: Suomalaisen Kirjallisuuden Seura. 155–168.
- 2015a: Nugae Indo-Uralicae. – *Journal of Indo-European Studies* 43: 368–375.
- 2015b: The language contact situation in prehistoric Northeastern Europe. – Robert Mailhammer, Theo Vennemann gen. Nierfeld & Birgit Anette Olsen (eds.), *The Linguistic Roots of Europe: Origin and Development of European Languages*. Copenhagen Studies in Indo-European 6; Copenhagen: Museum Tusulanum Press. 77–102.
- 2016: Historical phonology from Proto-Finnic to Proto-Livonian. – *Journal of Estonian and Finno-Ugric Linguistics* vol. 7, no 1: 39–65.
- 2018a: Ensitavun diftongit kantasuomessa. – Sampsa Holopainen & Janne Saarikivi (toim.), *Peri orthótētos etýmōn. Uusiutuva uralilainen etymologia*. Uralica Helsingiensia 11; Helsinki: Société Finno-Ougrienne. 251–268.
- 2018b: Vanhan kirjaeteläviron äännehistoriaa. – Jüvä Sullöv (toim.), *Valitsõmisjaotusõst keeleaoluuni*. Võro Instituudi toimõndusõq 33; Võro: Võro Instituut. 118–144.
- (manuscript) The Proto-Balto-Slavic Loanword Stratum in the Uralic Languages.
- Kannisto et al. 2013 = *Wogulisches Wörterbuch*. 2013. Gesammelt und geordnet von Artturi Kannisto. Bearbeitet von Vuokko Eiras. Herausgegeben von Arto Moisiso. Lexica Societatis Finno-Ugricae XXXV; Helsinki: Société Finno-Ougrienne.
- Kapović, Mate 2017: Proto-Indo-European phonology. – Mate Kapović (ed.), *The Indo-European Languages*. Second edition. Routledge Language Family Series; London and New York: Routledge. 13–60.

- Karsten, Tor Evert 1943/44: *Finnar och germaner. Folk- och kulturberöringar från tre årtusenden*. Folkmålsstudier 8/9. Helsingfors.
- Kassian, Aleksei & Mikhail Zhivlov & George Starostin 2015: Proto-Indo-European-Uralic Comparison from the Probabilistic Point of View. – *Journal of Indo-European Studies* 43. 301–347.
- Katz, Hartmut 1971: Ung. *ér, érik*. – *Nyelvtudományi Közlemények* 73: 143–146.
- 1978: *ἄρθος*. – *Études finno-ougriennes* 15: 179–188.
- 1981: Idg. *\*b<sup>h</sup>ágos* „Gott“? – *Freundliche Grüße zur Halbzeit*. Festschrift für Johannes Beckert. Bremen. 27–29.
- 1983a: Hehtitisch *hišša* – und zubehör. – Gabriella Frantz-Szabo (hrsg.), *v Festschrift Annelies Kammenhuber*. *Orientalia* 52/1; Rome: Pontificium Institutum Biblicum. 116–122.
- 1983b: Zu idg. *\*mr̥tó-*. – *Sprache* 29. 174–177.
- 1985: *Studien zu den älteren indoiranischen Lehnwörtern in den uralischen Sprachen*. Habilitationsschrift. (manuscript.)
- 1987: Rez. Rédei 1986. – *Ural-altaische Jahrbücher* 7: 252–63.
- 2003: *Studien zu den älteren indoiranischen Lehnwörtern in den uralischen Sprachen*. Aus dem Nachlass herausgegeben von Paul Widmer, Anna Widmer und Gerson Klumpp. Indogermanische Bibliothek. 3. Reihe; Heidelberg: C. Winter.
- Katzschmann, Michael 2005: Überlegungen zu einem proto-uralischen (Laryngal-)Phonem *\*x*. – Heiner Eichner, Peter-Arnold Mumm, Oswald Panagl, Eberhard Winkler (hrsg.) unter Mitarbeit von Roland Hemmauer, Susanne Knopp und Velizar Sadovski. *Fremd und eigen. Untersuchungen zu Grammatik und Wortschatz des Uralischen und Indogermanischen*. Wien: Edition Praesens. 93–116.
- Keresztes, László 1986: *Geschichte des mordvinischen Konsonantismus II. Etymologisches Belegmaterial*. *Studia Uralo-Altica* 26; Szeged: JATE finnugor tanszék.
- 1997: Questions of the linguistic relations between Lapps (Saamis) and their neighbours. – Sirikka-Liisa Hahmo, Tette Hofstra, László Honti, Paul Van Linde & Osmo Nikkilä (hrsg.), *Finnisch-ugrische Sprachen in Kontakt. Vorträge des Symposiums aus Anlaß des 30-jährigen Bestehens der Finnougristik an der Rijksuniversiteit Groningen 21.-23. November 1996*. Maastricht: Shaker. 145–150.
- 1998: Samisk påvirkning på nabospråkene. – László Keresztes & Tor Tveite (red.), *Finsk-ugriske dager 1995 : foredragssamling om finsk-ugriske emner*. Oslo: Universitetet i Oslo. 91–102.
- 2011: *Bevezetés a mordvin nyelvészethez*. Debrecen: Debrecen University Press.
- 2017: *Основы мордовского языкознания*. Debrecen: Debrecen University Press.
- KESK = Lytkin, V. I. & E. S. Gulyaev 1999: *Краткий этимологический словарь коми языка*. Syktyvkar: Коми книжное издательство.
- KEWA = Mayrhofer, Manfred 1956–1980: *Kurzgefasstes etymologisches Wörterbuch des Altindischen*. Heidelberg: Carl Winter.
- Kim, Ronald 2018: Slavic Iranian contacts. Linguistic relations. – *Encyclopædia Iranica*. [forthcoming [https://www.academia.edu/37845208/Slavic-Iranian\\_contacts\\_Linguistic\\_relations](https://www.academia.edu/37845208/Slavic-Iranian_contacts_Linguistic_relations)] [accessed 24 May 2019]
- Kiparsky, Valentin 1952: Terve ja terva. – *Virittäjä* 56: 94–95.
- Kloekhorst, Alwin 2008a: *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden Etymological Dictionary Series, Vol 5. Leiden – Boston: Brill.
- 2008b: Some Indo-Uralic aspects of Hittite. – *Journal of Indo-European Studies* 36: 88–95.
- Kluge, Friedrich 1975: *Etymologisches Wörterbuch des deutschen Sprache*. Bearbeitet von Walter Mitzka. 21., unveränderte Auflage. Berlin – New York: de Gruyter.

- 2012: *Etymologisches Wörterbuch des deutschen Sprache*. Bearbeitet von Elmar Seebold. 25. durchgesehene und erweiterte Auflage. Berlin/New York: de Gruyter.
- Kobayashi, Masato 2004: *Historical phonology of Old Indo-Aryan consonants*. Study of Languages and Cultures of Asia and Africa Monograph Series No. 42; Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Kogan, Anton 2005: *Дардские языки. Генетическая характеристика*. Москва: Восточная литература РАН.
- Koivulehto, Jorma 1976: Vanhimmista germaanista lainakosketuksista ja niiden ikäämisestä. – *Virittäjä* 80: 33–47, 247–290.
- 1979: Phonotaktik als Wegweiser in der Lehnwortforschung: die osfi. -str-Wörter. – *Finnisch-Ugrische Forschungen* 43. 67–79.
- 1983: Seit wann leben die Finnen im Ostseeraum? Zur relativen und absoluten Chronologie der alten idg. Lehnwortschichten im Ostseefinnischen. – Juha Janhunen, Anneli Peräniitty & Seppo Suhonen (toim.), *Symposium saeculare Societatis Fenno-Ugricae*. Mémoires de la Société Finno-Ougrienne 185; Helsinki: Société Finno-Ougrienne. 135–157.
- 1988a: Lapin ja itämerensuomen suhteesta. Ieur. -Tr-yhtymän korvautuminen lainoissa. – *Virittäjä* 92: 26–48.
- 1988b: Idg. Laryngale und die finnisch-ugrische Evidenz. – Alfred Bammesberger (hrsg.), *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*. Heidelberg: Winter. 281–297.
- 1991: *Uralische Evidenz für die Laryngaltheorie*, Veröffentlichungen der Kommission für Linguistik und Kommunikationsforschung, Heft 24; Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- 1999a: Varhaiset indoeurooppalaiskontaktit: aika ja paikka lainasanojen valossa. – Paul Fogelberg (toim.), *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*. Bidrag till kannedom av Finlands natur och folk 153; Helsinki: Suomen Tiedeseura. 207–236.
- 1999b: *Verba Mutuata. Quae vestigia antiquissimi cum Germanis aliisque Indo-Europaeis contactus in linguis Fennicis reliquerint*. Editio Klaas Ph. Ruppel. Mémoires de la Société Finno-Ougrienne 237; Helsinki: Société Finno-Ougrienne. 1–16.
- 2000a: Fest und Zyklus des Jahres: *Jul* und *kekri*. – *Neuphilologische Mitteilungen* 101: 235–252.
- 2000b: Reflexe des urbaltischen \*ā in baltischen Lehnwörtern des Ostseefinnischen. – *Linguistica Baltica* 8: 103–124.
- 2001a: Zum frühen iranischen und indoiranischen lexikalischen Einfluss auf das Finnisch-Ugrische. – Klaus Karttunen & Petteri Koskikallio (eds.), *Vidyārnavaṇḍanam: essays in honour of Asko Parpola*. Studia Orientalia 94; Helsinki: Finnish Oriental Society. 2001. 359–378.
- 2001b: The earliest contacts between Indo-European and Uralic speakers in the light of lexical loans. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: linguistic and archaeological considerations*. Mémoires de la Société Finno-Ougrienne 242; Helsinki: Société Finno-Ougrienne. 235–264.
- 2003: *Arjalaisia lainoja*. Handout of a presentation held at the meeting of the Finno-Ugrian Society 21.3.2003
- 2005: Noin jalot vietettiin hevonsuistajan Hektorin peijaat: *peijaiset* ja *peikko* -sanojen alkuperästä. – *Virittäjä* 109: 322–333.
- 2006: Wie alt sind die Kontakte zwischen Finnisch-Ugrisch und Balto-Slavisch? – Juhani Nuorluoto (ed.), *The Slavization of the Russian North: Mechanisms and Chronology*. Slavica Helsingiensia 27; Helsinki. 154–166.
- 2007: Saamen ja suomen 'poro'. – Jussi Ylikoski ja Ante Aikio (doaimmahan), *Sámit, sánit, sátnehámit. Riepmočála Pekka Sammallahtii miessemánu 21. beaivve 2007*.

- Mémoires de la Société Finno-Ougrienne 253; Helsinki: Société Finno-Ougrienne. 251–258.
- 2008: Frühe slavisch-finnische Kontakte. – Alexander Lubotsky, Jos Schaeken & Jeroen Wiedenhof (eds.), *Evidence and counter-evidence: essays in honour of Frederik Kortland. Vol. 1. Balto-Slavic and Indo-European linguistics*. Amsterdam: Rodopi. 309–321.
- 2009: Etymologisesti hämäriä -(is)tA-johdosverbejä, lainoja ja omapohjaisia. – *Journal de la Société Finno-Ougrienne* 92: 79–102.
- 2016: *Verba vagantur. Jorma Koivulehto in memoriam*. Red. Sampsa Holopainen, Petri Kallio & Janne Saarikivi. Mémoires de la Société Finno-Ougrienne 274; Helsinki: Société Finno-Ougrienne.
- manuscript. *Indoiranische Lehnwörter. Arjalaisia lainoja. Monografian luonnostelua*.
- Kochetov, Alexei & Alevtina Lobatova 2007: Komi-Permyak coronal obstruents: acoustic contrasts and positional variation. – *Journal of the International Phonetic Association* Vol. 37, No. 1: 51–82.
- Kölligan, Daniel 2018: The lexicon of Proto-Indo-European. – Jared Klein & Brian D. Joseph with Matthias Fritz (eds.), *Handbook of comparative and historical Indo-European linguistics : an international handbook*. Vol 3. Berlin: De Gruyter Mouton. 2229–2279.
- Koponen, Eino 1999: Itämerensuomalaisten kielten meteorologisen sanaston etymologiaa. – *Virittäjä* 103: 506–530.
- Koptjevskaja-Tamm, Maria 2008: Approaching lexical typology. – Martine Vanhove (ed.), *From Polysemy to Semantic Change. Towards a typology of lexical semantic associations*. John Benjamins. 3–52.
- Koreny, Éva 1970: Ara ‘Bruder der Mutter’. – *Nyelvtudományi Közlemények* 72: 423–480.
- 1972: *Iranische Lehnwörter in den obugrischen Sprachen*. Budapest: Akadémiai Kiadó.
- Korhonen, Mikko 1988: Uralilaisten kielten jälkitavujen vokaaliston historiaa. – *Virittäjä* 92: 8–25.
- Korn, Agnes 2005: *Towards a Historical Grammar of Balochi*. Wiesbaden: Reichert.
- Kortlandt, Fredrik 2010: Indo-Uralic and Altaic. *Studies in Germanic, Indo-European and Indo-Uralic*. 415–418.
- Kroonen, Guus 2013: *Etymological dictionary of the Proto-Germanic inherited lexicon*. Leiden Indo-European Etymological Dictionary Series, vol. 11; Leiden – Boston: Brill.
- Kulikov, Leonid 2007: Review of Southworth, Franklin C. 2005: *Linguistic archaeology in South Asia*. – *Acta Orientalia Vilniensia* 8: 173–177.
- 2017: Indo-Iranian. – Kapović (ed.), 205–213.
- Kulonen, Ulla-Maija 2010: *Fonesteemit ja sananmuodostus. Suomen kontinuaatiivisten u-verbijohdosten historiaa*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Kümmel, Martin 2007: *Konsonantenwandel*. Wiesbaden: Reichert.
- 2009: Verwandte des Indogermanischen? Zur Frage der „eurasiatischen“ und andere Makrofamilien. Vorläufige Fassung, erscheint in den Akten der Arbeitstagung Würzburg 2009. [https://www.academia.edu/342156/Verwandte\\_des\\_Indogermanischen\\_Zur\\_Frage\\_des\\_Eurasiatischen\\_und\\_anderer\\_Makrofamilien](https://www.academia.edu/342156/Verwandte_des_Indogermanischen_Zur_Frage_des_Eurasiatischen_und_anderer_Makrofamilien) [accessed in February 2018]
- 2014: Laryngeal Traces without Laryngeals in Vedic Metre? – Jared S. Klein, Elizabeth Tucker (eds.), *Vedic and Sanskrit Historical Linguistics. Papers of the 13th World Sanskrit Conference*, Vol. III, Delhi: Motilal Banarsidass. 25–44.
- 2015: Thoughts about pre-PIE Stop Systems. Presentation at the colloquium *The precursors of Proto-Indo-European: the Indo-Hittite and Indo-Uralic hypotheses*, Leiden, July 2015.

