

# Uralic studies, languages, and researchers

Edited by Sándor Szeverényi



## **Studia uralo-altaica 54**

### **Redigunt:**

Katalin Sipőcz

András Róna-Tas

István Zimonyi

Uralic studies, languages, and researchers

Proceedings of the 5<sup>th</sup> Mikola Conference  
19–20, September 2019

Edited by Sándor Szeverényi

Szeged, 2021

© University of Szeged,  
Department of Altaic Studies,  
Department of Finno-Ugrian Philology

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by other means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission in writing of the author or the publisher.

Printed in 2021.

Printed by: Innovariant Ltd., H-6750 Algyő, Ipartelep 4.

ISBN 978-963-306-803-8 (printed)

ISBN 978-963-306-804-5 (pdf)

ISSN 0133-4239 (Print)

ISSN 2677-1268 (Online)

## Table of contents

Foreword.....	9
<i>Sándor Szeverényi</i>	
Notes on Nicolaes Witsen and his Noord en Oost Tartarye.....	11
<i>Rogier Blokland</i>	
Undiscovered treasures: From the field research archive to the digital database.....	27
<i>Beáta Wagner-Nagy, Chris Lasse Däbritz, and Timm Lehmborg</i>	
On the language use of the first Finnish medical text .....	45
<i>Meri Juhos</i>	
Sajnovics, the responsible fieldworker .....	55
<i>Sándor Szeverényi</i>	
The life and work of the Saami theologian and linguist: Anders Porsanger .....	71
<i>Ivett Kelemen</i>	
The use and semantics of the Northern Mansi diminutive <i>-riś~rəś</i> .....	81
<i>Bernadett Bíró</i>	
The event of “giving” and “getting” in Siberian Uralic languages .....	99
<i>Katalin Sipőcz</i>	
A word-formational approach to neologisms in modern Northern Mansi .....	119
<i>Susanna Virtanen</i>	
Word and stem repetitions in the heroic epic songs collected by Antal Reguly .....	131
<i>Mária Sipos</i>	
The use of body part terms in expressing emotions in Udmurt .....	149
<i>Rebeka Kubitsch</i>	
The characteristics of responses given to compliments in Udmurt.....	173
<i>Zoltán Németh</i>	
On some Chuvash–Mari shared lexemes and Agyagási’s “Late Gorodets” hypothesis ..	185
<i>Christopher Culver</i>	

“Сувениры Севера” Minority identity and discourse. Representation of indigenous minorities of Northern Russia in the digital media. The case of Dudinka .....	201
<i>Zsuzsa Várnai and Ágnes Hámori</i>	
Reconsidering the Nganasan vowel system .....	229
<i>László Fejes</i>	
New aspects in the study of Mari, Udmurt, and Komi-Permyak: The <i>Typological Database of the Volga Area Finno-Ugric Languages</i> .....	255
<i>Erika Asztalos, Nikolett F. Gulyás, Laura Horváth, and Bogáta Timár</i>	
Ethnosyntax in Siberian Uralic Languages (a project report) .....	275
<i>Bernadett Bíró, Katalin Sipőcz, and Sándor Szeverényi</i>	

# On some Chuvash–Mari shared lexemes and Agyagási’s “Late Gorodets” hypothesis

Christopher Culver

## 1. Introduction

The Chuvash and Mari languages have been in contact for centuries and during that time loanwords have flowed in both directions, Chuvash–Mari and Mari–Chuvash (see e.g. Räsänen 1920; Fedotov 1990). Some of these shared lexemes can be traced to the Proto-Uralic language ancestral to Mari and were borrowed from Mari into Chuvash. Others were inherited by Chuvash from Proto-Turkic and then transmitted to Mari. Still others can be ultimately derived from Russian or Persian and were mediated through Chuvash to Mari.

There remain, however, a number of shared lexemes in Mari and Chuvash that have lacked a clear etymology. In a series of publications, Agyagási (2000, 2001, 2002) has posited that Mari and Chuvash were both in contact with an unknown third language, spoken by a people that was assimilated by the Mari and Chuvash after the large-scale population movements following the Mongol invasion in the 13th century. Agyagási initially identified this unknown language with the Late Gorodets archaeological population. Furthermore, she suggests that the exonym “Cheremis” – never used by the Mari themselves – originally referred to the people of the Late Gorodets archaeological population, and the ethnonym came to refer to the Mari only after the Late Gorodets population disappeared as a distinct people through assimilation. In a more recent publication on historical interaction between Chuvash and Mari, Agyagási (2019) no longer points to the Late Gorodets archaeological population specifically and, therefore, ceases to employ this nomenclature, but she continues to believe that there was a mysterious third population in contact with both Mari and Chuvash, and now she identifies this language as mediating lexical material from the West Baltic branch of Indo-European to Mari.

Agyagási bases her hypothesis of a third language in the region on a corpus of sixty-four shared Mari–Chuvash words that lack an etymology (Agyagási 2000, hereinafter referred to as the “Late Gorodets” material as a convenient label for the collection in spite of Agyagási’s more recent turn away from ascribing it to the Late Gorodets archaeological population). This corpus includes such etymological cruxes

as Mari *merañ* ~ Cv. *mārin* ‘hare’. On one hand, Chuvash *mārin* cannot be an inherited Turkic word due to the initial *m*- here (on the impermissibility of initial \**m*- in Proto-Turkic, see Tenišev 1984: 310ff.). At the same time, Mari *merañ* is difficult to explain as an inherited lexeme in Mari due to the irregular vowel and consonant correspondences among the Mari dialects. Thus, as an example, in Agyagási’s 2000 publication compiling the “Late Gorodets” material, she brings together the data on Mari *merañ* ~ Cv. *mārin* as follows:

23. V(Morg.) *mārin*, *mārik* ‘заяц’ (Ašm. 8. 317); (Kozm.)  
*mo’ran* ‘Hase’ (Munk.); (ohne Ortsangabe) *māren* ‘заяц’ (Ašm.uo.);  
 vgl. mar. mund. 1775: *мерангъ* [*merang*] ‘заяц’ (Sebeok and Raun 30); P *mer’añ*, *merañ* ‘Hase’ (Genetz); KB *moren* (Ramstedt); P B UP  
 USj C Č *merañ*, M *merañ*, MK *mer’añ*, UJ *merañe*, CČ *mor’an*,  
 JT *mōran*, JO *mōraṅə*, V *mōrāṅə* K *moren* 1. ‘nyúl; 2. fátyú,  
 zabigyerek’ (Beke). ☒ Räsänen 1920. 255; Fedotov 1968. 204;  
 Arslanov and Isanbaev 1984. 110.

Then, in a subsequent publication Agyagási (2001: 36) attempts to explain the unusual vowel correspondences between the various dialectal forms by positing borrowing into the already distinct Mari dialects from two different dialects of this mysterious neighbor: “Western Late Gorodets” contributed a form, *mor’*, to Hill Mari, while “Eastern Late Gorodets” contributed a different form, *mer’*, to Meadow Mari.

In spite of its convenient compilation of Mari and Chuvash words demanding etymological clarification, Agyagási’s corpus of “Late Gorodets” lexical material has drawn almost no commentary from other scholars as of yet. The “Late Gorodets” corpus merits special attention now, when research into the Mari lexicon and Mari historical phonology has made great advances in the intervening years. First, in compiling her wordlist Agyagási relied overwhelmingly on Beke’s dictionary, which was published in 1997–2001, and at the time of Agyagási’s initial research it was the only dictionary of Mari that drew together material from across the bulk of the Mari dialects. However, scholars today have a fuller picture of the Mari lexicon due to the publication of *Tscheremissches Wörterbuch* (TschWb) in 2008, a compilation of material collected by a number of late-19th century and early-20th century fieldworkers, which represents a resource as ample and valuable as Beke’s, and also the dictionary of Mari dialects of Udmurtia and Tatarstan edited by Veršinín (2011). One can also mention the 18th and 19th century manuscript wordlists of Mari dialects described by Sergeev (2002), which occasionally provide dialectal forms or meanings that are absent from later sources. Finally, though it was published three decades before Agyagási’s work, the Northwestern Mari dictionary by Ivanov and Tužarov is

not cited in Agyagási’s compilation of the “Late Gorodets” material, and yet in some cases it provides unique data unavailable elsewhere.

Second, Agyagási’s view of Mari historical phonology hews to the reconstructions proposed by Berezcki. More recently, however, Aikio (2014a) has proposed a new reconstruction of Proto-Mari vocalism that adheres to strict sound laws instead of so readily admitting “sporadic” sound changes. Aikio’s reconstruction of Proto-Mari also grounds Mari in the latest innovations in the historical phonology of its parent Proto-Uralic protolanguage, as found in recent publications by Aikio and Zhivlov (see Aikio forthcoming for a survey and the relevant literature).

By drawing on the modern state of the art in Mari historical phonology, dialectology, and lexicology, the present contribution challenges Agyagási’s ascription of several shared Mari and Chuvash words to a late shared substrate. I aim to show that a number of these words cannot serve as proof of contact with an unknown third language in the early second millennium CE.<sup>1</sup>

## 2. Mari and Chuvash historical phonology and the “Late Gorodets” material

It goes without saying that a shared substrate is not the only possible source of unetymologized lexical material found in both Mari and Chuvash. That is, the lexicons of each of these languages stem from a number of sources. Mari is a descendant of Proto-Uralic, and, therefore, it has retained lexical material dating from as far back as the era when Proto-Uralic was spoken; work has continually progressed in identifying new Uralic etymologies for Mari words that were formerly considered to be of uncertain origin (see e.g. Aikio 2014b, Metsäranta 2020). Additionally, Mari has undergone contact with other languages since the Proto-Uralic era. Besides possible contacts between pre-Proto-Mari speakers and prehistoric non-Uralic languages of European Russia that we can only speculate about, scholars have identified in the Mari lexicon early borrowings from Baltic (Gordeev 1967), various stages of Iranian (Joki 1973, Gordeev 1967) and Permian (Berezcki 1977; 1987; 1992–1994: II 97–129; 2005, see also Berezcki et al. 2013). In more recent centuries, Mari has been in contact with Chuvash, Tatar (Isabaev 1989–1994), and Russian (Savatkova 1969).

By the same token, Chuvash is a descendant of Proto-Turkic, and among those unetymologized items in its lexicon there may be words hitherto unidentified for their Proto-Turkic origin. Additionally, Chuvash may have undergone contact with

---

<sup>1</sup> I would like to thank Tapani Salminen and an anonymous reviewer for their comments which allowed me to improve on this paper. Naturally, the blame for any remaining errors falls solely on me.

unknown languages after its split from Proto-Turkic and during West Old Turkic speakers' migration from Asia to the southern Russian steppes and up to the Middle Volga, and scholars have identified borrowings from Persian and Arabic (Schermer 1977), Alanic (Dobrodomov 1980) and Permian (Rédei and Róna-Tas 1982). More recently, Chuvash too has been in contact with Russian and Tatar (see Fedotov 1996).

The first criticism of the “Late Gorodets” material as proof of a substrate shared by Mari and Chuvash is that, on phonological grounds, certain words can be shown to represent lexical material that was already present at the Proto-Mari stage and was then borrowed by Chuvash, or vice versa. Even if a given word still lacks an ultimate etymology, then that word may represent a borrowing in Mari or Chuvash, but it could have been borrowed into either of these languages before the population movements of the 13th century. Such words cannot serve as evidence for an unknown language of the Middle Volga in contact with both Mari and Chuvash, but rather they could be the result of other language contact situations that we know Mari and Chuvash participated in.

