ONE OF THE EARLIEST MAJOR APPEARANCES OF THE MARI LANGUAGE IN PRINT IS THE "LINGUARUM TOTIUS ORBIS VOCABULARIA COMPARATIVA" EDITED BY PETER SIMON PALLAS AND PUBLISHED IN TWO VOLUMES IN ST. PETERSBURG IN 1787 AND 1789. THIS COMPARATIVE WORDLIST CONSISTS OF TRANSLATIONS OF 273 RUSSIAN HEADWORDS INTO A LARGE NUMBER OF LANGUAGES. AS WITH OTHER LANGUAGES, THE MARI PORTION OF PALLAS'S AMBITIOUS WORK WAS DRAWN FROM A NUMBER OF MANUSCRIPT WORDLISTS COMPILED AT THE ORDER OF CATHERINE THE GREAT.

THE MARI ENTRIES IN PALLAS WERE EXTENSIVELY COMMENTED UPON BY SEBEOK (1960) WHO DESCRIBED THE ORIGIN AND CONTEXT OF PALLAS'S WORDLIST, PROVIDED AN ITEM-BY-ITEM ANALYSIS, SUMMARIZED THE DERIVATIONAL MORPHOLOGY AND STRESS PATTERNS, NOTED COMMON SOUND CORRESPONDENCES BETWEEN MARI DIALECTAL VARIANTS, AND PROVIDED AN ENGLISH INDEX. AS A LEXICAL REFERENCE, SEBEOK RELIED MAINLY ON MÓRIC SZILASI'S "CSTEREMISZ SZÓTÁR" OF 1901. THE ABSENCE OF CERTAIN ITEMS IN SZILASI LED SEBEOK TO LEAVE SEVERAL ITEMS FROM PALLAS UNCOMMENTED, DENOTING THEM ONLY WITH THE DESIGNATION "UNATTESTED".

DUE TO THESE LACUNAE AND SOME ERRORS IN SEBEOK'S ANALYSIS, ALHONIEMI (1979) TOOK UP THE MATERIAL AGAIN, PRODUCING HIS OWN EXTENSIVE GERMAN-LANGUAGE COMMENTARY THAT ANALYZES EACH OF THE ITEMS, COMMENTS ON THE ORTHOGRAPHY USED, AND TRACES THE INFLECTIONAL ENDINGS ATTESTED IN THE WORDLIST DATA. ONE DISTINCT CONCERN OF ALHONIEMI'S COMMENTARY IS THAT HE AIMED TO DETERMINE THE SPECIFIC PHONOLOGICAL TRAITS OF THE ITEMS IN PALLAS.

ALHONIEMI'S COMMENTARY CORRECTS MOST OF SEBEOK'S ERRORS AND MANAGES TO IDENTIFY A LARGER NUMBER OF MARI WORDS. BesideS SIMPLY DRAWING ON...
more lexical references than his predecessor, one technique by which Alhoniemi identifies items that Sebeok missed is that he takes into account the possibility of misunderstandings between the Russian-speaking wordlist compiler and the Mari informant. Thus, he recognizes e.g. Pallas’s ложош ‘муха’ as MariE ložaš ‘flour’, cf. Russian мука ‘flour’, муха ‘fly’ (an understandable mistake as /x/ is not a phoneme in Meadow Mari). Nevertheless, although Alhoniemi’s commentary is an improvement on Sebeok’s, he too was unable to identify certain items.

In the decades since, new information has come to light that calls for taking a fresh look at Pallas’s wordlist in order to fill in some of the gaps of previous studies. New Mari lexical resources have appeared, especially the "Tscheremissisches Wörterbuch" edited by Moisio and Saarinen (2008, hitherto referred to as TschWb) and the Mari-English Dictionary. Furthermore, Ceprees (2000; 2002) has examined many 18th-century manuscript wordlists held in Russian archives, including several commissioned for Pallas’s project. He has set out the principles by which Mari was usually represented in Russian orthography by early wordlist compilers, he notes cases of erroneous transmission, and he even identifies the geographical provenance of certain Mari word forms that would also appear in Pallas’s printed book.