- 2017: Etymological problems between Indo-Iranian and Uralic. Presentation held at May 2017 in Helsinki.
- 2018a: Substrata in Indo-Iranic and related questions. Handout of a presentation held at 5.6.2018 in the conference Loanwords and substrata in the Indo-European languages, Limoges.
- 2018b: The survival of laryngeals in Iranian. – Lucien van Beek, Alwin Kloekhorst, Guus Kroonen, Michaël Peyrot & Tijmen Pronk (eds.), *Farnah. Indo-Iranian and Indo-European Studies in Honor of Sasha Lubotsky*. Ann Arbor – New York: Beech Stave Press. 162–172.
- 2019: Early Indo-Iranic loans in Uralic: Sounds and strata. Presentation at the colloquium *Contacts Archaeology, genetics, languages Joining forces to shed light on early contacts (4000 BC – 1000 AD) between Indo-European and Uralic speakers*. Suomenlinna / Sveaborg 9.5.2019.
- Kuokkala, Juha 2012: *Jälkitavun labiaalivokaalit saamessa ja itämerensuomessa*. MA thesis. Helsinki. <http://urn.fi/URN:NBN:fi-fe201206085938> [accessed 27 May 2019]
- 2018: Finnic-Saamic labial vowels in non-initial syllables: An etymological evaluation. – Sampsa Holopainen & Janne Saarikivi (eds.), *Peri orthótētos etýmōn. Uusiutuva uraalilainen etymologia*. Uralica Helsingiensia 11; Helsinki: Société Finno-Ougrienne. 11–76.
- Kuryłowicz, Jerzy 1947–1948: Le degré long en indo-iranien. – *Bulletin de la Société de Linguistique de Paris* 44: 42–63
- Kylstra, Andries Dirk 1984: Nachträglicher Diskussionsbeitrag zu dem Aufsatz von Károly Rédei. – *Journal de la Société Finno-Ougrienne* 79: 1–8.
- LÄGLOS = A. D. Kylstra Sirkka-Liisa Hahmo, Tette Hofstra & Osmo Nikkilä, *Lexikon der älteren germanischen Lehnwörter in den Ostseefinnischen Sprachen I–III*. Amsterdam: Rodopi 1991–2012.
- Larsson, Lars-Gunnar 1984: Estnisch *piim* und finnisch *piimä* – ein baltisches Lehnwort? – *Fenno-Ugrica Suecana* 7: 129–140.
- Lehrman, Alexander 1987: Anatolian cognates for the Proto-Indo-European word for ‘wolf’. – *Die Sprache* 33: 13–18.
- Lehtimäki, Pekka 1972: *Oi- ja ei-nominit Länsi-Uudenmaan murteissa*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Lehtinen, Tapani 2007: *Kielen vuosituhannet*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Lehtiranta, Juhani 2001: *Yhteisaamelainen sanasto*. Second edition. Mémoires de la Société Finno-Ougrienne 200; Helsinki: Société Finno-Ougrienne.
- Lehtisalo, Toivo 1936: *Über die primären ururalischen Ableitungssuffixe*. Mémoires de la Société Finno-Ougrienne 72; Helsinki: Société Finno-Ougrienne.
- Lidén, Evar 1897: *Studien zur altindischen und vergleichenden Sprachgeschichte*. Uppsala: Almqvist & Wiksells.
- Ligeti, Lajos 1986: *A magyar nyelv török kapcsolatai a honfoglalás előtt és az Árpád-korban*. Budapest: Akadémiai Kiadó.
- Liimola, Matti 1957: Sananselyksiä. – *Virittäjä* 61: 191–203.
- 1971: Review of Károly Rédei: Die syrjänischen Lehnwörter im Wogulischen. – *Finnisch-Ugrische Forschungen* XXXIX: 88–95.
- 1973: Review of Korenchy 1972. – *Sananjalka* 15: 273–275.
- Liljegren, Henrik 2008: *Towards a grammatical description of Palula. An Indo-Aryan language of the Hindu Kush*. PhD dissertation. University of Stockholm.
- Lindeman, Fredrik Otto 1987: *Introduction to the Laryngeal theory*. Oslo: Instituttet for Sammenlignende Kulturforskning.
- 1993: Koivulehto Jorma. Uralische Evidenz für die Laryngaltheorie (Book Review). *Indogermanische Forschungen* 98: 309–312.

- Lipp, Reiner 2009: *Die indogermanischen und einzelsprachlichen palatale im Indoiranischen*. Indogerman. Bibl. 3. Reihe; Heidelberg: Winter. Band I & II.
- Liukkonen, Kari 1999: *Baltisches im Finnischen*. Mémoires de la Société Finno-Ougrienne 235; Helsinki: Société Finno-Ougrienne.
- LIV<sup>2</sup> = Helmut Rix. *Lexikon der indogermanischen Verben. Die Wurzeln und ihre Primärstambildungen*. Bearbeitet von Martin Kümmel, Thomas Zehnder, Reiner Lipp und Brigitte Schirmer. Zweite, erweiterte und verbesserte Auflage, bearbeitet von Martin Kümmel und Helmut Rix. 2001. Wiesbaden: Dr. Ludvig Reichelt Verlag.
- Lubotsky, Alexander 1988: *The System of Nominal Accentuation in Sanskrit and Proto-Indo-European*. Leiden – Boston: Brill.
- 1990: La loi de Brugmann et \*H<sub>3</sub>e. – Jean Kellens (ed.), *La reconstruction des laryngales. Actes du Colloque sur la Reconstitution phonologique organisé à Liège les 4-5 décembre 1987*. Ed. Jean Kellens. Université de Liège – Bibliothèque de la Faculté de Philosophie et Lettres. 129–136.
- 2001a: Reflexes of Proto-Indo-European \*sk in Indo-Iranian. – *Incontri linguistici* 24: 25–57.
- 2001b: The Indo-Iranian substratum. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: linguistic and archaeological considerations*. Mémoires de la Société Finno-Ougrienne 242; Helsinki: Société Finno-Ougrienne 301–317.
- 2018: The phonology of Indo-Iranian. – Jared Klein & Brian D. Joseph with Matthias Fritz (eds.), *Handbook of Comparative and Historical Indo-European Linguistics. An international handbook*. Vol. 1. Berlin: De Gruyter Mouton. 1875–1887.
- Lushnikova, Alla V. = Алла В. Лушникова 1990: *Стратификация Ирано-Уральских языковых контактов*. Unpublished dissertation. Moscow.
- Mägiste, Julius 1959: Gibt es baltische Lehnwörter im Tscheremissischen? – *Ural-Altäische Jahrbücher* 31: 169–175.
- Mallory, James & Adams, Douglas Q. 2006: *The Oxford introduction to Proto-Indo-European and the Proto-Indo-European world*. Oxford: Oxford University Press.
- Mallory, Jim 1991: *In search of the Indo-Europeans. Language, Archaeology and Myth*. London: Thames and Hudson.
- 2015: The problem of Tocharian origins. – *Sino-platonic papers* 259: 1–63.
- 2001: Uralics and Indo-Europeans: Problems of time and space. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: linguistic and archaeological considerations*. Mémoires de la Société Finno-Ougrienne 242; Helsinki: Société Finno-Ougrienne
- Malzahn, Melanie 2006: Kunstsprachliches und archaisches in der vedischen Metrik am Beispiel von distrahiert zu messendem e: neue Evidenz für alte Laryngalpräsenz. – Georges-Jean Pinault & Daniel Pétit (eds.), *La langue poétique indo-européenne*. Collection linguistique de la Société de linguistique de Paris, 91; Leuven – Paris: Peeters. 265–290.
- Martirosyan, Hrac 2010: *Etymological dictionary of the Armenian inherited lexicon*. Leiden Indo-European Etymological Dictionary Series, Vol. 10; Leiden: Brill.
- 2013: The place of Armenian in the Indo-European language family: the relationship with Greek and Indo-Iranian. – *Journal of Language Relationship* 10: 85–137.
- Matasović, Ranko & Tijmen Pronk & Dubravka Ivšić & Dunja Brozović Rončević 2016: *Etimološki rječnik hrvatskoga jezika. I. svezak A–Nj*. Zagreb: Institut za hrvatski jezik I jezikoslovlje.
- Matasović, Ranko 2005: The Centum Elements in Balto-Slavic. – Gerhard Meiser & Olav Hackstein (hrsg.), *Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft*. Wiesbaden: Reichert. 171–179.

- Mayrhofer, Manfred 1952: Zur Restproblematik des „Brugmannschen Gesetzes. – *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 70. Bd., 1./2: 8–19.
- 1970: Germano-iranica. – *Zeitschrift für Vergleichende Sprachforschung* 84: 224–230.
- 1984: *Lassen sich Vorstufen des Uriranischen nachweisen?* Wien: Verlag der österreichischer Akademie der Wissenschaften.
- 1989: Vorgeschichte der iranischen Sprachen; Uriranisch. – Rüdiger Schmitt (hrsg.), *Compendium linguarum iranicarum*. Wiesbaden: Dr. Ludwig Reichert Verlag, 4–24.
- 2003: *Die Personennamen in der Rgveda-Samhitā. Sicheres und Zweifelhafes*. München: Verlag der Bayerischen Akademie der Wissenschaften (in Kommission beim Verlag C.H. Beck).
- 2005: *Die Fortsetzung der indogermanischen Laryngale im Indo-Iranischen*. Österreichische Akademie der Wissenschaften. Philosophis-Historische Klasse. Sitzungsberichte, 730. Band; Wien: Verlag der österreichischer Akademie der Wissenschaften.
- McCone, Kim 1987: Hund, Wolf und Krieger bei den Indogermanen. – Wolfgang Meid (hrsg.), *Studien zum Indogermanischen Wortschatz*. Innsbruck: Institut für Sprachwissenschaft der Universität. 83–99.
- Melchert, Craig H 2000: Critical response to the last four papers. – Robert Drews (ed.), *Greater Anatolia and the Indo-Hittite language family. Papers presented at a colloquium hosted by the University of Richmond, March 18–19, 2000*. Journal of Indo-European Studies Monograph Series 38; Washington: Institute for the Study of Man. 229–235.
- Metsäranta, Niklas 2012: *Permiläisten kielten sanastosuhteista*. MA thesis. University of Helsinki.
- 2017: Päivitettyjä permiläisiä etymologioita ja rinnastuksia. – *Journal de la Société Finno-Ougrienne* 96: 211–241.
- 2018: Marin ja permiläisten kielten jakamista ääntenmuutoksista – aitoa konvergenssia vai kielikunnan sisäistä lainautumista? – Sampsa Holopainen & Janne Saarikivi (eds.), *Peri orthótētos etýmōn. Uusiutuva uralilainen etymologia*. Uralica Helsingiensia 11; Helsinki; Société Finno-Ougrienne. 91–134.
- [forthcoming]: *Uralilaisen kielikunnan sisäiset lainasanat – sanastontutkimusta ja etymologiaa painopisteenä permiläiset kielet ja mari*. PhD dissertation.
- Mészáros Edit 2001: A mordvin rokonságnevek eredet szerinti vizsgálata. – *Néprajz és Nyelvtudomány* 41: 169–184.
- Mikkola, Jooseppi Julius 1894: *Berührungen zwischen den westfinnischen und slavischen Sprachen: I: Slavische Lehnwörter in den westfinnischen Sprachen*. Helsingfors: Druck der finnischen Litteraturgesellschaft.
- 1902: Sananselytksiä. 1. Lohikäärme. 2. Marras. – *Virittäjä* 6: 70–72.
- Mikola, Tibor 2004: *Studien zur Geschichte der Samojedischen Sprachen*. Aus dem Nachlass herausgegeben Béata Wagner-Nagy. Studia uralo-altaica 45; Szeged: Szegedi Tudományegyetem.
- Milanova, Veronika & Sampsa Holopainen & Jeremy Bradley [forthcoming]: Contact phenomena in IE kinship and social terms and beyond. – Proceedings of the conference *Loanwords and substrata in the Indo-European languages*, Limoges, 4.–7.6.2018.
- Minissi, Nullo 1970: *Come utilizzare la documentazione finnougriistica nella linguistica indoeuropea*. Napoli: Istituto universario orientale.
- A. Molnár, Ferenc 2003: A magyar hosszúmássalhangzó-rendszer kialakulásáról és néhány finnugor eredetű szavunk hangtani fejlődéséről. – *Nyelvtudományi közlemények* 100: 187–198.