To illustrate this point, in this paper Mari historical phonology is viewed according to the recent reconstruction by Aikio (2014a). Briefly, Aikio's reconstruction differs from the reconstruction by Berezcki on which Agyagási generally relied in the following ways.

Berezcki (1992–1994: I 115) reconstructed the following first-syllable vowel system for the parent language of the Mari dialects:

i	ü	u
e	ö	o
		a

Berezcki explains the reduced labial vowels of the Mari dialects as the result of post-Proto-Mari reduction of what were full labial vowels in Proto-Mari; the correspondence Meadow Mari *o* ~ Hill Mari *a* is seen as the result of lowering of original \**o* in Hill Mari; and Hill Mari *ä* is seen as the result of fronting of original \**a*.

Aikio proposes, instead, the following vowel inventory for Proto-Mari:

i	ü		u
ĩ	ũ		ũ
e			o
ä		a	ǎ

Thus, key differences in Aikio’s reconstruction are that both full and reduced labial vowels must be reconstructed for Proto-Mari, and there was an opposition between \*ǎ and \*o and between \*ä and \*a. The subsequent developments in the Mari dialects are broadly as follows:

- PMari \*u > MariE NW W *u*
- PMari \*ü > MariE NW W *ü*
- PMari \*ũ > MariE *u* NW *ũ* W *ê*
- PMari \*ǎ > MariE *ü* NW *ũ* W *ə*
- PMari \*ǎ > MariE *o* NW W *a*
- PMari \*o > MariE NW W *o*
- PMari \*ä > MariE *a* NW W *ä*
- PMari \*a > MariE *a* NW W *a*

In the appendix to her recent monograph on Chuvash historical phonology, Agyagási (2019: 289ff) rejects Aikio’s reconstruction. Though Agyagási accepts \*ə as a member of the Proto-Mari vowel system, she attempts to revive the Bereczki-era reconstruction in terms of the high rounded vowels. She claims that Proto-Mari had solely full high labial vowels, and reduced labial vowels only arose later in Northwestern Mari, Hill Mari, and in the Upša subdialect of Meadow Mari. She goes on to propose various conditioning environments for this vowel reduction.

Nevertheless, Agyagási’s arguments against Aikio’s reconstruction are unsound. First, her survey of reduced labial vowels in the Mari dialects fails to take into account data from the Bol’saja Šija and Menzelinsk dialects in the Eastern Mari area as documented by Veršinín (2011); these dialects possess both full and reduced labial vowels (and broadly in the places we would expect from Aikio’s reconstruction) but they differ in other aspects from the Northwestern Mari–Hill Mari–Upša complex. Second, it is easy to find exceptions to the conditioning environments which Agyagási proposes. For example, although Agyagási (2019: 297) claims that the first-syllable vowel in Proto-Mari \**ü-ê* structures undergoes reduction, this is never the case with

MariE W NW *šüðö* ‘100’, which shows a full vowel across the Mari dialects. This in fact relates to one of the key insights afforded by Aikio’s 2014 article: among the Mari lexical material inherited from Proto-Uralic, reduced and non-reduced labial vowels go back to different PU vowels. The high front labial reflex of PU \*j does not appear as a reduced vowel anywhere across the Mari dialects, cf. PU \*šjta ‘100’ > MariE W NW *šüðö*. Additionally, in the Mari inherited lexicon we find a different treatment of high front labial vowels before *r*, i.e. in one set of words the reflex among the Mari dialects is *ö*, e.g. MariE W *mör* ‘strawberry’ from PU \*mürja, while other words do not show this lowering, e.g. MariE *βür* NW *βür* W *βär* ‘blood’ < PU \*wire (UEW 264–265, 576). Thus, the distinction between the two types of labial vowels represents an inherited feature in Mari that was already present when Mari and Chuvash came into contact.

There is wider consensus around the Chuvash historical vowel inventory than the Mari one (see Agyagási 2019), though as Savelyev (2018) has recently noted, the Northwestern dialect of Chuvash has been given insufficient attention and some of its distinctive sound shifts are relevant for detecting loans between Mari and Chuvash.

With a firm grounding in Chuvash and Mari historical phonology, evidence that the “Late Gorodets” words must represent material extraneous to Mari and Chuvash evaporates. Many of these words do not show irregularities between the Mari dialects at the post-Proto-Mari period, rather they can be traced back to Proto-Mari. One can then account for the data listed under many of the “Late Gorodets” items through simple borrowing of inherited Mari material into Chuvash.

As an example, consider No. 52 in the “Late Gorodets” material: MariW *βažar* ‘cross-grained (of wood)’ versus Cv. A *vušar* id. These words both regularly go back to Proto-Mari \*βāžar and Middle Chuvash \*vāšar, respectively. As their vocalism is identical, one word can simply represent a borrowing from the other, presumably from Mari into Chuvash as there is no regular source for Chuvash *š* in this environment.

A similar situation holds for MariE *mužâr* NW *müžür* W *mâžâr* ‘pair; spouse’ versus Cv. V *möšör* A *măšâr* id. The Mari forms regularly go back to PMari \*müžær,<sup>2</sup> while for the Chuvash forms we can reconstruct Middle Chuvash \*möšör. The

<sup>2</sup> Agyagási cites a Malmyž dialect form “mužün” from Beke’s dictionary, but this is in fact a typo on her part: we find expected *mužur* in Beke. She also cites a supposed Hill Mari form “mužor” from Budenz’s dictionary, but Budenz does not actually ascribe the word to Hill Mari, and, on the basis of every other Mari lexical reference, we should assume that Budenz documents a Meadow Mari form. With regard to the second-syllable vowel of Budenz’s “mužor”, Budenz uses a pre-modern transcription that often fails to correctly reflect Mari’s reduced vowels, and so the word should be read as the same form *mužâr/müžür* known from other Meadow Mari lexical resources.

protoforms in both languages are essentially identical, and therefore borrowing from Mari into Chuvash or vice versa is the most straightforward explanation. Again, this is probably a borrowing from Mari into Chuvash on the basis of Cv. *š* in this environment.