Thus, the present study aims to evaluate hitherto unrecognized items in Pallas in the light of newer references, bringing us a more complete understanding of what Pallas tells us about the Mari language in the late 18th century. Although the two volumes of the "Linguarum Totius Orbis Vocabularia comparativa" have been republished in recent decades (Pallas 1977, a photographic reproduction of the copy at the Staats- und Universitätsbibliothek Hamburg), for this study I have used a higher-quality scan derived from the copy at the Taylor Institution Library, Oxford, and made freely available on the internet. Unless otherwise mentioned, all Mari dialectal material is derived from TschWb and material from the MariE literary language is derived from the Mari—English Dictionary.

In the headings below, I give the Mari item from Pallas’s "Vocabularia comparativa", the Russian headword under which it is found, and the number of this headword.

Ирл ‘бол’ (61)

Sebeok marks this item as unclear but notes Szilasi’s jorla ‘poor’. The Cyrillic representation speaks against a round vowel, however. Alhoniemi (1979 : 212) is unable to identify it with any specific Mari word and only compares it to MariE ñrl̪ðwan ‘Masern’.

If one accepts the possibility of a misunderstanding between informant and compiler such that a verb was elicited instead of a noun, Pallas’s item can be identified as the 3 sg. pres. form of MariE ñrel ‘(Hund) knurren, (Mensch) wütend werden, murren’. Although Cyrillic u does not typically denote the vowel ñ (Ceprees 2002 : 102ff.), the Cyrillic representation does

1 The Mari-English Dictionary, located at http://dict.mari-language.com/, is an electronic dictionary of literary Meadow Mari, developed by the Department of Finno-Ugric Studies at the University of Vienna, which incorporates a number of previously published lexical sources as well as thousands of additional headwords.

2 The two volumes can be downloaded at https://commons.wikimedia.org/wiki/File:Linguarum_totius_orbis_1.pdf and https://commons.wikimedia.org/wiki/File:Linguarum_totius_orbis_2.pdf respectively.
match a form of this word attested from the dialect of Bolşoj Kilmez in which first-syllable i appears instead: irla.

Шўйдаўия 'власть' (66)

Sebeok (1960 : 332) correctly identifies the second element as Mari βуй 'head' but only speculates that the first element is šii 'neck', which is not acceptable. Alhoniemi (1979 : 227) only marks the item with a question mark.

The first sequence yи should be interpreted as an attempt to represent the front rounded high vowel ü in Mari. The placement of the stress tells us that the vowel denoted by Cyrillic a is likely the reduced vowel ё. Consequently, Pallas’s item can be identified with MariE lit. шўдувый 'сотник (lieutenant in Cossack army)'. Formed from šii '100' and βуй 'head', this Mari compound is a partial calque of Russian сотник < сто '100' + derivational suffix -ник. That this word known today from the MariE literary language can indeed traced back to 18th-century Mari is confirmed by the entry шудуўий 'сотник' in an 18th-century manuscript forming part of Damaskin’s dictionary that has been examined by Cepreev (2000 : 98; 2002 : 42). The difference in meaning between Russian власть and the Mari word can be explained as a misunderstanding where the wordlist compiler’s request for a noun 'power, authority' was answered with a term for a particular person holding power or authority.

Чюмышта 'ростъ' (69);

Alhoniemi (1979 : 210), who translates the Russian headword as 'Gestalt', was unable to identify the Mari word. It is unclear why he did not accept the interpretation of Sebeok (1960 : 299), who identified this as a derivation чүм-өст-а 'growth (lit. he stretches)'. While a derivation containing -өст- is not known from other sources, evidence for MariE čтем, W šтаем in TschWb supports Sebeok’s interpretation. The attested meanings 'spannen (Kleidungsstück, Leder), dehnen, vergrößern, weiter machen (Stiefel, Hut)' match the Russian headword. Furthermore, the Cyrillic letter ә is a common representation of the vowel ё, which is found in MariE (Upşa) čтем, and the vowel а as found in MariW šтаем (see Cepreev 2002 : 102—110).

Шитешь 'ростъ' (69)

Sebeok (1960 : 320) accepts this word as it is, but Alhoniemi (1979 : 226) only lists it with a question mark. This can be identified as a deverbal noun in -өст from MariE šтем W Nw šěтә '(aus)keimen, hervorkeimen (Getreide)', probably from the Western or Northwestern Mari areas due to the use of Cyrillic letters denoting front vowels and the final soft sign. Such a derivational form is attested in MariE lit. иштыш 'shoot, sprout'.