- Morgenstierne, Georg 1938: *Indo-Iranian frontier languages. II. Iranian Pamir languages (Yidgha-Munji, Sanglechi-Ishkashmi and Wakhi)*. Oslo: Instituttet sammenlignende kulturforskning.
- 2003: *The new etymological vocabulary of Pashto*. Edited by J. H. Elfenbein, D. N. MacKenzie, Nicholas Sims-Williams. Beiträge zur Iranistik 23; Wiesbaden: Reichert.
- Moór, Elemer 1956: Die Ausbildung des ungarischen Volkes im Lichte der Laut- und Wortgeschichte. – *Acta Linguistica Academiae Scientiarum Hungaricae* No. 7. 3/4: 341–377.
- 1965: Úr és ország. – *Nyelvtudományi Közlemények* 67: 139–42.
- MSzFE = Lakó, György (főszerk.) 1971–1978: *A magyar szókészlet finnugor elemei*. Budapest: Akadémiai Kiadó.
- Munkácsi, Bernát & Béla Kálmán 1986: *Wogulisches Wörterbuch*. Budapest: Akadémiai Kiadó.
- Munkácsi, Bernát 1901: *Árja és kaukázusi elemek a finn-magyar nyelvekben*. Budapest: Magyar Tudományos Akadémia.
- 1903: Verschiedenheit in den arischen Lehnwörtern der finnisch-magyarischen Sprachen. – *Keleti Szemle* 4: 374–384.
- MWb = Heikki Paaonen, Kaino Heikkilä & Martti Kahla 1990–1999: *H. Paasonens Mordwinisches Wörterbuch*. Lexica Societatis Fenno-Ugricae XXIII. (Kotimaisten kielten tutkimuskeskuksen julkaisuja 59); Helsinki: Société Finno-Ougrienne.
- Napol'skikh, Vladimir 2001: Tocharisch-uralische Berührungen: Sprache und Archäologie. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: linguistic and archaeological considerations*. Mémoires de la Société Finno-Ougrienne 242; Helsinki: Société Finno-Ougrienne
- 2002: Beziehungen zwischen Finno-Ugriern und zentraleuropäischen Indogermanen. – Rogier Blokland & Cornelius Hasselblatt (eds.), *Finno-Ugrians and Indo-Europeans: linguistic and literary contacts. Proceedings of the Symposium at the University of Groningen, November 22–24, 2001*. Studia Fenno-Ugrica Groningana 2; Maastricht. 265–271.
- Napol'skikh, Vladimir = Напольских, В.В. 2014 2010: Уральско-арийские взаимоотношения: история исследований, новые решения и проблемы. – Е.В. Яровой (ответственный редактор), *Индоевропейская история в свете новых исследований: сборник научных статей (Сборник трудов конференции памяти профессора В.А. Сафронова)*. Москва: Московский государственный областной университет. 229–242.
- 2014: Проблема начала финно-угорского-иранских контактов. – *Ананьинский мир: истоки, развитие, связи, исторические судьбы*. Археология евразийских степей. Вып. 20; Казань. 76–89.
- 2015: К происхождению названий соли в финно-пермских языках. – *Linguistica Uralica* 51, Issue 3: 161–176.
- Németh, Dániel 2017: Historical comparison of the Hungarian etnonim's etymologies [sic!]. Presentation held in IFUSCO 2017, 18–22 September, University of Warsaw.
- Nikkilä, Osmo 1988: Aho, Sirkka-Liisa Hahmo, Tette Hofstra, A.D. Kylstra, Osmo Nikkilä. *Omaa vai lainattua: itämerensuomen germaanisiin lainasanoihin liittyviä kirjoitelmaa*. Helsinki: Suomalaisen Kirjallisuuden Seura. 135–140.
- 1901: Auf den Spuren des Pferdes. Ist ostseefinnisch *hepo* doch ein germanisches Lehnwort? – *Ural-Altäische Jahrbücher Neue Folge* 10: 1–8.
- 1994: Germanischer Einfluß auf das Ostseefinnische. – *Incontri linguistici* 17: 103–113.
- Nikolaev, Alexander 2014: Greek εἰαμενή, Vedic *yávasa-*. – *Münchener Studien zur Sprachwissenschaft* 68: 127–140.
- NIL = Wodtko, Dagmar S., Britta Irslinger & Carolin Schneider 2008: *Nomina im indogermanischen Lexikon*. Heidelberg: Winter.

- Nirvi, Ruben Erik 1952: *Synonyymitutkimuksia sukulaisnimistön alalta*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Nowicki, Helmut 1976: *Die neutralen s-Stämme im indo-iranischen Zweig des indogermanischen*. Diss. Würzburg.
- Nussbaum, Alan 1986: *Head and horn in Indo-European*. Berlin – New York: de Gruyter.
- Nuutinen, Olli 1987: Balttilaisten lainojen substituutiotaipauksia. – *Virittäjä* 91: 52–70.
- Oberlies, Thomas 2017: The evolution of Indic. – Jared Klein & Brian D. Joseph with Matthias Fritz (eds.), *Handbook of Comparative and Historical Indo-European Linguistics. An international handbook*. Vol. 1. Berlin: De Gruyter Mouton. 447–470.
- Ofitsch, Michaela 1995: Zu den anlautenden Laryngalen im Hethitischen. – *Historische Sprachforschung* 108: 16–29.
- Olsen, Birgit Anette 1988: *The PIE noun-suffix \*-tlom and its variants*. Copenhagen: The Royal Danish Academy of Sciences and Letters.
- 2005: On Iranian dialectal diversity in Armenian. – Gerhard Meiser & Olav Hackstein (hrsg.), *Sprachkontakt und Sprachwandel, Akten der XI. Fachtagung der Indogermanischen Gesellschaft, Halle an der Saale, 17. - 23. September 2000*. Wiesbaden: Reichert. 473–482.
- 2009: Martinet's rule of laryngeal hardening: a reappraisal. – Stephanie W. Jamison, H. Craig Melchert & Brent Vine (eds), *Proceedings of the 21 Annual UCLA Indo-European conference*. Bremen: Hempen Verlag. 207–222.
- Paasonen, Heikki 1896: Kielellisiä lisiä suomalaisten sivistyshistoriaan. – *Suomi*, III jakso, 13. osa, 4.
- 1907: A finn és a cseremisiz isten-névről. – *Nyelvtudományi közlemények* 37: 14–21.
- 1908–1910a: Zur Frage von den finnisch-ugrischen s-Lauten im Wortanlaut. – *Finnisch-Ugrische Forschungen VIII*: 66–71.
- 1908–1910b: Ein vorarisches Lehnwort im Mordwinischen. – *Finnisch-Ugrische Forschungen VIII*: 75–77.
- 1909: Über ein altes arisches lehnwort. – *Finnisch-Ugrische Forschungen IX*: 120–122.
- 1917: Alkuper. *str-* ja *tr-*yhtymän edustus länsisuomalaisissa kielissä. – *Virittäjä* 24: 66–67.
- 1926: *Ostjakisches Wörterbuch nach den Dialekten an der Konda und am Jugan*. Zusammengest., neu transkr. und hrsg. von Kai Donner. Lexica Societatis Fenno-Ugricae II; Helsinki: Société Finno-Ougrienne.
- Palunčić, Filip 2013: On Ossetic metathesis. – *Studia Iranica* 42: 23–41.
- Parpola, Asko 1999: Varhaisten indoeurooppalaiskontaktien ajoitus ja paikannus kielellisen ja arkeologisen aineiston perusteella. – Paul Fogelberg (toim.) *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*. Bidrag till kännedom av Finlands natur och folk 153; Helsinki: Suomen Tiedeseura. 179–206.
- 2005: The Nāsatyas, the Chariot and Proto-Aryan Religion. – *Journal of Indological Studies*, Nos. 16 & 17: 1–63.
- 2010: New etymologies for some Finnish words. – Klaus Karttunen (ed.), *Anantam Sāstram. Indological and linguistic studies in honour of Bertil Tikkanen*. Studia Orientalia 108; Helsinki: Finnish Oriental Society. 305–318.
- 2012: Formation of the Indo-European and Uralic (Finno-Ugric) language families in the light of archaeology. – Riho Grünthal & Petri Kallio (eds.), *A linguistic map of prehistoric Northern Europe*. Mémoires de la Société Finno-Ougrienne 266; Helsinki: Société Finno-Ougrienne. 119–184.
- 2015: *The Roots of Hinduism*. Oxford: Oxford University Press.
- 2017: Finnish *vatsa* ~ Sanskrit *vatsá* and the formation of Indo-Iranian and Uralic languages. – *Journal de la Société Finno-Ougrienne* 96: 245–286.
- 2019: *Jumal'avita!* – Santeri Junntila & Juha Kuokkala (toim.), *Petri Kallio rocks!* Helsinki. 139–152.

- Peters, Martin 1980: *Untersuchungen zur Vertetung der indogermanischen Laryngale im Griechischen*. Veröffentlichungen der Kommission für Linguistik und Kommunikationsforschung, Heft 8; Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Peyrot, Michaël 2018: Tocharian *etswe* ‘mule’ and Eastern Iranian. – Lucien van Beek, Alwin Kloekhorst, Guus Kroonen, Michaël Peyrot & Tijmen Pronk (eds.), *Farnah. Indo-Iranian and Indo-European Studies in Honor of Sasha Lubotsky*. Ann Arbor – New York: Beech Stave Press. 270–283.
- Pinault, Georges-Jean 2001: Le type latin *uorāgō*: un reflet d'un suffixe indo-européen. A reflection of an Indo-European suffix. – *Glotta. Zeitschrift für griechische und lateinische Sprache* 77 (1–2): 85–109.
- 2007: Review of Katz (2003). *Bulletin de la Société Linguistique de Paris* 102/2: 249–253.
- 2012: Remarks on PIE amphikinetic and hysterokinetic nouns. – Benedicte Nielsen Whitehead, Thomas Olander, Birgit Anette Olsen and Jens Elmegård Rasmussen (eds.), *The Sound of Indo-European. Phonetics, Phonemics and Morphophonemics*. Copenhagen: Museum Tusulanum Press. 399–424.
- 2016: Increase in flourishing: Vedic *púramdhi-*, Av. *parəndi-*. – *Historische Sprachforschung* 129: 83–107.
- Polomé, Edgar C. 1990: Etymology. – Edgar C. Polomé (ed.) *Research guide on language change*. Berlin – New York: Mouton de Gruyter. 415–440.
- Ронарьядов = Понарьядов, В. В. 2012. Насколько достоверный современные реконструкции прафинно-пермского вокализма? – *Вопросы финно-угорской филологии* 3 (8): 78–91.
- Pooth, Roland 2015: A typological overview of Proto-Indo-European. Manuscript. 2015-07-23. [https://www.academia.edu/11855666/A\\_typological\\_overview\\_of\\_Prot Indo-European](https://www.academia.edu/11855666/A_typological_overview_of_Prot Indo-European) [accessed 24 May 2019]
- Pronk, Tijmen 2013: Several Indo-European Words for ‘Dense’ and Their Etymologies. – *Journal of Indo-European Studies* 41, 1&2: 1–19.
- Pystynen, Juho 2015: Semivowel losses and assimilations. Presentation in CIFU XII, August 19 2015. Oulu. [https://www.academia.edu/15172786/Semivowel\\_losses\\_and\\_assimilations\\_in\\_Finnic\\_and\\_beyond](https://www.academia.edu/15172786/Semivowel_losses_and_assimilations_in_Finnic_and_beyond) Accessed September 2016 [accessed 24 May 2019]
- 2017: The State of the Finno-Permic Lexicon. Presentation at the conference Contextualizing historical lexicology, 25 May 2017, University of Helsinki. <https://blogs.helsinki.fi/hist-lexicology/files/2017/.../chl-pystynen.pdf> [accessed 24 May 2019]
- 2018: *Itämerensuomen pitkien vokaalien alkuperä*. Master's thesis. Helsinki. [https://helda.helsinki.fi/bitstream/handle/10138/236213/Pystynen\\_Juho\\_Pro\\_gradu\\_2018.pdf](https://helda.helsinki.fi/bitstream/handle/10138/236213/Pystynen_Juho_Pro_gradu_2018.pdf) [accessed 27 May 2019]
- 2019: Kantasuomen keskivokaalit: vahvempia ja heikompiä tuloksia. – Santeri Junttila & Juha Kuokkala (toim.), *Petri Kallio rocks*. Helsinki. 42–47.
- Ramat, Paolo & Anna Giacamo Ramat (eds.) 1998: *The Indo-European languages*. Routledge Language Family Series. London & New York: Routledge.
- Rau, Jeremy 2009: *Indo-European Nominal Morphology: The Decads and the Caland System*. Innsbrucker Beiträge zur Sprachwissenschaft 132; Innsbruck: Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Rauhala, Ilona 2019: Kuvussa ja kuopalla – suomen sanan *kupa* ja saamen sanan *gohpi* etymologista tarkastelua. – Santeri Junttila & Juha Kuokkala (toim.), *Petri Kallio rocks*. Helsinki. 161–171.
- Ravila, Paavo 1966: *Johdatus kielihistoriaan*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Rédei, Károly 1983: Die ältesten indogermanischen Lehnwörter der uralischen Sprachen. – Juha Janhunen, Anneli Peräniitty & Seppo Suhonen (toim.), *Symposium saeculare*

- Societatis Finno-Ugricae*. Mémoires de la Société Finno-Ougrienne 185; Helsinki: Société Finno-Ougrienne.
- 1986a: Antwort auf den Diskussionsbeitrag von A.D. Kylstra. – *Journal de la Société Finno-Ougrienne* 80: 263–267.
- 1986b: Antwort auf eine ”Kritik”. – *Kratylos* 31: 217–221.
- 1986c: *Zu den indogermanisch-uralischen Sprachkontakten*. Veröffentlichungen der Kommission für Linguistik und Kommunikationsforschung, Heft 16; Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- 1988a: Die ältesten indogermanischen Lehnwörter der uralischen Sprachen. – Denis Sinor (ed.) *The Uralic languages: description, history and foreign influences*. Handbuch der Orientalistik 8: Handbook of Uralic studies 1; Leiden: E. J. Brill. 638–664.
- 1988b: Mordvin népnév eredetéről. – Péter Domokos and János Pusztay (szerk.), *Bereczki emlékkönyv*. Uralistikai tanulmányok 2; Budapest: ELTE BTK. 315–317.
- 2002: Etimológia a nagy szótári vállalkozások után. – *Nyelvtudományi közlemények* 99: 216–234.
- 2004a: Szófejtések. – *Nyelvtudományi közlemények* 101: 130–136.
- 2004b: Postume Etymologien (review of Katz 2003). – *Finnisch-Ugrische Forschungen* 58: 318–325.
- Reshetnikov, Kirill & Mikhail Zhivlov 2011: Studies in Uralic vocalism. – *Journal of language relationship* 5: 96 – 109.
- Riese, Timothy 2001: *Historische Nominalderivation des Wogulischen*. Studia Uralica 10; Wiesbaden: Harrasowitz.
- Ringe, Don 1995: ”Nostratic” and the Factor of Chance. – *Diachronica* 12: 55–74.
- 2004: Reconstructed ancient languages. – Roger D. Woodard (ed.), *The Cambridge Encyclopedia of the World’s Ancient Languages*. Cambridge: Cambridge University Press. 1112–1128.
- Rintala, Päivi 2003: *Iha köyhän laihakin. Tutkimus itämerensuomen iha-kantaisista sanoista*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Ritter, Ralf-Petter 1977: Zur Frage der finnischen Evidenz für die Sieverssche Regel im Germanischen. – *Die Sprache* 23: 171–179.
- RIVELEX = Thomas Krisch et al. 2006–2012: *Rigveda-Lexikon*. Bd. 1–2. Graz: Leykam.
- Rix, Helmut 1991: Urindogermanisch \**gheslo-* in den südindogermanischen Ausdrücken für ”1000”. – Lambert Isebaert (ed.), *Studia Etymologica Indoeuropaea. Memoriae A.J. van Windekens dicatae*. Leuven: Peeters & Departement Orientalistik. 225–232.
- Róna-Tas 2017: Vítás magyar etimológiák. Válasz Honti Lászlónak. – *Nyelvtudományi Közlemények* 113: 37–84.
- Rudnyčkyj, Jaroslav B. 1974: Slavic terms for ‘god’. – Manfred Mayrhofer, Wolfgang Meid, Bernfried Schlerath, Rüdiger Schmitt (hrsg.), *Antiquitates Indogermanicae. Studien zur Indogermanischen Altertumskunde und zur Sprach- und Kulturgeschichte der indogermanischen Völker. Gedenkschrift für Hermann Güntert*. Innsbrucker Beiträge zur Sprachwissenschaft 12; Innsbruck: Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. 111–113 .
- Räsänen, Martti 1920: *Die tschuwassischen Lehnwörter im Tscheremissischen*. Mémoires de la Société Finno-Ougrienne 48; Helsinki: Société Finno-Ougrienne.
- 1956: Einige iranische Entlehnungen in den finnischugrischen und türkischen Sprachen. – Felix Tauer (ed.), *Charisteria Orientalia praecipue ad Persiam pertinentia Jan Rypka*. Praha: Československa Akademie Ved. 274–279.
- 1969: *Versuch eines etymologischen Wörterbuchs der Türksprachen*. Lexica Societatis Fenno-Ugricae 17:1; Helsinki: Société Finno-Ougrienne.
- Saarikivi, Janne & Mika Lavento 2012: Linguistics and Archaeology: A Critical View of an Interdisciplinary Approach with Reference to the Prehistory of Northern Scandinavia. –