Let us next turn to the case of the PMari \*a and \*ä. Proto-Uralic \*ä and \*a were both raised in pre-Proto-Mari, and consequently any Mari word with a first-syllable vowel going back to Proto-Mari \*ä or \*a must be a loanword instead of inherited Uralic material. The majority of these words are known to be relatively recent borrowings from Chuvash. However, as noted by Bereczki (2005: 193), we find \*a and \*ä even in Iranian borrowings in Mari that are presumably much older than the 13th century – an example is the Mari ethnonym \*mari itself (Joki 1973). PMari \*ä is also found in the earliest layer of loanwords from Permian, cf. MariE *jakte* NW W *jäfte* ‘pine tree’ < Komi *jag* ‘Kiefernwald (auf Sandboden)’, Udm. *jag*, *lag*, *d’ag*. ‘Fichtenwald auf Sandboden’ (Bereczki 1992–1994: II 98).

Consequently, a word common to Mari and Chuvash that shows PMari \*ä or \*a and has no Chuvash etymology need not be a borrowing into both Mari and Chuvash from a third language in which they were both in contact beginning from the 13th century. Consider the “Late Gorodets” item consisting of MariE *kače* NW *käčə* ‘young man’ versus Cv. A *kaččä* id. (No. 5). The geminate consonant in the Chuvash word is a feature generally found in loanwords, including those from Mari, cf. Cv. V *merčče-* ‘чахнуть’ < MariE *merčem* W *mercem* id. (Fedotov 1990: 309). The dialectal forms underlying the Mari counterpart to Cv. *kaččä* regularly go back to PMari \*käčə. Therefore, it cannot be excluded that the Mari word is a pre-13th-century borrowing from some unknown source,<sup>3</sup> which was then borrowed by Chuvash.

For words where the Chuvash material is known only from the northwestern Chuvash area, where contacts with Mari had been maintained until recent centuries, we must also take into account distinctive Hill Mari sound shifts. Thus, in the case of MariE *nol’*, *nol’o* ‘sediment’ and W *nal’ə* ‘red clay’ versus Cv. V *nal’ə/nay* ‘тина’ (No. 25), we find a word that can be reconstructed as PMari \*näl’ə, but in Chuvash the word is attested only in dialects in the northwest of the Chuvash-speaking area, in close proximity to Hill Mari. Consequently, the most parsimonious explanation is that Viryal Chuvash borrowed the word subsequent to the shift of PMari \*ä > Hill Mari *a*.

A similar objection can be made to the ascription of MariE *tuškem* NW *tüşkem* W *tăškem* ‘смешивать, перемешивать’ ~ Cv. *tăška* id. (No. 44) to a shared substrate. The Mari words regularly go back to PMari \*tüşkem. While the dialectal source of

<sup>3</sup> Gordeev (1967: 196) in fact adduces Iranian parallels.

Ašmarin's word is unclear,<sup>4</sup> the most parsimonious explanation is that this represents material from the Viryal dialect, and the back unrounded reduced vowel is simply due to Viryal Chuvash borrowing the word from Hill Mari subsequent to the delabialization of PMari \*ũ > MariW â. (Again, \*š in this position would be normal for inherited Proto-Mari vocabulary, but irregular in Chuvash.)

Finally, let us consider the case of Cv. *páčă, păčkata* ~ MariE *pič* NW W *pəc* 'stuffy; dark' (No. 29–30). These items were first included in the "Late Gorodets" corpus. Then, in a more recent work Agyagási (2019: 280–281) has argued that the Mari and Chuvash words are not related as such, but nevertheless they are borrowings from the "Cheremis substrate" which supposedly transmitted West Baltic lexical material into both Chuvash and Mari. Thus, she traces Cv. *páčă* to a West Baltic dial. \*pũčio 'suffocate, stifle; blow up', while Mari *pič* is ultimately a derivation of a West Baltic verb \*pũ 'blow up' with a nominal derivational suffix -č.<sup>5</sup>

Agyagási's argumentation that these words must have been borrowed independently into Mari and Chuvash from a third, substrate language is flawed on phonetic grounds. First, she reconstructs the Proto-Mari form for the Mari material as \*pič. In fact, on the basis of the correspondence MariE č̣ NW W c, the Mari word originally featured the consonant \*č and not \*č̣, and therefore we must reconstruct PMari \*pəč̣.

PMari \*č̣ has a regular historical source. On the one hand, it represents PU \*č in word-final or pre-consonantal position (intervocalically pre-PMari \*-č̣- > PMari \*-č̣-). On the other hand, it can appear as the result of a shift PU \*t > PMari \*č̣ after PU \*i, cf. Mari *βič̣* '5' < PU \*witti (UEW 557). Indeed, the vowel and consonant correspondences of Mari *pič̣* are completely identical across the Mari dialects to those of Mari *βič̣* '5', and therefore Mari *pič̣* need not be any kind of late loan from an unknown substrate language, rather it may well go all the way back to Proto-Uralic.

On the Chuvash side, by contrast, the č̣ in Cv. *păčkata* has no regular source; Cv. č̣ results regularly only from the affricatization of \*t before \*i. Cv. č̣ would, however, be expected as a reflex of PMari \*č̣ in borrowings from Mari (e.g. Cv. *merčče* 'чакнуть' mentioned above). Consequently, here too the most parsimonious explanation is that Chuvash borrowed material from Mari that had already existed in the latter language since some unknown era.