Чумратырмышь 'шаръ' (74)

Sebeok (1960 : 299) transliterates Pallas’s Cyrillic item as ҹumраtәrmәш and interprets it as consisting of two combined words, of which the first reflects MariE ҹумраš 'rund' and the second "perhaps a root tertэ- 'round'".

Alhoniemi (1979 : 210) maintains the same interpretation of this item as a combination of two words, and he also identifies the first component with MariE ҹумраš, but he provides no explanation for the second part. However, Alhoniemi misreads the item from Pallas, mistakenly citing it as
'Чумрăшьрмешьь'. Alhoniemi evidently had access to only a low-resolution reproduction of Pallas’s work, but from the high-resolution scan consulted for the present study, it is clear that the word is actually 'Чумрăшьрмешьь' as in Sebeok’s commentary, and Alhoniemi was led astray by the similar letter forms for ш and т in Pallas’s typeface.

After establishing that Pallas’s printed book presents a form тырмышь, we can further suppose a mistake during the typesetting phase by which a manuscript’s т was mistakenly replaced with м. Thus, we can identify this item with MariE тăртăш 'Ball, kugelförmig’, or possibly its MariW form тăртăш if the final soft sign represents frontness of the second-syllable vowel.

Пылăмьр ’буря’ (81)

Sebeok (1960 : 315) interpreted this as a compound, seeing the first part as MariE пăл ~ W пăл ‘sky’, but only conjecturing that the second part is related to MariE амăрĕшк ‘dirt’. Alhoniemi (1979 : 222) only marks it with a question mark.

This item can be identified with MariE пуламьр ‘Alarm, Störung, Lärm, Trubel, Schrecken’, attested from the Morki region in TschWb. This is not a compound involving ‘sky’ at all, but rather a borrowing from Tatar (cf. Tat. lit. болама ‘мешанина, путаница’). The word subsequently passed into the Meadow Mari literary language as пуламыр ‘agitation, rebellion, unrest, disorder, revolt, turmoil, commotion, stir, panic, etc.’.

Садиги ’паръ’ (111)

Alhoniemi (1979 : 224) only marks this with a question mark, but Sebeok (1960 : 317) was already on the right track when he wrote for this word, translated as ‘steam’ in Russian, ”probably a misunderstanding, for the word should mean ‘like that’”. Sebeok did not mention the specific Mari form he had in mind, but we are clearly dealing here with MariE сăдăγе. Of the five forms of this word in TschWb, all have stress marked on the initial syllable. For two dialects, stress is denoted on both the first and last syllables, but for the other three dialects (Birsk, Болшój Kilmez, Upšа) the stress falls solely on the first syllable, and the final vowel in the Birsk and Upšа dialects is partly reduced. The Cyrillic letter у is a common means of denoting a weak front vowel in 18th-century wordlists (see Ceprees 2002 : 106—108).

Муней ’колъ’ (129)

Sebeok (1960 : 311, 312) proposes that this is a variant muńej of MariE менье, W мăнкă, but no such variant is attested elsewhere. Alhoniemi (1979 : 220) marks the word with a question mark but also compares it to MariE мене.

Because of the overwhelming phonological similarity, this item should be identified with MariE мунăчи ‘Kröte’, namely the form мунеj attested from the dialect of Болшój Kilmez. As the two words мене and мунеj would have followed closely together in an alphabetically arranged manuscript wordlist, we can assume that at some point in the preparation of the material for press, the Russian translation of Mari ’stake’ was mistakenly associated with the Mari word for ‘toad’.
These items should be examined together. With regard to "возъ", Sebeok (1960: 309) marks this as unattested, but Alhoniemi (1979: 218) identifies the word as MariE (Birsk) kunžo 'Last'. The word is not found in TschWb, but it is present as kuńdžo in Paasonen’s dictionary of Eastern Mari, based on the Birsk dialect.

The other item, 'судно', is similarly marked as unattested by Sebeok (1960: 308), while Alhoniemi (1979: 218) only marks it with a question mark. Sinor (1961: 172—173) compared this item in Pallas to Mongolian günże 'radeau, canot' and similar items in the Tungusic languages, but as Middle Mongolian loans in Mari were typically mediated through Turkic (see e.g. Róna-Tas 1982), such a connection remains fanciful without Tatar or Chuvash evidence.