- Janne Saarikivi & Charlotte Damm (ed.): *Networks, Interaction and Emerging Identities in Fennoscandia and Beyond*. Mémoires de la Société Finno-Ougrienne 265; Helsinki: Société Finno-Ougrienne.
- Saarikivi, Janne 2007: Uusia vanhoja sanoja. – Jussi Ylikoski ja Ante Aikio (doiammahan), *Sámit, sánit, satnehámit. Riepmočala Pekka Sammallahtii miessemanu 21. beaivve 2007*. Mémoires de la Société Finno-Ougrienne 253; Helsinki: Société Finno-Ougrienne. 325–347.
- 2011: Saamelaiskielet. Historiaa ja nykypäivää. – Irja Seurujärvi-Kari, Petri Halinen & Risto Pulkkinen (toim.), *Saamentutkimus tänään*. Tietolipas 234; Helsinki: Suomalaisen Kirjallisuuden Seura. 77–119.
- 2014: Reconstruction of culture and ethnicity. Remarks on the methodology of historical-comparative linguistics (on the basis of Western Uralic Languages). – Eldar Heide & Karen Bek-Pedersen (eds.), *New Focus on Retrospective Methods: Resuming Methodological Discussions–Case Studies from Northern Europe*. FF Communications 307; Helsinki: Academia Scientiarum Fennica. 186–216.
- 2018: Finnic and other Western Uralic borrowings in Permian. – Sampsa Holopainen & Janne Saarikivi (toim.), *Peri orthótetos etýmōn. Uusiutuva uralilainen etymologia*. Uralica Helsingiensia 11; Helsinki; Société Finno-Ougrienne. 269–356.
- Salminen, Tapani 1999: Euroopan kielet muinoin ja nykyisin. – Paul Fogelberg (ed.), *Pohjan poluilla: suomalaisten juuret nykytutkimuksen valossa*. Bidrag till kännedom av Finlands natur och folk 153; Helsinki: Suomen Tiedeseura. 13–26.
- 2001: The rise of the Finno-Ugric language family. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: linguistic and archaeological considerations*. Mémoires de la Société Finno-Ougrienne 242; Helsinki: Société Finno-Ougrienne. 385–396.
- 2002: Problems in the taxonomy of the Uralic languages in the light of modern comparative studies. – *Лингвистический беспредел: сборник статей к 70-летию А. И. Кузнецовой*. Москва: Издательство Московского университета. 44–55.
- Sammallahti, Pekka 1988: Historical phonology of the Uralic languages with special reference to Samoyed, Ugric and Permian. – Denis Sinor (ed.), *The Uralic languages: description, history and foreign influences*. Handbuch der Orientalistik 8: Handbook of Uralic studies 1; Leiden: E. J. Brill. 478–554.
- 1998: *The Saami languages. An introduction*. Kárásjohka: Davvi Girji.
- 1999: Saamen kielen ja saamelaisten alkuperästä. – Paul Fogelberg (toim.), *Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan*. Bidrag till kännedom av Finlands natur och folk 153; Helsinki: Suomen Tiedeseura. 70–90.
- 2001: The Indo-European loanwords in Saami. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: linguistic and archaeological considerations*. Mémoires de la Société Finno-Ougrienne 242; Helsinki: Société Finno-Ougrienne. 397–416.
- Schalin, Johan 2014: Scandinavian-Finnish language contact in the Viking age in the light of borrowed names – Joonas Ahola, Frog, Clive Tolley (eds.), *Fibula, Fabula, Fact – the Viking Age in Finland*. Studia Fennica Historica 18; Helsinki: Suomalaisen Kirjallisuuden Seura. 399–436.
- Schindler, Jochem 1963: Zu einigen Lehnwörtern im Finnischen. – *Die Sprache* 9: 203–206.
- 1972: *Das Wurzelnamen in Arischen und Griechischen*. PhD thesis. Julius-Maximilians-Universität zu Würzburg.
- Schmeja, Hans 1974: Griechen und Iranier. – Manfred Mayrhofer, Wolfgang Meid, Bernfried Schlerath, Rüdiger Schmitt (hrsg.), *Antiquitates Indogermanicae. Studien zur Indogermanischen Altertumskunde und zur Sprach- und Kulturgeschichte der indogermanischen Völker. Gedenkschrift für Hermann Güntert*. Innsbrucker Beiträge zur Sprachwissenschaft 12; Innsbruck: Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. 377–389.

- Schmid, Wolfgang P 1979: Zur Frage der Datierung Iranischer Lehnwörter in den Finnisch-Ugrischen Sprachen. – Christoph Gläser & János Puszta (hrsg.), *Festschrift für Wolfgang Schlachter zum 70. Geburtstag*. Veröffentlichungen der Societas Uralo-Altaica 12; Wiesbaden: Harrassowitz. 265–270.
- Schmitt, Rüdiger (herausg.) 1989: *Compendium linguarum iranicarum*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- 2009: Iranian Archaisms vs. Vedic Innovations -- and the Indo-Iranian Unity. – Angela Marcantonio (ed.), *The Indo-European Language Family: Questions About Its Status*. Journal of Indo-European Studies Monograph Series, Monograph 55; Washington: Institute for the Study of Man.
- Schneider, Carolin 2010: *Die Maskulinen Stämme auf -man und -iman im Altindischen. Wortbildung, Funktion und indogermanische Grundlagen*. Hamburg: Verlag Dr. Kovac.
- Schrijver, Peter 1991: *The Reflexes of the Proto-Indo-European Laryngeals in Latin*. Leiden: Brill.
- 2001: The lost languages of Northern Europe. – Christian Carpelan, Asko Parpola & Petteri Koskikallio (eds.), *Early Contacts between Uralic and Indo-European: linguistic and archaeological considerations*. Mémoires de la Société Finno-Ougrienne 242; Helsinki: Société Finno-Ougrienne. 417–425.
- 2016: Sound Change, the Italo-Celtic Linguistic Unity, and the Italian Homeland of Celtic. – John T. Koch & Barry Cunliffe (eds.), *Celtic from the West 3 – Atlantic Europe in the Metal Ages: Questions of Shared Language*. Oxford: Oxbow Books. 489–502.
- Schulze, Wilhelm 1933: *Kleine Schriften von Wilhelm Schulze*. Zum 70. Geburtstag am 15. Dezember 1933 herausgegeben von Indogermanischen Seminar der Universität Berlin. Göttingen: Vandenhoeck & Ruprecht.
- Schwartz, Martin 2008: Iranian \*L, and Some Persian and Zaza Etymologies. – *Iran and the Caucasus* 12:2: 281–287.
- Schwyzer, Eduard 1939: *Griechische Grammatik*. Band 1. München: Beck.
- Seldeslachts, Erik 2006: Prakrit-like developments in Old Indo-Aryan. Testing the ‘Kölver-principle’. – Bertil Tikkanen & Heinrich Hettich (eds.), *Themes and Tasks in Old and Middle Indo Aryan Linguistics: Papers of the 12 World Sanskrit Conference*. New Delhi: Motilal Banarsidass. 121–141.
- Setälä, Emil Nestor 1900: I. N. Smirnow’s Untersuchungen über die Ostfinnen. – *Journal de la Société Finno-Ougrienne* 17: 1–52.
- 1912: Eine arische Bezeichnung des Meeres in der finnischen Volkspoesie. – *Festschrift Vilhelm Thomsen zur Vollendung des siebenzigsten Lebensjahres am 25. Januar 1912 dargebracht von Freunden und Schülern*. Leipzig: Harrassowitz. 188–191.
- 1928: Einige vor- und uralische -er und -r- Wörter in den finnisch-ugrischen Sprachen. – *Ungarische Jahrbücher* VIII: 298–308.
- Siegl, Florian 2014: Wanderwörter, areality and borders – the word for ‘bread’ in selected areas of Eurasia. – Nobufumi Inaba, Jorma Luutonen & Arja Hamari (toim.), *Juuret marin murteissa, latvus yltää Uraliin. Juhlakirja Sirkka Saarisen 60-vuotispäiväksi 21.12.2014*. Mémoires de la Société Finno-Ougrienne 270; Helsinki: Société Finno-Ougrienne. 331–339.
- Simon, Zsolt [in press]: Urindogermanische Lehnwörter in den uralischen und finno-ugrischen Grundsprachen: eine Fata Morgana? – *Indogermanische Forschungen*.
- Sims-Williams, Nicholas 1983: Chotano-Sogdica. – *Bulletin of the School of Oriental and African Studies* 46 (1): 40–51.
- 1998: Iranian. – Paolo Ramat & Anna G. Ramat (eds.), 125–153.
- 2000: *Bactrian Documents from Northern Afghanistan, Vol I: Legal and Economic Documents*. Oxford; New York: The Nour Foundation in association with Azimuth Editions and Oxford University Press.

- 2016: *A Dictionary: Christian Sogdian, Syriac and English*. Wiesbaden: Reichert Verlag.
- Sinor, Denis (ed.) 1988: *The Uralic languages: description, history and foreign influences*. Handbuch der Orientalistik 8: Handbook of Uralic studies 1; Leiden: E. J. Brill.
- Sipos, Mária 2002: Az obi-ugor alapnyelv lexikális innovációi I. – *Nyelvtudományi Közlemények* 99: 7–56.
- 2003: Az obi-ugor alapnyelv lexikális innovációi II. – *Nyelvtudományi Közlemények* 100: 245–263.
- Sivula, Jaakko 1986: *Lintu soidessa sokea*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- SKES = Y.H. Toivonen, Erkki Itkonen et al. (eds.) 1955–1981: *Suomen kielen etymologinen sanakirja*. Osat 1–7. Lexica Societatis Finno-Ugriensis XII; Helsinki: Société Finno-Ougrienne.
- Sköld, Hannes 1925: *Die ossetische Lehnwörter in Ungarischen*. Lund: Gleerup. Leipzig: Otto Harrassowitz.
- Sköld, Tryggve 1955–1957: Finn. *osa*, ein indoiranisches Lehnwort? – *Språkvetenskapliga Sällskapets i Uppsala förhandlingar* 1955–1957: 17–26.
- 1983: On the origin and chronology of Saamish (Lappish) words. – Louise Bäckman & Åke Hultkrantz (eds.), *Saami Pre-Christian Religion: Studies on the oldest traces of religion among the Saamis*. Uppsala: Almqvist & Wiksel. 61–62.
- 1984: Fi. *mato*. – *Fenno-Ugrica Suecana* 7: 157–164.
- 1999: Finnisch *pohja*, ein indoiranisches Lehnwort. – *Historische Sprachforschung* 112: 62–68.
- 2002: Finnisch *naaras*, ein Indoiranisches Lehnwort. – Rogier Blokland & Cornelius Hasselblatt (eds), *Finno-Ugrians and Indo-Europeans: linguistic and literary contacts: proceedings of the symposium at the University of Groningen, November 22-24, 2001*. Studia Fenno-Ugrica Groningana 2; Maastricht: Shaker Publishing; Maastricht: Shaker. 350–352.
- de Smits, Merlijn (manuscript): *Murky etymologies*. A draft paper. <https://www.academia.edu/38218892/murkyetymologies.pdf> [last accessed 24 May]
- SSA = Erkki Itkonen & Ulla-Maija Kulonen (päätoimittajat) 1992–2000: *Suomen sanojen alkuperä. Etymologinen sanakirja*. I–III. Helsinki: Kotimaisten kielten tutkimuskeskus & Suomalaisen Kirjallisuuden Seura.
- Steer, Thomas 2015: *Amphikinese und Amphigenese: morphologische und phonologische Untersuchungen zur Genese amphikinetischer Sekundärbildungen und zur internen Derivation im Indogermanischen*. Wiesbaden: Reichert.
- Steinitz, Wolfgang 1944: *Zur Geschichte des finnisch-ugrischen Vokalismus*. Uppsala: Applebergs Boktryckeriaktiebolag.
- Stüber, Karin 2000: Zur Bedeutung und Etymologie von altindisch *śaras-*, iranisch *harah*, und griechisch *ἔλος*. – *Historische Sprachforschung* 113: 132–142.
- 2002: *Die primären s-Stämme des Indogermanischen*. Wiesbaden: Reichert.
- Szemerényi, Oswald 1960: *Studies in the Indo-European System of Numerals*. Heidelberg: Winter.
- 1966: Iranica II. – *Die Sprache* 12: 192–226.
- 1968: The development \*s > \*h in the Indo-European languages. – *Die Sprache* 14: 161–163.
- 1977a: Besprechung von Joki A. J. Uralier und Indogermanen. – *Ural-altaische Jahrbücher* 49: 128–135.
- 1977b: *Studies in the kinship terminology of the Indo-European languages with special reference to Indian, Iranian, Greek and Latin*. Textes et mémoires 7; Acta Iranica 17. Leiden: Brill; Teheran, Liège: Bibliothèque Pahlavi.