The adduced examples above are likely loans from Mari into Chuvash. However, among the "Late Gorodets" material there are also words where borrowing from one

<sup>4</sup> Ašmarin gives the source as "CIIBB. BA", but this abbreviation is not found in the list of abbreviations at the end of the dictionary and has apparently never been deciphered by researchers of Chuvash.

<sup>5</sup> This solution is extremely *ad hoc*, as no derivational suffix -č̣ is known.

of these languages into the other can be supposed, but the direction of borrowing is unclear. In the comparison in No. 27, Chuvash V *ora* A *ura* ‘haystack’ regularly goes back to Middle Chuvash \**âra*, while its Mari counterpart E *ora* NW W *ara* id. regularly goes back to PMari \**âra*. The words, therefore, show no irregularities for which we must posit a third language in contact with both Mari and Chuvash, but the task is left to future etymological research to determine if this represents inherited Chuvash material transmitted into Mari or inherited Mari material transmitted into Chuvash.

### 3. Etymologizable words

Besides the general objection that the “Late Gorodets” words may represent inherited, albeit still unetymologized material in Mari borrowed into Chuvash or vice versa, there are cases where more concrete etymologies can be proposed.

One such etymology has in fact already been published. The inclusion of MariW *pâjârka*, *pâjârt* versus Cv. V *pâyârt*, *pât’ârt* ‘a little bit’ in the “Late Gorodets” material (No. 31) overlooks a convincing Russian loan etymology that had already been established decades earlier by Rédei and Róna-Tas (1983: 37). As an aside in their discussion of the mistaken comparison of Komi *parga* ‘in der Flachsheckel zurückgebliebener flockenförmiger, reiner Abfall vom gehechelten Flachs’ with Chuvash *parga* ‘Büschel’, Rédei and Róna-Tas identify this Mari and Chuvash material – as well as Tat. dial. *payarka* – as adoptions of the Russian *ноярок* ‘шерсть ягнят (первой стрижки)’ with the semantic development ‘small heap of wool’ → ‘small heap, bundle’.<sup>6</sup>

Another instance where a clear acceptable etymology can be proposed is the equation Cv. *kănăš* ‘rubbish’ ~ Mari *kunăž* id. (No. 7). The Chuvash forms here V *kônôš* A *kănăš* regularly go back to earlier \**kônôš*. On the Mari side, the correspondence MariE *u* ~ NW *ũ* ~ W *ê* is regular and goes back to PMari \**ũ* according to Aikio (2014a), and therefore Aikio traces the Mari word back to PMari \**künəž*.<sup>7</sup>

<sup>6</sup> See the entry for *ноярок* in Dal’s dialectal dictionary for the full range of semantics in Russian. When lambs are shorn for the first time, they produce a quite small amount of wool, and the example sentences that Beke’s dictionary gives for Mari *pajârka* suggest the word was mainly applied to small amounts of material (wool/straw/bast).

<sup>7</sup> Agyagási cites a Mari dialectal form *kunăš* from the *Черемусско-русский словарь* of V. M. Vasil’ev, published in Kazan in 1911. This form is also found in Szilasi’s dictionary (drawn in fact from Troitskij’s earlier dictionary) and ascribed to the Hill Mari variety. The most parsimonious explanation for this form is that Hill Mari reborrowed the word from Viryal

The vocalism of the reconstructed Proto-Mari and Middle Chuvash forms matches exactly, and borrowing in either direction could be posited on that basis without the need to involve a third, unknown language. Furthermore, as Alexander Savelyev (p.c.) points out, there is another definition of this word besides the ones that Agyagási cites. Namely, the dictionary of Sergeev glosses Cv. *kānāš* as ‘всякая всячина в избе, лишние предметы’, and we can also compare this word to Cv. *kān-kan* ‘там и сям торчащие в разные стороны предметы’. In Savelyev’s view, this appears to be none other than a derivation in *-āš* of the same root found in Cv. *kān-* ‘вытягиваться, разгибаться’, *kāntāran-* ‘торчать, вытягиваться’. The root *kān* in turn goes back to Proto-Turkic \**kön* (Fedotov 1996: I 254). Thus, an original root meaning ‘stick out’ was extended to refer to assorted junk around one’s home that got in one’s way, and then its meaning broadened to ‘rubbish; filth’ in general.

Savelyev’s proposed etymology is attractive, because an earlier sense ‘encumbrance’ is attested for Mari *kunǎž* as well: Beke’s dictionary attests MariNW *kūnǎžlanem* ‘im Wege sein’, with the Mari informant defining the word as *ere aβašt βelen pižāt, okoltep pašaš* ‘sie hängen fortwährend an der Mutter, lassen sie nicht arbeiten’.

Also, Savelyev’s explanation has predictive power for other Mari material that has hitherto been unetymologized. In his dictionary of Mari dialects of Udmurtia and Tatarstan, Veršinín attests MariE *kunčó* ‘надоедливый компаньон (тж. нежеланный и неуходящий, несмотря на намёки, гость); ребенок, ни на шаг не желающий отходить от своей матери’. Sergeev (2002: 177, 32–33) documents *кунзю* ‘воз’ from an 18th century manuscript prepared for the *Vocabularia comparativa* of Peter Simon Pallas that reflects mainly the Malmyž dialect, and here we are probably dealing with the same word. This Mari word can be derived from the same Cv. *kān*-root with a different derivational suffix *-čǎ*,<sup>8</sup> i.e. Middle Chuvash \**kǎnčǎ* was borrowed into Mari as \**kūnčǎ*, and this ultimately became expected *kunčó* in Eastern Mari. All of the above-mentioned words would, thus, be examples of the same polysemy found in Russian *торчать* ‘stick out; hang around, be in the way, be bothersome’.