Сергеев (2002: 32—33, 177) has mentioned that a form куңзіо 'воз' is attested in a manuscript wordlist compiled for Pallas that shows clear traits of the Malmyž dialect. The item spelled Кунзя in Pallas’s printed book may be seen as another representation of the same word: the final я can be interpreted as the reduced vowel a but also marking palatalization of the preceding consonant, i.e. kunža. The difference in meaning between 'cart' and 'boat' is understandable, as besides the simple fact that a boat was also a common means of conveying a load in this region, the use of one and the same word for 'cart' and 'boat' is a feature of the neighboring and strongly influential Tatar language, namely Tat. көймә 'лодка, судно, кибитка'.

Чипталмаш 'брать' (185)

Sebeok (1959: 298) breaks this word down into the components čipt-al-m-aš, though he doesn’t comment on the root čipt- that he sees at the heart of the word. Alhoniemi (1979: 210), on the other hand, sees here a verb root plus the infinitive suffix -aš, but he does not identify the root.

The nomen actionis suffix -maš has been identified by both Sebeok and Alhoniemi in other items from Pallas, and the marked productivity of this suffix in 18th-century Mari records in general has been noted by Иванов (1975: 228) and Сепреев (2000: 84; 2002: 83, 141). This item in Pallas can be identified as such a derivational form stemming from a verb root Ėörtal-. Such a verb with this meaning is known from the MariE literary language: чипталаш 'to cover with a mat, to cover with a bast mat; (figuratively) to attack, to fall upon, to charge'. The Mari-English Dictionary views both senses as ultimately stemming from MariE lit. чыпта 'mat, bast mat, matting'.

TschWb attests Ėörta 'Bastmatte, locker gewebte Pferdedecke aus Bast' from several MariE dialects (Birsk, Sernur, Morki) and the first-syllable vowel is denoted, presumably under the influence of the initial Ė-, as fronted. Consequently, the use of Cyrillic u to represent the first-syllable vowel is to be expected.

Пилышь 'победа' (188)

Sebeok (1960: 315) mistook the Cyrillic letter н for б and thus incorrectly reads this as pilаš. He then marks this item as unattested. Alhoniemi (1979: 223) reads Pallas’s Cyrillic correctly as Пилышь, but he only lists the item with a question mark.
The Mari-English Dictionary offers the verb *пылнаш* 'to become frail, to become sickly, to grow weak; to be depressed, to be dispirited; to lose interest, to grow cold towards', which in form and meaning closely resembles another MariE lit. word *пылнаш* 'to come to ruin, to grow poor'. It was presumably the latter word in the MariE literary language that led the compilers of TschWb to list a headword *pîlînem*, but the sole attested form (from the Volga dialect of MariE) that TschWb offers under this headword is *pölnem*. Pallas’s item can be explained as a deverbal abstract noun in -*ş* from *pölnem*.

The semantic link between Pallas’s ‘victory’ and the meaning ‘become frail’ is highlighted by the existence of a derived transitive verb MariE lit. *пылныктараш* 'to stifle, to overwhelm'. That is, the victory of one side of a conflict is the weakening of the other side. As confirmation of this link, the same root is seen in *пылненъ* 'победа' in a late 18th-century manuscript wordlist kept in the state archives of the Kirov district (see Сергеев 2000: 39—42, 141).

### Шуръ ’китъ’ (198)

Sebeok (1960 : 321) marks this as unattested and Alhoniemi (1979 : 227) provides no comment for it in his list. Located far from any ocean, the Mari would supposedly not have had their own word for the marine animal. Indeed, Сергеев (2000 : 77; 2002 : 76) notes that an 18th-century Mari wordlist collected by one Mendier Bekdorin follows Russian *кит* with “оной названия у них нет”. In another 18th-century wordlist collected for Pallas (Сергеев 2000 : 144) the Mari informant appears to have answered with the Russian word: "китъ — китъ".

As the notions ‘whale’ and ‘sea monster’ were very closely connected in earlier eras, we may suppose that the word which the Mari informant provided was MariE *şərt* NW *şərt* ‘böser Geist’. This supernatural creature was identified with bodies of water and could catch those who had gone to the water to swim or fish (Sebeok, Ingemann 1956).