- SW = Janhunen, Juha 1977: *Samojedischer Wortschatz: gemeinsamojedische Etymologien*. Castrenianumin toimitteita 17; Helsinki: Helsingin yliopisto.
- Takács, Gábor 1997: Note on the Origin of PIE \*pes 'Penis'. – *Journal of Indo-European Studies* 25: 371–382.
- P. Tálos Endre 1984: Vogul + osztjak/ 2. – *Nyelvtudományi Közlemények* 86/1: 89–99.
- 1987: On the vowels of Proto-Uralic. – Rédei Károly (hrsg.), *Studien zur Phonologie und Morphologie der Uralischen Sprachen*. Studia Uralic 4; Wien: Verband der wissenschaftlichen Gesellschaften Österreichs. 70–80.
- Thieme, Paul 1938: *Der Fremdling im Rgveda: eine Studie über die Bedeutung der Worte ari, arya, aryaman und ārya*. Abhandlungen für die Kunde des Morgenlandes XXIII; 2. Leipzig: Deutsche morgenländische Gesellschaft.
- 1953: *Die Heimat der indogermanischen Gemeinsprache*. Abhandlungen der Akademie der Wissenschaften und der Literatur zu Mainz, Geistes- und Sozialwissenschaftliche Klasse; Wiesbaden: Verlag der Akademie in Kommission bei Franz Steiner Verlag.
- 1963: "Jungfrauengatte" Sanskrit *kaumarah patih*, hom. *koyrídios pósis*, lat. *maritus*. – *Historische Sprachforschung* 78: 161–248.
- 1965: Drei rigvedische Tierbezeichnungen. – *Historische Sprachforschung* 79: 211–223.
- Thomsen, Vilhelm 1869: *Den gotiske Sprogklasses Indflydelse paa de finske*. København: Gyldendalske boghandel.
- 1890: *Beröringer mellem de finske og de baltiske litauisk-lettiske Sprog*. København: Blanco Lunos.
- TESz = Benkó Loránd (főszerk.) 1967–1976: *A magyar nyelv történeti-etimológiai szótára*. I–III. Szerkesztők Lajos Kiss, László Kubinyi & László Papp. Budapest: Akadémiai Kiadó.
- Teubner, Johann Karl 1974: Einige chwaresmische Tiernamen. – Manfred Mayrhofer Wolfgang Meid, Bernfried Schlerath & Rüdiger Schmitt (hrsg.), *Antiquitates Indogermanicae. Studien zur Indogermanischen Altertumskunde und zur Sprach- und Kulturgeschichte der indogermanischen Völker. Gedenkschrift für Hermann Güntert*. Innsbrucker Beiträge zur Sprachwissenschaft 12; Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Thordarson, Fridrik 2009: *Ossetic grammatical studies*. Institutet for sammenlignende kulturforskning, Ser. B, Skrifter, 131; Sitzungsberichte / Österreichische Akademie der Wissenschaften, Philosophisch-historische Klasse, Bd. 788; Veröffentlichungen zur Iranistik, Nr 48.; Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Tichy, Eva 2004: Vedisch *tiráti*. – Adam Hyllested, Anders Richardt Jørgensen, Jenny Helena Larsson & Thomas Olander (eds.), *Per aspera ad asteriscos. Studia Indogermanica in honorem Jens Elmegård Rasmussen sexagenarii Idibus Martiis anno MMIV*. Innsbrucker Beiträge zur Sprachwissenschaft 112; Innsbruck: Institut für Sprachen und Kulturen der Universität. 565–571.
- Tischler, Johann 2002: Bemerkungen zur Urheimatfrage. – Matthias Fritz & Susanne Zeilfelder (hrsg.), *Novalia Indogermanica. Festschrift für Günther Neumann zur 80. Geburtstag*. Graz: Leykam. 475–485.
- Toivonen, Y.H. 1917–1920: Huomioita lainasanatutkimukseen alalta. – *Journal de la Société Finno-Ougrienne* XXXIV: 1–47.
- 1928: Zur Geschichte der finnisch-ugrischen inlautenden Affrikaten. – *Finnisch-Ugrische Forschungen* XIX: 1–270.
- 1956: Über die syrjänischen Lehnwörter im Ostjakischen. – *Finnisch-Ugrische Forschungen* XXXII: 1–32.
- Tremblay, Xavier 1998: Sur *parsui* du Farhang-i-Ōim, *ratu-*, *prtū-*, *pitu-* et quelques autres thèmes avestiques en *-u-*. Essais de grammaire comparée des langues iraniennes III. – *Studia Iranica* XXVII: 187–204.

- 2005a: Irano-Tocharica & Tocharo-Iranica. – *Bulletin of the School of Oriental and African Studies, University of London*, Vol.68, No. 3: 421–449.
- 2005b: Bildeten die iranischen Sprachen ursprünglich eine genetische Familie oder einen Sprachbund innerhalb des indo-iranischen Zweiges? Beiträge zur vergleichenden Grammatik der iranischen Sprachen V. – Gerhard Meiser & Olav Hackstein (Hrsg.), *Sprachkontakt und Sprachwandel, Akten der XI. Fachtagung der Indogermanischen Gesellschaft, Halle an der Saale, 17.–23. September 2000*. Wiesbaden: Reichert. 673–688.
- TschWB = Arto Moisiö & Sirkka Saarinen 2008: *Tscheremissisches Wörterbuch*. Aufgezeichnet von Volmari Porkka, Arvid Genetz, Yrjö Wichmann, Martti Räsänen, T. E. Uotila und Erkki Itkonen. Lexica Societatis Fenno-Ugricae XXXII; Helsinki: Société Finno-Ougrienne, Kotimaisten Kielten tutkimuskeskus.
- Tunkelo, E. A. 1913: Wortgeschichtliche Beiträge. – *Finnisch-Ugrische Forschungen* 13: 74–119.
- UEW = Károly Redei 1988: *Uralisches Etymologisches Wörterbuch*. Band I-II: Unter Mitarbeit von Marianne Bakró-Nagy, Sándor Csúcs, István Erdélyi †, László Honti, Éva Korenchy †, Éva K. Sal und Edit Vértes, Band III: Register Zusammenestellt von Attila Dobó und Éva Fancsaly. Wiesbaden: Harrassowitz.
- Uotila, Eeva 1973: *Zur Etymologie von finn.* sampa. Euroasiatica. 2, Folia philologica AION–SI suppleta 4; Napoli: Istituto Universitario Orientale.
- De Vaan, Michiel 2000: The Indo-Iranian animal suffix \*-ācā-. – *Indo-Iranian Journal* 43: 279–293.
- 2008: *Etymological dictionary of Latin and the other Italic languages*. Leiden Indo-European Etymological Dictionary Series, vol. 7; Leiden – Boston: Brill.
- Vaba, Lembit 1997: Ostseefinnisches meri ‘Meer’ – doch ein baltisches Lehnwort. – Sirkka-Liisa Hahmo, Tette Hofstra, László Honti, Paul van Linde, Osmo Nikkilä (Hrsg.), *Finnisch-ugrische Sprachen in Kontakt. Vorträge des Symposiums aus Anlaß des 30-jährigen Bestehens der Finnougristik an der Rijksuniversiteit Groningen 21.–23. November 1996*. Shaker: Maastricht. 225–31.
- Vainik, Ene 2014: Jumala jälgi ajamas. – *Mäetagused* 58: 7–34.
- Vasmer, Max 1953–1958: *Russisches etymologisches Wörterbuch*. Heidelberg: Winter.
- Vértes, Edit 2003: Széljegyzetek az indo-európai-finnugor kutatásokhoz. – Zoltán Molnár & Gábor Zaicz (eds.), *Permistica et Uralica. Ünnepi könyv Csúcs Sándor tiszteletére*. Fenno-Ugrica Pázmániensia I. Piliscsaba: Pázmány Péter Katolikus Egyetem. 272–278.
- Vilkuna, Kustaa 1933: Ein frühurarisches lehnwort, fi. *ola* ‘feuerstein, kiesel’. – *Finnisch-Ugrische Forschungen* 21: 160–162.
- Vine, Brent 1991: Indo-European and Nostratic. – *Indogermanische Forschungen* 96: 9–35.
- 1998: Indo-European and Nostratic: some further comments. – Joseph C. Salmons and Brian D. Joseph (eds.), *Nostratic: sifting the evidence*. Current Issues in Linguistic theory 142; Amsterdam & Philadelphia: John Benjamins. 85–105.
- Viredaz, Rémy 2017: Trois étymologies ossètes: *rasyg*, *ævzist*, *ærx<sup>o</sup>y*. – *Studia Iranica* 46: 31–48.
- VKS = *Vadja Keele Sõnaraamat*. Toimetanud Elna Adler, Silja Grünberg ja Merle Leppik. 1990–2011. Eesti Keele Instituut. <http://www.eki.ee/dict/vadja/> [last accessed 24 May 2019]
- VMS = *Väike murdesõnastik*. Toimetanud Valdek Pall. Eesti Keele Instituut. <http://www.eki.ee/dict/vms/> [Last accessed 24 May 2019]
- Watkins, Calvert 1957: Latin *marítus*. – E. Pulgram (ed.), *Studies Presented to Joshua Whatmough*. The Hague: Mouton. 277–281.
- 1965: Evidence in Italic. – Werner Winter (ed.), *Evidence for laryngeals*. London: Mouton. 181–188.
- 1995: *How to kill a dragon?* Oxford: Oxford University Press.

- 2011: *The American heritage dictionary of Indo-European roots*. Third edition. Boston and New York: Houghton Mifflin Harcourt.
- Weiss, Michael 2010: Observations on the littera rule. Handout, East Coast Indo-European Conference XXIX, Cornell University, June 19, 2010. [http://conf.ling.cornell.edu/weiss/Observations\\_on\\_the\\_littera\\_rule.pdf](http://conf.ling.cornell.edu/weiss/Observations_on_the_littera_rule.pdf) [last accessed in November 2019]
- Werba, Ludwig 1997: *Verba Indoarica. Die primären und sekundären Wurzeln der Sanskrit-Sprache. Pars I: Radices Primariae*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Widmer, Paul 2002: *Nugae uralo-tocharicae. – Finnisch-Ugrische Mitteilungen* 24/25:171–178.
- Windfuhr, Gernot 2009: Dialectology and topics. – Windfuhr (ed.), 5–42.
- Windfuhr, Gernot 2009 (ed.): *The Iranian Languages*. Routledge language family series. London: Routledge. 5–42.
- WOT = Róna-Tas, András & Árpád Berta (with the assistance of László Károly) 2011: *West Old Turkic. Turkic loanwords in Hungarian*. Wiesbaden: Harrassowitz.
- Winkler, Eberhard 2014: Zum ererbten Wortschatz des Tscheremissischen. – Nobufumi Inaba, Jorma Luutonen & Arja Hamari (toim.), *Juuret marin murteissa, latvus ylittää Uraliin. Juhlakirja Sirkka Saarisen 60-vuotispäiväksi 21.12.2014*. Mémoires de la Société Finno-Ougrienne 270; Helsinki: Société Finno-Ougrienne. 401–414.
- 2017: The comparative method in the Uralic linguistics. – Jared Klein & Brian D. Joseph with Matthias Fritz (eds.), *Handbook of comparative and historical Indo-European linguistics: an international handbook*. Berlin: De Gruyter Mouton. 98–105.
- WOLD = *World Loanword Database*. Edited by Haspelmath, Martin & Tadmor, Uri. <http://wold.cld.org> [last accessed 24 May 2019]
- Ylikoski, Jussi 2019: Kuopimisen alkuperästä. – Santeri Junntila & Juha Kuokkala (toim.), *Petri Kallio rocks*. Helsinki. 191–199.
- Zaicz, Gábor 1998: Mordvin. – Abondolo (ed.), 184–218.
- Zaicz, Gábor (főszerkesztő) 2006: *Etimológiai szótár. Magyar szavak és toldalékok eredete*. Budapest: Tinta Könyvkiadó.
- Zgusta, Ladislav 1955: *Die Personennamen griechischer Städte der nördlichen Schwarzmeerküste: Die ethnischen Verhältnisse, namentlich das Verhältnis der Skythen und Sarmaten, im Lichte der Namenforschung*. Československá akademie věd. Monografie orientálního ústavu 16; Praha: Nakladatelství československé Akademie Věd.
- Zhivlov 2006 = Живлов, Михаил: *Реконструкция праобско-угорского вокализма*. Unpublished dissertation. Moscow.
- Zhivlov, Mikhail 2010: Studies in Uralic vocalism I: A more economical solution for the reconstruction of the Proto-Permic vowel system. – *Journal of Language Relationship* 4. 167–176.
- 2013: Андроновский арийский язык. – Ю.Б. Коряков, А.А. Кибри (редакционная коллегия) *Реликтовые индоевропейские языки Передней и Центральной Азии*. Языки мира. Москва: Academia. 217–220.
- 2014: Studies in Uralic vocalism III. – *Journal of Language Relationship* 12: 113–148.
- 2015: Неиндоевропейский субстрат в финно-волжских языках. X традиционные чтения памяти С. А. Старостина. Handout of a presentation, 27 March 2015. [www.academia.edu/11689161/Неиндоевропейский\\_субстрат\\_в\\_финно-волжских\\_языках](http://www.academia.edu/11689161/Неиндоевропейский_субстрат_в_финно-волжских_языках) [accessed 27 May 2019]
- 2016a: The origin of Khanty retroflex nasal. – *Journal of Language Relationship* 14/4: 293–302.
- 2016b: The origin of Indo-European ablaut in the light of Indo-Uralic comparison. *Illich-Svitych Nostratic Seminar*, 22 December 2016. Handout.

- [https://www.academia.edu/36341248/The\\_origin\\_of\\_Indo-European\\_ablaut\\_in\\_the\\_light\\_of\\_Indo-Uralic\\_comparison](https://www.academia.edu/36341248/The_origin_of_Indo-European_ablaut_in_the_light_of_Indo-Uralic_comparison) [accessed 27 May 2019]
- 2017: Отражение ностратических огубленных гласных в индоевропейском. – М. Н. Саенко (ed.), *Памяти В. М. Иллич-Свитыча: материалы круглого стола*. Москва: Институт славяноведения РАН.
- Zimmer, Stefan 1990: *Urvolk, Ursprache und Indogermanisierung. Zur Methode der Indogermanischen Altertumskunde*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- 1999: Comments on a great book: The Encyclopedia of Indo-European Culture. Review of Mallory and Adams 1997. – *Journal of Indo-European Studies* 27: 105–163
- 2012: Wie kommt slav. *bogъ* zur Bedeutung "Gott"? – Irina Podtergera (hrsg.) *Schnittpunkt Slavistik Ost und West im wissenschaftlichen Dialog Festgabe für Helmut Keipert zum 70. Geburtstag Teil 3: Vom Wort zum Text*. V & R Unipress; Bonn University Press.
- 2017: Culture of the speakers of Proto-Indo-European. – Jared Klein & Brian D. Joseph with Matthias Fritz (eds.), *Handbook of comparative and historical Indo-European linguistics : an international handbook*. Berlin: De Gruyter Mouton. 75–85.