With regard to Cv. *yapala* ~ Mari *jaβala* ‘thing’ (No. 55), an explanation for the Chuvash word was given by Ašmarin in the early 20th century in his massive Chuvash dictionary. In the entry for the Chuvash interrogative word *yepε* ‘как?’, Ašmarin

---

Chuvash – which preserved a labial vowel – at a time subsequent to the Hill Mari unrounding of Proto-Mari \**ǎ*.

<sup>8</sup> For a parallel, cf. Mari *olǎǎrčó* ‘скальница’ (part of a weaving loom) < Chuv. *xultǎrčǎ* id., which in turn is considered to be a derivation from a lost verb \**xultǎr-* ‘дрожать, трястись’ reconstructed on the basis of Common Turkic \**qaltir* (Fedotov 1996: II 356).

(IV 278) writes “отсюда происх[одит] слово йапала”. Viewing Cv. *yapala* ‘thing’ as a derivation from *yepe* ‘how?’ is reasonable inasmuch as the sheer bulk of usages of Cv. *yapala* documented by Ašmarin (IV 204–206) suggest that the word originated as a metasyntactic expression which was ultimately lexicalized, cf. English *whatsit* ‘thing’ < *what is it* and *whatchamacallit* ‘thing’ < *what you may call it* for parallels.

Ašmarin’s explanation for Cv. *yapala* was overlooked by the Chuvash etymological dictionaries of Egorov (1964) and Fedotov (1996), and therefore the word may have seemed more mysterious than it in fact is. Accepting the Chuvash etymology from *yepe* ‘how?’ allows viewing the Mari word as a borrowing from Chuvash; the Mari word is, after all, found only in the Volga dialect, which maintained contact with Chuvash as late as the 20th century. There are no phonological obstacles to borrowing from Chuvash into Mari in this case, though the chronology of the borrowing is ambiguous. The existence of forms *yepete* and *yapala* in Chuvash points to original front vocalism. Thus, the word must either have been borrowed prior to the Meadow Mari backing of PMari \*ä, i.e. Middle Chuvash \*yäpälä > Mari \*jäβälä (> modern Volga Mari *jaβala*), or instead the Chuvash word was borrowed into Mari subsequent to the backing of Chuvash \*ä > a, i.e. Cv. *yapala* > Volga Mari *jaβala*.

While the etymology of the Chuvash interrogative adverb *yepe* ‘how?’ (along with variants such as *leple* and *neple* with differing initial consonantism) remains unclear, again this Mari–Chuvash lexical parallel does not serve as proof of contacts between both Mari and Chuvash and a third language of the region, and an explanation should be sought within the separate history of Chuvash.

Finally, with regard to Mari *jokrok* ~ Cv. *yäkraka* ‘boring’ (No. 57), it should be noted that, across the Mari dialects, the meanings ‘empty; creepy; lonely’ are also attested for the Mari word. In his recent dialectal dictionary Veršinín (2011: 117) has documented the word *joŋrok* ‘один-одинёшенек’ from the Verkhñjaja Iž-Bob’ja dialect of Eastern Mari, and we can assume that this represents the same word due to the matching semantics. Consequently, we can reconstruct PMari \*joŋrok ‘boring; lonely’, with the bulk of Mari dialects undergoing cluster simplification to *jokrok*.<sup>9</sup> This form, *jokrok*, with *k* instead of earlier \*ŋ would then have been borrowed from Mari into Chuvash; the opposite direction of borrowing is not possible, as Chuvash shows a different resolution of the *-ŋr-* cluster, cf. Cv. *tură* ‘god’ < Proto-Turkic \*teŋri (Fedotov 1996: II 252–253). The variant Hill Mari forms *jăk(ă)rikă* and *jăk(ă)raka* can be explained through reborrowing from Chuvash.

<sup>9</sup> Clusters involving *ŋ* followed by a consonant have been unstable in the history of Mari, cf. for example MariE *umdo* W *əŋgəđđə* ‘spear’ < PMari \*ŋŋđđ.

#### 4. Onomatopoeia/sound symbolism

As a third objection, some of the items in the “Late Gorodets” corpus can be shown to be of an onomatopoeic nature or cases of sound symbolism, and in such words irregular phonetic correspondences are commonplace and expected. On the basis of solely such irregular phonetic correspondences, these words do not serve as proof of borrowing from a third language.

An example of this is Agyagási’s comparison (as No. 17) of Cv. V *lõncõ, lëncë* ‘слабый, нетуго натянутый’ with MariE *lančâra* NW *lânzâra* W *lanzâra* ‘schwach, weich (Mensch, Pferd); dünngewebt (Stoff); zerfetzt (Kleid)’. If we draw on the full breadth of Mari lexical resources cited herein, we find the same root ‘soft’ in a highly varied number of forms: MariE *lančâra, luñčâra, lazâra, lãnčë-lañčë* NW *lânzärä, lünzâra* W *l’änzärä, l’äzärä, lanzâra, lânzâra*. Confronted by this wild difference in vocalism and the presence or absence of the nasal, it would be far more straightforward to consider this a sound-symbolism root where either \*ũ, \*ä or \*ə could be plugged into the consonantal skeleton \*IV(n)ć-, than to consider this a post-1200 CE borrowing into both Mari and Chuvash. Agyagási’s attempt to argue for contact with dialects of an unknown third language by divergent forms from one Mari dialect to another (see her treatment of Mari *merañ* above) breaks down when we consider that even within the same dialect we find a variety. For example, speakers from the Bol’šoj Kil’mez dialect provided both *lančâra* and *lazâra*. The word in Chuvash must be a loanword owing to the initial *l-*, but the word can be explained as a borrowing from Mari, and the above-cited Mari forms with a labial vowel serve to explain the back labial reduced vowel in Viryal Chuvash *lõncõ*.