Though the use of the Cyrillic letter *у* in Pallas’s representation Шуръ seems to suggest a back vowel (see Сергеев 2002 : 106), the MariNW form with its front vowel remains a possible source — in his analysis of another 18th-century attestation of Mari, namely the poem in honor of Catherine the Great, Veenker (1981) notes that the Cyrillic letter *у* can denote MariNW *ə* after *ş*; the final soft sign in Pallas may also suggest front articulation.

### Бяи ’въ’ (265)

Sebeok (1960 : 298) marks this as unattested and Alhoniemi (1979 : 208) only marks it with a question mark. Because Mari does not have initial *b*-, we would have to suppose that the Cyrillic letter *ё* represents *β*, as it in fact does in several other items in Pallas’s book, e.g. Бурсимашъ ’оранъ’ = MariE *bursmaš* id. (185), Бытьонъ ’волны’ = MariE *bütøngə* id. (101). After a labial consonant, Cyrillic *я* typically represents the front low vowel *â*. No Mari word *βäj* is known, however.

If we examine the purportedly Mari item бяи in the context of the full entry in Pallas, we see that is nearly identical to the Udmurt item бай two lines below it. Тепляшина (1966) considers the purportedly Udmurt бай to represent Udm. vaj ‘give! (imperative)’, mistakenly reflected among words...
for ‘in’. One can view the Mari entry as simply an erroneous duplication of the Udmurt entry. This would not be the only such mix-up in Pallas; Alhoniemi (1979:217) has pointed to how Mordvin koda ‘how’ (represented as кода) was mistakenly placed under Mari in Pallas’s item 270 ‘какъ’.

Всё в ‘послѣ’ (269)

Sebeok (1960:328) marks this as unattested, while Alhoniemi (1969:209) only marks it with a question mark. The item as it appears in Pallas (listed after вары corresponding to Mari E вара ‘danach, dann, später’) is best regarded as a misprint for MariE baraše ‘letztere(r/s)’.

Умсысь ‘безъ (кромѣ)’ (375)

Sebeok (1960:325) marked this word as unattested, while Alhoniemi (1979:230) only lists it with a question mark.

While Pallas’s book spells the item with a final soft sign, suggesting a palatalized sibilant, Сергеев (2002:35, 183) cites the word as умсысь, with a final hard sign pointing to a non-palatalized s, from one of Pallas’s source manuscripts that represents a dialect transitional between the Krasnoufimsk and Kungursk varieties of Eastern Mari. The final soft sign in Pallas’s book can therefore be regarded as a misprint.

Due to the item’s Eastern Mari provenance and the final hard sign that Сергеев established, this item can be identified with умсыз ‘малоумный’ found in the dictionary compiled by Вершинин (2011) on Mari dialects of Tatarstan and Udmurtia. This Mari word consists of Russian ум ‘mind’ followed by the caritative suffix -сыз of Tatar origin. It is thus a partial calque of Russian безумный. We can suppose a misunderstanding between the wordlist compiler and the Mari informant such that the compiler’s request for an item ‘without’ was answered with a Mari word whose Russian translation is a compound containing без- ‘without’.

Items that remain unclear

The following items in Pallas remain unexplained, but some comments can nevertheless be made about certain items:

Си́ть ‘жена’ (10), Сергеев (2002:35, 181) notes that this word is present in a manuscript wordlist, reflecting the Krasnoufimsk dialect of Eastern Mari, that was compiled for Pallas, but he only assumes this is a Mari word that has fallen out of use;

Сапъ ‘слухъ’ (48); Сю́къ ‘шумъ’ (56); Кусташь, Кужашь ‘трудъ’ (62); Тырга ‘кругъ’ (73); Пешпорь ‘воздухъ’ (110); Югорь ‘кора’ (135);

Аракарму ‘виноградъ’ (143), both Sebeok (1960:297) and Alhoniemi (1979:208) identify the first component as MariE araka ‘Branntwein, Wein’, but the second part remains unclear and Sebeok only suggests a metathesis of MariE мör ‘Erdbeere’;

Чунлык ‘червь’ (145), Сергеев (2000:27—28, 98; 2002:34—35, 131, 184) notes that this item is found in a manuscript wordlist reflecting the Malmyž dialect and containing a large number of errors, but he only assumes that this is a Mari word that has fallen out of use;

Важикъ ‘звѣрь’ (147); Ромъ ‘тать, воръ’ (182); Елня ‘китъ’ (198); Сюн-зысь ‘благъ’ (219); Хась ‘ зло’ (220);
Тумпашь 'возить' (236), if this is not the typesetter’s misreading of MariE toltaš ‘bring’ (not in TschWb but known from MariE lit. толташ ‘to convey, to transport’ and attested in the 1926 dictionary of Васильев under толтём ‘везу’), as the two sequences of letters could very easily be confused with one another when written in cursive Cyrillic script.