## ABBREVIATIONS OF LANGUAGES

Av	Avestan
: OAv	Old Avestan, YAv = Young Avestan
E	Enets
: FE	Forest Enets
Est	Estonian
Fi	Finnish
Hu	Hungarian
Izh	Izhorian
Ka	Karelian
Kam	Kamas
Kh	Khanty
Li	Livonian
Ma	Mari
Mat	Mator
Md	Mordvin
MI	Middle Iranian
Ms	Mansi
: E	Erzya, M = Moksha
N	Nenets
: FN	Forest Nenets, TN = Tundra Nenets
NP	New Persian
OI	Old Indic (Sanskrit)
OP	Old Persian
Oss	Ossetic
Pfi	Proto-Finnic
PFU	Proto-Finno-Ugric
PIA	Proto-Indo-Aryan
PIE	Proto-Indo-European
PII	Proto-Indo-Iranian
PI	Proto-Iranian
PKh	Proto-Khanty
PMs	Proto-Mansi
POUg	Proto-Ob-Ugric
PSa	Proto-Saami
PS	Proto-Samoyedic
PU	Proto-Uralic
Pug	Proto-Ugric
PWU	Proto-West-Uralic
Sa	Saami
: Kld	= Kildin Saami, L = Lule Saami, N = North Saami, P = Pite Saami S =
South	Saami, Sk = Skolt Saami, T = Ter Saami, U = Ume Saami
SEst	South Estonian (võro-seto)

Slk	Selkup
Sogd	Sogdian
Ve	Veps
Vo	Votic

# APPENDIX I: LATER IRANIAN LOANS IN THE PERMIC LANGUAGES

As many of the loans into Permic languages are clearly late borrowings, they provide little information for the main research questions of this work. Here these loans are dealt with only cursorily, and it has to be stressed that they deserve a later, much larger study. My main purpose here is to show why they are late and were not included in the discussion of the main part of the etymological analysis. The main criteria for late loans are 1) the word has clearly been borrowed after the sound changes that led to the Proto-Permic sound system (the extensive changes in the vocalism, denasalization, the shift *\*ks* > *\*s*, see Sammallahti 1988, Bartens 1990 and Csúcs 2005 for a complete listing of Proto-Permic sound changes); and 2) there clear indications that the word is borrowed from a form akin to modern Ossetic.

Various different versions of Proto-Permic vowel reconstruction have been suggested (Harms 1967; Sammallahti 1988; Csúcs 2005; Zhivlov 2010), the Proto-Permic reconstructions used here follow the system of Csúcs (2005), if no other source is mentioned. This is mainly due to the fact that even though the reconstruction of Csúcs is not necessarily the best, he has presented the most complete listing of Proto-Permic vocabulary so far.

The earlier loanwords that were borrowed into Pre-Permic or earlier stages are treated in the main part of this work.

## LIST OF CONVINCING ETYMOLOGIES

**Ko** *art*, *artal*- < PP *\*art*-

(Csúcs 2005: 321)

See the entry *\*arwa*

**Ko** *žožeg*, **Ud** *žážeg* ‘goose’ < PP *žážeg* ‘goose’ (Csúcs 2005: 327)

← **Indo-Iranian**, cf. PII *\*žhans-á-* > OI *hamsá-* ‘goose, swan’ < PIE *\*g<sup>h</sup>ans-*

(Koivulehto 2001d: 286)

This has to be a late loan due to *\*a*, although the word-initial affricate points to an earlier donor, making the dating of the loan uncertain. The word-initial palatal sibilant is unexpected, but a similar substitution appears in other loans into Permic as well. Koivulehto explains this through influence from the word-initial palatal affricate.

Technically, the Permic word could equally well be explained from Baltic \**žansis* (the same problems with the affricate are present here, though). The reflex of the IE word is not attested in Iranian. The Indo-Iranian etymology is not mentioned by Csúcs (2005: 327).

**Ko *jendon*, *jemdom* ‘steel’; Ud *andan* id. < PP \**andan***

← Alanic \**andon-*, > Oss *andan*, *ændon* ‘steel’ < PII \**sam-dāna-* (Cheung 2002: 161)

(KESK: 99; Joki 1973: 249–250; Rédei 1986c: 65; Csúcs 2005: 320)

This has to be a late borrowing, because it does not show the results of the denasalization. The Komi forms are irregular, probably due to folk etymology.

**Ko *beris* ‘Linde’; Ud *beriz*, *beriz̄* id.**

< PP \**beriz̄*

← Iranian \**barza-* (< PII \**b<sup>h</sup>r̥Hǰa-*), > Oss *bærzæ* ‘birch’ (Cheung 2002: 173; EWAia II: 269–270, s.v. *bhūrjā-*)

(Joki 1973: 255; Rédei 1986c: 66; Lushnikova 1990: 192–193; Csúcs 2005: 321)

Here the vowel \**ä* can point to an Ossetic-type donor language. Cheung mentions that there are very few regular Iranian reflexes of \**b<sup>h</sup>r̥Hǰa-*, and this makes it even more obvious that the Permic word has to be borrowed from a form akin to Ossetic, not from some other Iranian language. It is unclear why Permic has palatal *ž* here. *b* usually appears in later loans, although examples of *b* as a reflex of earlier *p* also appear.

**Ko *čers*, Ud *čers* < PP \**čers* ‘spindle’ (Csúcs 2005: 323)**

< PI \**častra-* ‘spindle’ (see \**kečrā* for references)

(Metsäranta 2012: 36–37)

The Permic word is suspiciously similar to the Indo-Iranian word for ‘spindle’. The affricate *č* is unexpected, and it remains unclear why Indo-Iranian \**č* would have been substituted by Permic \**č*. The consonant-cluster *-rs* has to result from metathesis. Here it is suspicious that Permic has *s* and not \**š*. Perhaps the \**-rš* cluster would not have been allowed. Despite these problems, the etymology is convincing, but many details still need further investigation. \**kečrā* is an earlier loan into PU/PWU from a Pre-Indo-Iranian source (see the entry). According to SSA, already Thomaschek had connected the western Uralic \**kečrā* forms with Peric \**čers*, but they cannot, of course, be regular cognates.

**Ko *dar* ‘Löffel, Tasse’, Ud *duri* id. < PP \**dari/dori* (Csúcs 2005: 330; Zhivlov 2013: PP \**dāri*)**

← cf OI *dārvi-* ‘Löffel’ (EWAia II: 704, s.v. *dārvi-*)

(KESK: 87; Rédei 1986c: 68; Csúcs 2005: 330; Zhivlov 2013: 219; Csúcs 2005: 330)  
 EWAia does not comment on the Iranian forms Par. *durī*, *durē* ‘großer Löffel’ or Nuristani Kati *dur* ‘Löffel’ which are mentioned by Rédei as cognates of the Indic word. The Indo-Iranian borrowing has to be relatively late because the vowel correspondence Komi *a* ~ Udmurt *u* does not appear in inherited Uralic words as the reflex of PU \**a* (Sammallahti 1988: 531). This vowel correspondence could reflect an earlier \**e*-*i*-stem, but it is obvious that the Permic word cannot have been borrowed into this stem type. The Indic word probably had a reflex in Iranian, from which the Permic word was borrowed, or the Permic word is a Wanderwort borrowed via some intermediary.

**Ko, Ud *das* ‘ten’ < PP \**das***

← I \**das*, compare Oss. *dæs*; < PII \**daća-*, > Av *dasa-*

(KESK: 87; Joki 1973: 257; Rédei 1986c: 67; Csúcs 2005: 215–216)

This is probably a late loan as has been usually assumed. Permic voiced *d* corresponds to Iranian voiced *d*. If the loan was Pre-Permic, we would expect similar reflexes of Iranian \**a* that are attested in the Uralic vocabulary. Rédei does not comment on the vowel substitution in any way. A similar relationship between Permic *a* and Iranian *a* is attested in many later loans, probably pointing to the substitution \**ä* ← \**a*.

**Ko *ęksi* ‘emperor’; Ud *eksei*, *äksej* ‘emperor, duke’ < PP \**ekse(j)***

← ? Alanic \**æxsaina-* ‘lady, mistress’, cf. Oss *æxsin-* (Cheung 2002: 16)

(KESK: 211; Joki 1973: 258; Csúcs 2005: 331)

This is probably an Ossetic-type borrowing, although some details require commenting. The Permic word-initial vocalism reflects the Alanic prothetic vowel *æ-*. Rédei derives the word from an Ossetic form related to *æxsin* ‘mistress’, which he considers as related to OP \**xšāyaθiya-* ‘king’, but according to Cheung (2002: 16) *æxsin* is derived from \**xšaita-* ‘glowing’. This \**xšaita-* is reflected by Ossetic *æxsid*, *æxsed* ‘dawn’. The Permic word could be borrowed from *æxsid*, but semantically this is not obvious; perhaps a masculine equivalent of *æxsin* once existed in Alanic, and the Permic word is borrowed from it. Csúcs gives a Middle Iranian form \**äxsē*, but it is unclear how this form has been reconstructed. A direct borrowing from an earlier form \**xšāyaθiya-* is rather unlikely, as the prothetic vowel in Permic would be more difficult to explain, and also because of the lack of the sound change \**ks* > *s* the word cannot be dated to Pre-Permic.

**Ko *gen*, Ud *gon* ‘body-hair’ < ? PP \**gón***

← Iranian \**gun-*, cf. Av *gaona-* ‘hair’, Ossetic *q.yn*, *ğun* ‘hair, wool’, Pashto *yūna* ‘body-hair’ < PII \**gawna-* (Cheung 2002: 215)

(Rédei 1986c: 69; Csúcs 2005: 333)

The Permic word was probably borrowed from a late Iranian word with \**u* (this is assumed by both Rédei and Csúcs) rather than an earlier one with the diphthong \**aw*. Csúcs reconstructs the Proto-Permic form with \**o*, but in the new reconstruction of Zhivlov (2010), this vowel is \**ä*, which makes the substitution of Iranian \**u* phonetically unexpected. It is possible that the Komi and Udmurt words are parallel loans.

**Ko, Ud *gjn* ‘felt’ < PP \**gjn***

(Csúcs 2005: 333)

This is a parallel borrowing from Iranian \**gun-*, see above for the Iranian source. The word can also be a rather early loan, as Permic \**ġ* can reflect Uralic \**u*.

**Old Permic *ideg* ‘angel’ < PP \**ideg***

← **Alanic \**idawæg*, > Oss *dawæg*, *idawæg* ‘angel’ < ? PI \**wi-tāva-ka-*** (Cheung 2002: 178)

(Joki 1973: 264; Rédei 1986c: 70; Helimski 1997b:120; Csúcs 2005: 335)

This word is an obvious Alanic borrowing, with Permic voiced stops reflecting the Ossetic voiced stops and Permic *i-* reflecting the *i* that has been lost in Iron. For alternative Iranian etymologies of the Ossetic word, see Cheung (op. cit.).

**Ko *kam* ‘Balz, Balzeit, Balplatz’ < ? PP \**kam***

< **Iranian \**kām*, cf. OI *kāma-* ‘Wunsch, Begehren, Verlangen’, *kāmín-* ‘Lust habend, verlangend’ MP *kām* ‘Wunsch’ ← PII\**kaHma-*** (EWAia I: 338–339, s.v. *kāma-*, Cheung 2007: 227–228, s.v. \**kahm*)

(a new etymology)

The Komi word is connected to Uralic \**keji-ma-* (> *kiima* ‘rut’) as an uncertain reflex in the UEW, but this is phonologically impossible. The Komi word can be explained as a late loan from < PII \**kaHma-*. Semantically and phonologically this explanation is unproblematic. The word is found only in the Udora dialect of Komi, which makes this assumption slightly uncertain, however.

**Ko, Ud *kar-* ‘to do’ < PP \**kari-***

← **I \**kar-*, Av *kar-*, *kərənaomi* 1SG, MP *kwn* ‘to do’, Parthian *kr-* ‘to do, make’** (Cheung 2007: 236–238, s.v. \**kar-*)

(KESK: 117; Joki 1973: 267; Rédei 1986c: 70; Lushnikova 1990: 214; Csúcs 2005: 341; Zhivlov 2013: 220)

The view that is found in Joki and Rédei, that the loan is a very late one, is correct because of the vocalism *a*. Zhivlov mentions this word among the Andronovo Aryan etymologies,

but there is no obstacle to deriving the word from Iranian. However, the word is unlikely to be from Alanic, as at least in modern Ossetic the stem with *-r* has become lost: the Ossetic verb is *kanñ/kanun*. We do not know, of course, what the situation in prehistoric Alanic was, but there is no way to prove that the stem *kar-* existed in a predecessor of Ossetic.

**Ko *keñ*, Ud *kenem* ‘hemp’ < PP \**ken* < ? Pre-P \**känä***

← Iranian, cf. Oss *gæñ*, *gænæ*, Sogdian *kynb* ‘hemp’

(Joki 1973: 270–271; Rédei 1986c: 70; Helimski 1997b: 121)

Helimski (1997b: 121) considers this a Wanderwort. Even so, Iranian source is the most probable for the Permic word, although the precise donor form can hardly be determined. Due to the Permic *k-* the word does not look like a late borrowing from Alanic. According to Thordarson (2009: 62–63), forms with both *g-* and *k-* are found in modern Ossetic (*k-* is found in *kættæg* ‘linen’, which reflects earlier \**kæntæg*), and this could reflect early dialectal differences, perhaps resulting from the fact that the Ossetic words have been borrowed from two different sources. *g-* in Ossetic is in any case secondary, so prehistoric Alanic could have had the form with *k-*. The word is attested already in Herodotus’s list of Scythian words.

The Mari word *kəñe*, *kiñe* that is mentioned by Joki has to be a parallel borrowing from Iranian. Rédei assumes that the Mari word is borrowed from Permic. The two words are certainly not cognates.

EWAia (II: 605, s.v. *šana-*) considers the Iranian forms and OI *šana-* ‘hemp’ cultural words borrowed from some unknown source. For words of similar shape in other Indo-European languages, see Beekes (2010: 636–637, s.v. *κάνναβις*).

**Ko *ńameđ*, *ńameť* ‘toerag’ < PP \**ńamät***

← Iranian \**namad* ‘felt’, Av *nəmata-*, MP *namad* ‘felt’ < PI \**namata-*

(Rédei 1986c: 72; Lushnikova 1990: 237–238)

The Permic vocalism points to a late loan. It is unclear why the word has palatal *ń-*, but this is not a hindrance to the etymology. Semantically the etymology is convincing. Hungarian *nemez* is a well-known parallel borrowing from the same Iranian word (EWUng s.v. *nemez*).