Similarly, the comparison Cv. *käsäya* ~ Mari *kisa* ‘сница’ (No. 10) is found in the “Late Gorodets” corpus. While the forms attested in the Mari and Chuvash dialects show irregular correspondences, the tit bird has a distinctive call that lends itself well to onomatopoeia.

#### 5. Conclusions

The objections raised in the course of this work ultimately affect approximately 45 items of the 64 in the “Late Gorodets” wordlist. Some of the Mari words may well go back to Proto-Uralic, inasmuch as phonologically nothing speaks against tracing them to a Proto-Uralic ancestor. In other cases, we might seek an origin in post-Proto-Uralic contacts with other languages of the region. The Chuvash material in most cases can be explained as loans from Mari. This in fact represents a return to the *status quo ante*, as Lukojanov (1974) and Fedotov (1990) prior to Agyagási’s “Late Gorodets” hypothesis were inclined to treat many of these words as just like any other Mari loans into Chuvash.

Consequently, for those remaining items where irregularities do appear among the Mari dialectal forms, other explanations should be sought. With regard to Mari *merañ* ‘hare’, we might revive the suggestion by Räsänen (1933: 360) that this word is a compound word consisting of Mari \*mü ‘earth’ (reconstructed as the root of attested *mlande* id.) and perhaps a lost Viryal Chuvash noun \*xorañ that can be reconstructed on the basis of cognates in other Turkic languages, cf. Tatar *kujan*, Altai *kojan*, etc., and consequently the Mari dialectal forms are the result of differing processes of vowel contraction in the post-Proto-Mari period.

While the exact identity of the Cheremis people attested in historical sources remains an outstanding problem of the history of the Volga–Kama region, evidence is lacking that the material collected in the “Late Gorodets” corpus must necessarily be traced to the language spoken by this people instead of to other possible sources for Mari and Chuvash vocabulary.

## Abbreviations

- MariE Eastern and Meadow Mari
- W Western (Hill) Mari
- NW Northwestern Mari
- Cv. Chuvash
- A Anatri dialect of Chuvash
- V Viryal dialect of Chuvash
- Tat. Tatar

## References

- Aikio, Ante 2014a. On the reconstruction of Proto-Mari vocalism. *Journal of Language Relationship*, 11: 125–157.
- Aikio, Ante 2014b. Studies in Uralic etymology III: Mari etymologies. *Linguistica Uralica*, 2014(2): 81–93.
- Aikio, Ante. forthcoming. Proto-Uralic. In: Marianne Bakró-Nagy, Johanna Laakso, and Elena Skribnik (eds.), *The Oxford guide to the Uralic languages*. Oxford: Oxford University Press.
- Agyagási, Klára 2000. Der sprachliche Nachlaß der Spät-Gorodec Bevölkerung in den tschuwaschischen und mariischen Mundarten. *Folia Uralica Debrenciensia*, 7: 3–26.
- Agyagási, Klára 2001a. Die Spuren der Sprache der Spät-Gorodec Bevölkerung in den tschuwaschischen und mariischen Mundarten. *Congressus nonus internationalis Fenno-ugristarum Pars IV*. 35–39.

- Agyagási, Klára 2001b. Nyelvi kapcsolatok történeti típusainak vizsgálata a csuvas–cseremis nyelvviszonyban [An examination of the types of language contact in relation to Chuvash and Cheremis]. *Folia Uralica Debreceniensis* 8: 27–33.
- Agyagási, Klára 2002. Название черемис в средневековых источниках. In: И. В. Тараканов (ed.), *Пермистика-9. Вопросы пермской и финно-угорской филологии*. Ижевск: Изд. дом “Удм. ун-т”, 102–109.
- 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*. Wiesbaden: Harrassowitz.
- Ašmarin = Ашмарин, Н. И. 1928–1950. *Словарь чувашского языка 1–17*. [Dictionary of the Chuvash language] Казань & Чебоксары: Издание Народного комиссариата по Просвещению Чувашской Автономной Советской Социалистической Республики и Чувашского Государственного Университета.
- Beke, Ödön 1997–2001: *Mari nyelvjárás szótár (Tschermisssches Dialektwörterbuch) 1–9*. Bibliotheca ceremissica 4. Szombathely: Berzsényi Dániel Tanárképző Főiskola.
- Bereczki, Gábor 1977. Permi–cseremis lexikális kölcsönzések [Permian–Cheremis lexical borrowings]. *Nyelvtudományi Közlemények*, 79: 57–77.
- Bereczki, Gábor 1987. Пермско-марийские лексические связи: Сущность, развитие и функции языка. Москва: Наука, 112–115.
- Bereczki, Gábor 1992–1994: Grundzüge der tscheremissischen Sprachgeschichte I–II. Debrecen: Universitas Szegediensis de Attila József nominata.
- Bereczki, Gábor 2003. Adalékok a permi–cseremis nyelvi kapcsolatokhoz. In: Zoltán Molnár and Gábor Zaicz (eds.), *Permistica et Uralica: Köszöntő könyv Csúcs Sándor tiszteletére* [Permistica et Uralica: Festschrift for Sándor Csúcs]. Fenno-Ugrica Pázmániensia 1. Piliscsaba: Pázmány Péter Katolikus Egyetem, 39–41.
- Bereczki, Gábor 2004. Újabb adalékok a permi–cseremis lexikális kölcsönhatáshoz [New data on Permian–Cheremis lexical contact effects]. In: Eszter Várady and Márta Csepregi (eds.), *Permieik, finnek, magyarok: Írások Szíj Enikő 60. születésnapjára* [Permians, Finns, and Hungarians: Festschrift for Enikő Szíj on the occasion of her 60th birthday]. Budapest: ELTE BTK Finnugor Tanszék, 14–17.
- Bereczki 2005 = Березки, Габор 2005. Пермско-марийские лексические совпадения – заимствования или общие субстратные элементы? *Linguistica Uralica*, 41(3): 187–200.

- Bereczki, Gábor et al. 2013. *Etymologisches Wörterbuch des Tscheremissischen (Mari). Der einheimische Wortschatz. Nach dem Tode des Verfassers herausgegeben von Klára Agyagási und Eberhard Winkler.* Veröffentlichungen der Societas Uralo-Altaica 86. Wiesbaden: Harrassowitz.
- Dobrodomov = Добродомов, И. Г. 1980. *О некоторых гиппологизмах и созвучных словах: Из аланского пласта иранских заимствований чувашского языка.* In: *Проблемы исторической лексикологии чувашского языка.* Труды ЧНИИ 97. Чебоксары: Научно-исследовательский институт языка, литературы, истории и экономики при совете министров чувашской АССР, 21–29.
- Fedotov 1990 = Федотов, М. Р. 1990. *Чувашско-марийские языковые взаимосвязи.* Саранск: Издательство саратовского университета, саранский филиал.
- Fedotov 1996 = Федотов М. Р. 1996. *Этимологический словарь чувашского языка I–II.* Чебоксары: Чувашский государственный институт гуманитарных наук.
- Gordeev 1967 = Гордеев Ф. И. 1967. *Балтийские и иранские заимствования в марийском языке.* In: *Происхождение марийского народа.* Йошкар-Ола: Марийское книжное издательство, 180–202.
- Isanbaev = Исанбаев, Н. И. 1989–1994. *Марийско-тюркские языковые контакты I–II.* Йошкар-Ола: Научно-исследовательский институт языка, литературы и истории им. В. М. Васильева.
- Ivanov and Tužarov = Иванов, И. Г. and Тужаров, Г. М. 1971. *Словарь северо-западное наречия марийского языка.* Йошкар-Ола: Марийский научно-исследовательский институт.
- Joki, Aulis 1973. Uralier und Indogermanen. Die älteren Berührungen zwischen den uralischen und indogermanischen Sprachen. Mémoires de la Société Finno-ougrienne 151. Helsinki: Suomalais-Ugrilainen Seura.
- Lukojanov = Лукоянов, Г. В. 1974. *Марийские заимствования в чувашском языке.* Чебоксары: Чувашское книжное издательство.
- Metsäranta, Niklas 2020: *Läpiä päähän – huomioita itämerensuomalaisesta läpi-sanueesta.* In: Sampsa Holopainen et al. (eds.), *Scripta miscellanea in honorem Ulla-Maija Forsberg.* Mémoires de la Société Finno-ougrienne 275. Helsinki: Suomalais-Ugrilainen Seura, 228–232.
- Räsänen, Martti 1920. *Die tschuwassischen Lehnwörter im Tscheremissischen.* Mémoires de la Société Finno-ougrienne 48. Helsinki: Société Finno-ougrienne.

- Räsänen, Martti 1933. *Tscheremissische u.a. Etymologien. – Liber semisaecularis Societatis Fenno-Ugricae*. Mémoires de la Société Finno-ougrienne 67. Helsinki: Suomalais-Ugrilainen Seura.
- Rédei, K. & Róna-Tas, A. 1982. Proto-Permian and Votyak loan-words in Chuvash. In: András Róna-Tas (ed.), *Studies in Chuvash etymology I*. Studia Uralo-Altaica 17. Szeged: Universitas szegediensis de József Attila nominata, 158–179.
- Rédei, Károly and Róna-Tas, András 1983. Early Bulgarian loanwords in the Permian languages. *Acta Orientalia Academiae Scientiarum Hung.*, 37(1–3): 3–41
- Savatkova = Саваткова А. А. 1969. *Русские заимствования в марийском языке*. Йошкар-Ола: Марийское книжное издательство.
- Savelyev 2018 = Савельев, А. В. 2018. К уточнению сценария чувашско-марийских контактов. In: А. М. Иванова & Э. В. Фомин (eds.) *Языковые контакты народов Поволжья и Урала. Сборник статей XI Международного симпозиума*. Чебоксары: Чувашский государственный университет имени И. Н. Ульянова, 95–104.
- Scherner, Bernd 1977. *Arabische und neupersische Lehnwörter im Tschuwaschischen*. Wiesbaden: Franz Steiner Verlag.
- Sergeev = Сергеев, О. А. 2002. *Истоки марийской письменности*. Йошкар-Ола: Марийское книжное издательство.
- Tenišev = Тенишев Э. Р. (ed.) 1984. *Сравнительно-историческая грамматика тюркских языков: Фонетика*. Москва: Наука.
- TschWb = Arto Moisio and Sirkka Saarinen (eds.) 2008. *Tscheremissisches Wörterbuch*. Aufgezeichnet von Volmari Porkka, Arvid Genetz, Yrjö Wichmann, Martti Räsänen, T. E. Uotila und Erkki Itkonen. Lexica Societatis Fenno-Ugricae 32. Helsinki: Suomalais-ugrilainen Seura – Kotimaisten kielten tutkimuskeskus.
- UEW = Rédei, Károly 1988–1991. *Uralisches etymologisches Wörterbuch*. Budapest: Akadémiai Kiadó.
- Veršinín 2011 = Вершинин, В. И. 2011. *Словарь марийских говоров Татарстана и Удмуртии*. Йошкар-Ола: Научно-исследовательский институт языка, литературы и истории им. В. М. Васильева.