**Ko *ńań*; Ud *ńań***

< PP \**ńań* ‘bread’

← Iranian \**nān*, > MP *nān*, Sogd *nyn* ‘bread’ < PI ? \**nayna-*

(KESK: 202; Joki 1973: 292; Rédei 1986c: 73; Lushnikova 1990: 238–239; Csúcs 2005: 311, 313, 361)

The etymology is convincing both phonologically and semantically. The reason for the palatal nasal in Permic remains uncertain. The word cannot be very old because of the vocalism *a* in Permic, and also because the Permic loans resemble the Middle Iranian forms like MP *nān*. From Permic the word was borrowed into various Siberian languages, such as Khanty and Mansi (see Siegl [2014] for a recent discussion of the Siberian words).

**Ko *nébeg* ‘book; a piece of birch-bark with house marks’ < PP \**nébeg***

← Iranian \**nibēg-*, MP *nibēg* ‘writing, scripture, book’, Av *nipišta-* ‘written down’ < PI \**ni-pišta-*, a verbal adjective from the root \**pays-* ‘to write’

(KESK: 187; Joki 1973: 290; Rédei 1986c: 73–74; Csúcs 2005: 362)

The disyllabic form in Permic, the vowels and the voiced *b* all point to a late borrowing.

**Ko *pad-vež*; Ud *pad-vož* ‘crossroads’**

< PP \**pad* < Pre-P ? \**panta*

← Iranian \**panta-* or \**panda*, MP *pand*, Ossetic *fænd* ‘path’, *fændæ* ‘way’, YAv *pantā-* ‘path’ < PI \**pantHā-* : *paθā-* (Cheung 2002: 185–186)

(KESK: 214; Joki 1973: 292; Rédei 1986c: 74; Csúcs 2005: 365)

This is a semantically convincing etymology, but an alternative option would be to derive the word from a form akin to Av *paða-*, OI *padá-* ‘Tritt, Fußstapfe, Schritt’, Oss *fæd*. The vowels (Permic \**a* corresponding to PI \**a*) point to a late loan, but if the word was borrowed from \**panta-*, the consonantism *d* < *nt* points to borrowing before denasalization. If the word is indeed rather from \**padá-* ‘footprint’ or its later Alanic reflex, there is no reason to assume that the word was borrowed before the Proto-Permic era.

**Ko *pod* ‘foot’; Ud *pid* id. < ? PP \**pod***

← I \**pad-* < PI \**pad-*, Av *pad-* ‘foot’

(KESK: 223; Joki 1973: 303; Rédei 1986c: 75; Lushnikova 1990: 239–240; Csúcs 2005: 368)

This is a very obvious etymology. The borrowing cannot be very old, as the Permic vowels cannot reflect a Uralic \**a-a*-stem, and *d* as the substitute of Iranian *d* also points to a late loan. However, the Permic vowel correspondence here probably points to a slightly earlier borrowing than with words like \**das*, where Permic *a* corresponds to Iranian \**a*. It is possible that the word was borrowed while the Proto-Permic vowel changes were taking place.

**Ko *ram* ‘peaceful, still’ < PP \**ram***

← Iranian \**rām-*, MP *rām* ‘peace’, Av *rāman-* ‘peace, calmness’ < PI \**rāman-*

(Joki 1973: 306–307; Rédei 1986c: 76; Lushnikova 1990: 246; Csúcs 2005: 373)

The vocalism *a* in Permic points to a late borrowing.

**Ko *vurd*; Ud *vudor* ‘otter’ < PP \**vudrV***

← I \**udra-*, YAv *udra-* ‘otter’, Oss *wyrd*, *urdæ id.* (Cheung 2002: 244)

(KESK: 70; Joki 1973: 347–348; Rédei 1986c: 80; Lushnikova 1990: 269–270)

This is a late loan because of Permic *u* corresponding to Iranian *u*. It remains unclear why the Permic forms have secondary *v-*; Metsäranta (personal communication) notes that the word for ‘water’, Ko *va*, Ud *vu* could have influenced the word, or that an analogical *v* based on Russian *выдра* ‘otter’ (cognate to the Iranian word) can be assumed. The Komi word has gone through metathesis. Perhaps the words in Udmurt and Komi are parallel borrowings, as the Komi word could have been borrowed from a form akin to Ossetic where a similar metathesis took place.

**Ko *variš* ‘hawk’, Ud *variš*, *variš* ‘hawk’ < PP \**variš* ‘hawk’** (Csúcs 2005: 293)

← Iranian, cf. MP *varišak/vāriš*, Sogd *w’ryn’k*, Chwarezmian *w’ryruk*, Oss *uāri*, YAv *vārən-gan-* ‘name of a bird’ (Gharib 1995: 398; Abaev 1958–1989)

(Saarikivi 2018: 318)

Although the Permic word is very similar to the Iranian word and looks like an obvious borrowing, the etymology has been published only recently by Saarikivi, who seems have gotten the idea from an anonymous referee of his article. The Permic word has to be a late loan because of the vocalism and the retained non-initial syllable. The word resembles especially the Middle Persian word, which means that the Permic word is probably not an Alanic loan. Perhaps the word for ‘falcon’ has spread as a some kind of a Wanderwort/luxury loan from Middle Persian.

**Mari \**wāräkš* ‘hawk’** (E *warasš*, V U *warakš*, NW *wāräkš*, W *wāräsš*) is also considered an Iranian loan by Saarikivi, but it has been argued that the Mari word has to be a borrowing from Permic instead (Bereczki 1994: 90; Aikio 2014: 140). There is no compelling reason to assume that the Mari word was borrowed from Permic, as it could equally well be borrowed from the Iranian source, but due to the uncertain donor language and the Wanderwort nature of the loan it is possible that the Mari word has been borrowed through Permic. The precise etymology of the Mari word requires further study.

As a side note it can be said that the etymology of the Iranian word is not clear, as it has no secure cognates in Indo-Iranian (AiWb: 1412, s.v. *vārə-gan-*, *vārən-gan*). De Vaan (2003: 183, footnote 137) supports the idea that the Sogdian word reflects the same form as the Avestan word, and that the Iranian has been possibly retained in the loanword *varužan* ‘male bird’ in Armenian. De Vaan supports the old explanation that the Avestan

word is a cognate of *vār-* ‘defence’ and the verb *jan-* ‘to kill’, meaning ‘the one who breaks the defence’.

**Ko *varneŝ* ‘one-year-old lamb’**

< PP *\*varnäs*

← *\*warnā-*, Av *varənā* ‘wool’ (< PI *\*HwrHnaH-*), MP *warrag* ‘lamb, ram’, Pashto *wari*

(Joki 1973: 336; Rédei 1986c: 79; Lushnikova 1990: 276; Zhivlov 2013: 220)

The word has to be late because of the Permic vocalism. Also semantically the word is closer to the attested Middle Iranian words than the Old Iranian or Old Indo-Aryan words that denote ‘wool’. See *\*wurun*.

**Ko *sarid’z* ‘see’; Ud *zarid’z* id.**

← *\*zraya-*

(KESK: 249; Joki 1973: 349; Rédei 1986c: 81–82)

See *\*sar*

**Ko *zarńi*; Ud *zarńi* ‘gold’ < PP *\*zarńi***

← I *\*zaraniya-*, cf. Oss *zæran* ‘gold’

(KESK: 104; Rédei 1986c: 82; Csúcs 2005: 399)

This is obviously a late loan due to the vocalism and the sibilant *z* which does not appear in inherited Pre-Permic words. See also Metsäranta (forthcoming).

**Ko *zon* (: *zonm-*) ‘son’ < PP *\*zɔn***

← Iranian *\*zana-*, Oss *zænæg* ‘child’, Av *zana-* ‘mankind’ < PII *\*z’anHa-*

(KESK: 106; Joki 1973: 349–350; Rédei 1986c: 82; Lushnikova 1990: 291; Csúcs 2005: 399)

The Komi word has a dubious *nm*-stem, but the etymology is still convincing. The word is obviously a relatively late loan because of the sibilant *z*. It is uncertain why Proto-Permic why *\*o* is the vowel substitution here and not *\*a/ä*, but this can possibly be explained by assuming an earlier borrowing like *pod* ‘foot’. The word could have been borrowed while the Proto-Permic vowel changes were taking place. See also Milanova, Holopainen & Bradley (forthcoming).

## LIST OF DUBIOUS ETYMOLOGIES

**Ko** *amiš*, *ameć* ‘Pflugschar’; **Ud** *ameз*, *amež* id. < PP \**ameć/ameз*

← \**āmāč* > **NP** *amāč*, *āmāč* ‘Pflugschar’

(KESK: 211; Rédei 1986c: 64; Csúcs 2005)

Rédei also mentions the Turkic words *amač* ‘plough’, and refers to Doerfer’s view that the Iranian word is borrowed from Turkic, and the Permic word from Iranian. Rédei notes that borrowing from Turkic is not probable because the word is not attested in Chuvash. This is not a very good argument, however. The word is attested only in Persian and in Yaghnobi as a loan, and it is not at all certain that in an Iranian contact language of the Permic languages this word existed too. The Turkic word may simply have disappeared from Chuvash. Helimski (1997b: 121) assumes that the Permic word is a Wanderwort. In any case the word is a late loan because of the *a* vocalism.

**Ud** *ana* ‘ohne’

← **Iranian** \**ana-*, **Oss** *ænæ* ‘ohne’ <, ? \**an-* (< PIE \**ṇ-*)

(Rédei 1986c: 64–65)

This is one of the very few examples of borrowed particles or adverbs. The etymology is in principle convincing, but it is also possible that the resemblance of the Permic and Iranian words is accidental. Also Rédei considers the etymology as uncertain. It is very unlikely that the Permic word is borrowed from the negative prefix \**an-* (< PIE \**ṇ-*) but rather from the Ossetic preposition *ænæ*. Permic \**a* corresponds to Iranian \**a* in the same way as in other examples like \**das*.

**Ud** *ba(d')d'žin*, *ba(d')d'žim* < ? PP \**bazd'in*

← **Iranian**, cf. **Ossetic** *bazgyn*, *bazgin* ‘thick’

(Joki 1973: 255; Rédei 1986c: 65; Csúcs 2005: 321)

The etymology is convincing from the semantic point of view, but the idea that the Permic affricate could result from a metathesis of the Ossetic consonant-cluster \**-zg-* is dubious. Joki argues that the donor form was Pre-Ossetic or Scythian \**baz-d'yn-*, but there is no substance in this reconstruction: how can Ossetic *zg* reflect earlier \**zd'*?

**Ko** *eziš* ‘silver’, **Ud** *uzveš* < PP \**üz-veš*

(KESK: 331–332)

The etymology of this word has long been debated. The word is somehow related to Hungarian *ezüst* ‘silver’ and Ossetic *ævzist*, *ævzæstæ* ‘silver’, but it remains unclear what the direction of borrowing is. The Permic word has also been explained as a compound of the Uralic metal terms *\*äsä-wäškä* (Csúcs 2005: 321; J. Häkkinen 2009: 25–28). For a recent discussion of the Ossetic word, see Viredaz (2017).

**Ud *erži*, *erži* ‘Adler’ < PPK *\*erži***

← ? Iranian *\*rzi-* < *\*rdzi-*, cf. Av *ərəzi-fīa-* ‘Adler’, OI *rjipyá-* ‘Beiwort des Adlers, der Marut’ (EWAia I: 250–251, s.v. *rjipyá-*)

(Rédei 1986c: 68–69; Csúcs 2005: 331)

The etymology is semantically convincing, but due to the fact that the word is not attested in any later Iranian language, the explanation remains uncertain. The affricate here can substitute for an earlier Iranian affricate, but the palatalization is unclear. The Udmurt vocalism is the same as in *beriz*, *beriz̥* ‘Linde’ (see above).

**Ko *mēs* ‘cow, female animal’, Ud *mes* id. < ? PP *\*mes* (Csúcs 2005: 356)**

← ? Iranian, cf. Munji *məškāy* ‘a two-year old calf’ (Morgenstierne 1938: 229)

(KESK: 175)

This etymology is semantically dubious, as the primary meaning of the Permic word seems to be ‘female animal’, whereas the Munji word refers to a ‘male calf’. The Permic word does resemble the word in Munji, but this form has no cognates in other Iranian languages. Morgenstierne derives the Munji word (cautiously) from a pre-form *\*ham-huškaka-*, but because of the lack of formal parallels elsewhere in Iranian, it remains uncertain whether the Munji word can reflect such a Proto-Iranian form. This form could hardly have produced the Permic words.

**Ko *poda* ‘Fußgänger; Vieh’; Ud *pudo* ‘cattle’ < PP *\*poda***

← ? Iranian, cf. Tajik *poda* ‘herd’

(Joki 1973: 304; Rédei 1986c: 75–76; Csúcs 2005: 370)

The etymology is formally unproblematic, but the limited distribution of the word in Iranian raises doubts. Both Joki and Rédei mention that Yaghnobi *padá* might be borrowed from Tajik. If the etymology is correct, the word has to be a late borrowing, because in earlier loans the Iranian *d* has become lost (see *\*gu-*).

Another option that is also mentioned by both Joki and Rédei is that the Permic word is derived from Permic *pad* ‘foot’. Joki is sceptical (“ganz nicht überzeugend”) towards the idea that the word ‘cattle’ could have been derived from ‘foot’.

**Ko *reŋižti* ‘wiederkauen’; Ud *žomešti*- *id.* < PP \**romi/romi***

← ? Iranian \**rōmast*, Waxi *ramöt* ‘Wiederkauen’, Balochi *rōmast id.* < PII \**rawmantHa-* (EWAia II: 470–471, s.v. *romantha-*)

(Joki 1973: 309–310; Rédei 1986c: 66; Lushnikova 1990: 247–248)

The development of the vocalism of the Permic words is unclear (Rédei 1986). The resemblance to the Iranian word can also be mere accident.

**Ud *šaer*, *šair*, *šajir* ‘land, country, foreign’**

← Iranian \**šahr*, cf. MP *šahr* ‘land, country, city’ < PI \**xšaθra-*

(Joki 1973: 319; Rédei 1986c: 77; Lushnikova 1990: 264–265; Csúcs 2005: 381)

The Permic two-vowel sequence is difficult to explain from the Iranian form. If the etymology is correct, the Permic word is obviously a late borrowing due to the *a* vocalism. Lushnikova and Helimski (1997b: 121) consider this word as a Wanderwort. As similar forms are also attested in Turkic languages, such as Tatar *šəhər*, it is possible that the Udmurt word is a Turkic loan, and also the distribution limited to Udmurt speaks in favour of this.

**Ko *tašti* ‘Tasse’; Ud *tušti* ‘Schlüssel, Teller’ < PP *tašti***

← I \**tašta-*, cf. Av *tašta-* ‘Tasse, Schale’, MP, NP *tašt id.*

(KESK: 278; Joki 1973: 331; Rédei 1986c: 78; Lushnikova 1990: 267; Csúcs 2005: 389)

The palatal sibilant is unclear. The Permic vowel correspondences do not point to any Pre-Permic vowel combination. This may be a Wanderwort.

**Ko *veřt* ‘Weberschätze, Litzen’; Ud *vřrt* ‘Rietblatt’ < PP \**vört***

← Iranian?, cf. Oss *urd* ‘Webekette, Schaft’

(Rédei 1986c: 80; Csúcs 2005: 396)

The etymology is formally unproblematic, but there is no certain Indo-Iranian etymology for the Ossetic word, meaning that the Permic and Ossetic words can also be separate borrowings from some third source.

## LIST OF UNCONVINCING ETYMOLOGIES

**Ko** *asiṽ* ‘morning’; **Ud** *asiḷ* ‘morning, east’ < PP ? \**asiḷ*

← **Iranian ?**

(KESK: 34)

This etymology is not mentioned by Joki (1973) nor Rédei (1986). KESK compares the Permic word with Sogdian and Chwarezmian *ās-* ‘high’. Semantically the etymology is not very good and it is unclear which Sogdian and Chwarezmian word is actually meant here, as the word is not found in the Sogdian dictionary of Gharib (1995).

**Ko, Ud** *uri* ‘world; peace’ < PP? \**uri*

← ? **Iranian, cf. OAv** *urvādah-* ‘Freude, Wonne’

(KESK: 298; Rédei 1986c: 78–79)

Rédei considers this etymology as dubious because the Permic word reflects only the first syllable of the Iranian word. Rédei also does not give any exact donor form, but considers the word “Middle Iranian” even though the Avestan word does not have attested cognates in the Middle Iranian languages (AiWb: , s.v. *urvādah-*). The etymology is doubted also by Helimski (1997b: 121). The resemblance of the Permic and Iranian words is probably coincidental.

## APPENDIX II: POSSIBLE URALIC LOANWORDS IN INDO-IRANIAN

Here the Indo-Iranian words with tentative Uralic etymologies are presented, and the etymologies are briefly evaluated. A more detailed survey of these loans remains a task for future research. The possible late loans from Hungarian into Ossetic are not discussed here.

**Fi *Ilmari* ‘the sky god’; Udm *Inmar* id.**

< PU ? \**Ilmari*

→ PII \**Índra-* > OI *Índra-* ‘war-god’, Mitanni *<sup>d</sup>In-da-ra-*, ? YAv *Indra-* ‘name of a *Daeva*’

(Parpola 2015: 66, 109; 2019)

There are various competing Indo-European etymologies for the name of Indra, but none are convincing or widely accepted; see EWAia (I: 192–, s.v. *Índra-*) and RIVELEX (II: 160, s.v. *Índra-*) about the various explanations. One possible etymology that is mentioned by EWAia and RIVELEX is nasal-infix derivation from the PIE root \**h<sub>2</sub>eyd-* ‘schwollen’. Parpola suggests that this problematic deity name could be an early borrowing from Uralic.

There are several problems with Parpola’s explanation. First of all, forms with *r* are attested only in Finnic and Udmurt, and it is difficult to reconstruct a form \**Ilmari* or the like for Proto-Uralic or to a language stage from where the Indo-Aryans could have borrowed the word at a very early period (the borrowing has to be early for both areal reasons and for the fact that Indra appears in the earliest Vedic hymns). Saarikivi (2018: 329–330) has claimed that the Udmurt word might be borrowed from Finnic, and in the light of other Finnic-Permic loanwords this is a possible explanation, although one has to note that not many such loans are found in Udmurt, but mostly in Komi. However, because *r* is not a transparent derivational suffix in Udmurt, the word *Inmar* is etymologically opaque and it does not stand in a clear relationship with the word *inm* ‘sky, heaven’.

A further complication is that fact that the basic PU word for ‘sky’, from which Finnic *ilma* and Udmurt *inm* are derived is \**jelma* in modern reconstruction (cf. Aikio 2014: 9), meaning that even if a predecessor of Finnic *Ilmari* could be reconstructed for PU, it would rather be something akin to \**jelmari*, and it is unlikely that it would have been borrowed as *Indra* into Proto-Indo-Aryan.

Also Parpola’s arguments for the word-medial consonantism of the Indo-Aryan word are unconvincing. While an epenthetic *d* does appear in languages like Greek, this cannot be used as a proper parallel for the Indo-Aryan development. Furthermore, the consonant development that is attested in Udmurt could have hardly occurred so early that it would have been present in the donor language of *Indra*.

Not a Uralic loanword in Indo-Iranian

**Fi *kala*; SaN *guolli* (PSa < \**kuolē*); Mo E, M *kal*; Ma *kol*; Hu *hal* (: *hala-*); PMs \**kūl*; PKh \**kūa*; PS \**kālā***

**PU \**kala* ‘fish’**

→ ? PI \**kara-* > Av *kara-* ‘a mythical fish’

(Kümmel 2018a)

Kümmel assumes a Uralic etymology for the Avestan word, as it does not have a clear Indo-European etymology. IEW (s.v. \*(a)sk<sup>w</sup>al-, \*k<sup>w</sup>al-) connects the word with various Indo-European words meaning various kinds of fish or ‘whale’, such as Lat *squalus*, OPr *kalis*, and Germanic \**hwali-* ‘whale’, but the Avestan word is not connected to these other Indo-European words in modern etymological sources (De Vaan 2008; Kroonen 2013). The Avestan word does not denote a simple fish but a mythological creature, which makes the etymology questionable. In any case the Avestan word is isolated in Iranian, making a chance resemblance with Uralic quite likely. Formally this etymology is possible, as \**l* became \**r* in Iranian.

Schrijver (2001) assumes that the Uralic word and the Indo-European words mentioned above were borrowed from the same substrate source in Northern Europe. Due to geographical reasons, it does not seem very likely that the Avestan word was borrowed from the same substrate language, but this cannot be ruled out for certain. Formally Iranian \**kara-* and the Latin, Baltic and Germanic words could be cognates (the Iranian word could reflect \**k(w)el-o-* or \**k(w)al-o-*).

Blažek (2005: 176–177) has suggested that the Indo-Iranian word could rather be borrowed from Proto-Uralic \**kārV* ‘sterlet’ (UEW s.v. \**kārə*). This etymology is semantically likewise unproblematic, but it is dubious whether such a fish name can be reconstructed to Proto-Uralic. The Uralic reconstruction is considered as uncertain by UEW. The set of cognates consists of Udmurt *karei* ‘sterlet’, Nenets (Tundra) *χiri* ‘kleiner Stör’, (Forest) *kiři* ‘Sterlet’ and Ob-Ugric forms that are reconstructed under POUg \**kārV* by Zhivlov (2006: 134–135): PMs \**kārkaĵ* ‘sterlet’ (> Ms North *kāraĵ*, East *kōrĵaj*, West *kārkaĵ*), PKh \**keri* (> Kh North *kāri*, East *kōri*, South *kārə*) id. The Ob-Ugric vowel-combination does not point to any Uralic source, and Udmurt *a* is also not the regular reflex of PU \**ä*. It remains thus uncertain whether the word can be reconstructed to Proto-Uralic, and also the idea that the word could have been borrowed from Proto-Uralic to (Indo-)Iranian is likewise uncertain.

? Uralic loanword in Indo-Iranian

**Pre-P \**kaða-***

> PP \**gu-* > Ko *gu-* ‘to steal’

→ Iranian \**gada-* > Av *gaða-* ‘thief’

(Katz 1985: 288; 2003: 230; Kümmel 2018a)

See the entry \**gu-*

Not a Uralic loanword in Indo-Iranian

**PU \*kota ‘chum’**

→ **Iranian \*kata > Av kata- ‘house’**

See the entry \*kota

Not a Uralic loanword in Indo-Iranian

**PU (?) “\*kǎñčá” (> Mo E kañčt’) ‘Hanf’ → PII \*ganžā- > OI gañjā- ‘hemp’**

**PU (?)” \*kənü’ (> Ma kəne) ‘Hanf’ → PII \*ćana- > OI śaṇá- ‘eine Hanf-Art’;  
→ Oss gænæ**

**POUg (?) “kəñá” ‘Hanf’ → PI \*kanaba- > Khotanese kaha- ‘hemp’**

(Katz 2003: 143–144)

The words for ‘hemp’ in the Uralic and Indo-European languages are Wanderwörter, and there is no reason to assume that the Indo-Iranian words are borrowed from Uralic. For the etymology of the Iranian and possibly related IE words for ‘hemp’, see EWAia and Beekes. Katz assumes three different borrowings here, but none of the forms he mentions can be reconstructed for Proto-Uralic or any proto-language which would be early enough to provide loanwords into PII or PI.

Not Uralic loanwords in Indo-Iranian

**SaN gádʹfi ‘female stoat’; Mns East kōäl’, North kāl’ ‘female’ (< PMnss \*kāl’);  
Kh East kōjəŋ, kō̃̃jəγ ‘female animal’(< PKh \*kājəγ); ? Hung hölgy ‘lady,  
dame’; Kam šüjü ‘female animal, capercaillie hen’, Mat kejbe ‘mare’ (< PS  
\*kājwä)**

**PU \*käd’wä ‘female animal’ (UEW s.v. kad’wa; Sammallahti 1988: 545; Aikio 2015b:  
63)**

→ **PI \*gadwā- > YAv gaδwā- ‘bitch’**

(Kümmel 2018a)

Semantically this is a plausible etymology. However, in Avestan the word is attested in both masculine (*gaδwa-* ‘dog’) and feminine (*gaδwā-* ‘bitch’) forms (AiWb: 490, s.v. *gaδwa-*), which makes the etymology dubious, as the Uralic word clearly denotes a female animal, and not specifically a dog. It is, of course, possible that the word was borrowed into Iranian in the female word, and the masculine form is derived from it, but

this is impossible to formally prove. The word is isolated in Avestan, so it is questionable whether it reflects a Proto-Iranian formation.

? Uralic loanword in Indo-Iranian

**PU \*kād'wā 'female animal'** (see above)

→ **PI \*kaθwā- > YAv kaθwā- 'female donkey'**

(Kümmel 2018a)

Like **\*gadwā-**, this is also a possible etymology, although it is dubious to assume that the same Uralic word was borrowed into Iranian twice with different substitutions. For this Avestan word, only the female form is attested (AiWb: 435, s.v. *kaθwā-*).

? Uralic loanword in Indo-Iranian

**Md E pango, M panga 'Pilz, Schwamm; Mari pongo, pongâ 'Pilz, Fliegenpilz, Baumschwamm, Meltau; dumm' (< PMA \*pongə); Mansi East pēŋk, West pēŋk, North pāŋx 'Fliegenpilz; Rausch' (< PMs \*p̄ŋkə), South pēŋl- 'betrunken sein' (< PMs (der.) \*p̄ŋl); Kh East pāŋk, South paŋx 'Fliegenpilz' (< PKh \*pāŋk), East paŋkəl- 'singen, nachdem man Fliegenpilze gegessen hat'; Nganasan h<sup>u</sup>aŋkud'a 'betrunken sein' (< PS \*peŋkâ-)**

< **PU \*piŋka 'psychedelic mushroom** (UEW s.v. *\*paŋka*; Zhivlov 2014: 119, 133; Aikio 2015b: 59)

→ **PII ? \*b<sup>h</sup>anga- > ? Av baŋha- 'name of a plant and a narcoticum that is made of it'** (EWAia II s.v. *bhaŋgá-*)

(Katz 1985: 172–173; 2003: 142–143; Kümmel 2018a)

Usually this has been considered a loan in the opposite direction, but the whole Iranian comparison is uncertain because of phonological problems and scarce attestation, see the entry on **\*piŋka**.

Not a Uralic loanword in Indo-Iranian

**Fi salko 'long pole or rod'; SaN čuolggu 'rod for pushing nets under ice', Erzya šalgo 'stick, thorn, stinger'; Komi šuv, Udmurt ded'i-šul 'sled runner' (< PP PP šul); East Khanty sayəλ 'lath' (< PKh \*sāyəλ); North Mansi sāyla 'lath' (< (PMs \*s̄ylā); Hungarian szál-fa 'lumber wood; long log'**

< **PU \*čilka-j ~ \*šilka-j/\*čilka-w ~ \*šilka-w** (UEW: 460–461, No. 929; Zhivlov 2014: 132; Aikio 2015nb: 60)

→ PII \**ćalka-* > OI *śalka-* ‘Span, Abschnitzel’

(Katz 1985: 308; 2003: 246–247)

Since the Uralic word itself is a probable borrowing from some form of Indo-European (Balto-Slavic?), the idea that it would be an early borrowing to Indo-Iranian looks implausible (see the entry \**ćilkaj*).

Not a Uralic loanword in Indo-Iranian

**Hu *szak*; Kh *jöŋ-säk* ‘Eisbrei (im Frühling vom Eis abgebröckelt)’**

< PU ? \**ćik(k)a* (Katz: \**šēkka*)

→ PII \**ćaka-* > OI *śaka-* ‘Stüchen’

(Katz 2003: 284–285)

No such PU form can be reconstructed. UEW (s.v. \**sakkə*) gives an uncertain Proto-Ugric reconstruction. Furthermore, *śakala-* has a possible Indo-European etymology (see the entry \**ćilkaj* in the main text).

Not a Uralic loanword in Indo-Iranian

**Fi *veitsi* ‘knife’; Ko *vež-* ‘перерубить наискось’ < PP \**veži-*; Hu *vés* ‘meißeln, eingraben, stemmen, gravieren’; Kh East *wānt’-*, South *wānt’-*, North *waš-* ‘aufschneiden, spalten’**

< PU \**wəŋci* ‘knife’ (UEW s.v. \**wāŋcV*)

→ \**wāćī-* ‘axe, pointed knife’ > OI *vāćī-*, ? YAv *vāsī-* ‘pointed knife (?)’ (Lubotsky 2001b: 313)

(Parpola 2012: 161; Kümmel 2018a, 2019)

The etymology is semantically and phonologically possible, although the fact that a number of different pre-forms can be postulated for the Indo-Iranian word makes the etymology more hypothetical. A cluster of a nasal and \**ć* could also have been a possible substitution for Uralic \*-*ŋć-*, so the borrowing could have resulted in a form like \*\**wanćī-*.

? Uralic loanword in Indo-Iranian