Conclusion

Thus a deeper understanding of the Mari data in Pallas, going beyond Sebeok and Alhoniemi’s respective commentaries, can be reached by using the richer array of lexicographical resources which were not available to those earlier scholars, as well as well as by taking the peculiarities of Pallas’s typeface and 18th-century manuscript handwriting into account. Where the Cyrillic representation of the item in question closely matches the phonology of an attested Mari form even when the meaning differs, as in the case of Муней, Шьидаибьи, and Садиги, then we should assume a misunderstanding between wordlist compiler and informant.

If, as Sebeok notes in his commentary, Pallas’s ambitious work has drawn criticism since virtually the moment it was published for its overly hasty execution and the fact that much in it remains unusable, it nevertheless remains impressive that the overwhelming majority of Mari items can be recognized from more recent references, and the work provides vital evidence for reconstructing 18th-century Mari. Furthermore, even if errors in compilation or transmission might render it impossible to fully explain all items, further breakthroughs may be possible through e.g. the full publication of the manuscripts languishing in Russian archives. The present article has offered a contribution to the study of the Mari words in Pallas, but the material continues to merit scrutiny.

Address

Christopher Culver
Str. Bărsei 1/10, 400605 Cluj-Napoca, Romania
E-mail: christopher.culver@alumni.helsinki.fi
Phone: +40747371341

Abbreviations

MariE — East Mari; MariNw — Northwest Mari; MariW — West Mari; Tat — Tatar; Udm — Udmurt.

Sources of Lexical Data

Szilasi M. 1901, Cseremisz szótár (Vocabularium čeremisicum), Budapest.
Асылгараев Ш. Н., Ганиев Ф. А., Закиев М. З., Миннуллин К. М., Рамазанова Д. Б. 2007, Татарско-русский словарь, Казань.
Васильев В. М. 1926, Мари мутэр, Москва.
Вершинин В. И. 2011, Словарь марийских говоров Татарстана и Удмуртии, Иошкар-Ола.
REFERENCES


Iванов Й. Г. 1975, История марийского литературного языка, Йошкар-Ола.


Тепляшина Т. И. 1965, Памятники удмуртской письменности XVIII века, Москва.

КРИСТОФЕР КУЛВЕР (Хельсинки—Клуж-Напоца/Коложвар)

О НЕКОТОРЫХ РАНЕЕ НЕ ОПОЗНАННЫХ МАРИЙСКИХ СЛОВАХ В СЛОВАРЕ П. С. ПАЛЛАСА

«Сравнительные словари всех языков и наречий» П. С. Палласа (1787—1789 гг.) — это выдающаяся оранжерия записей марийского языка, содержащая марийские переводы 273 русских заглавных слов. Материал был исследован Т. А. Себеоком в его обширном комментарии, опубликованном в 1960 г., а затем А. Алхониеми, почти двадцать лет спустя. Однако оба ученых не смогли распознать всех марийских слов в словаре. С помощью современных лексических источников по марийскому языку, а также благодаря изучению рукописных словарей, послуживших источником для Палласа, автор статьи расшифровал некоторые из ранее не идентифицированных слов: Ирл амьрь 'буря' = марВ пуламрэ 'беспорядок, смута, раздор'; Садиги шар 'шаръ' = марВ тəртəш 'шар'; Всерсе арасэ 'последний (только появившийся)'; Умсысь безъ 'безъ (кромь) = марВ умзэш 'безумны'. В статье также отмечено, что бяи — это, возможно, удмуртское слово, ошибочно упомянутое как марийское. Результат данного исследования позволит лучше понять словарь Палласа, а также марийский язык XVIII века. Правда, 17 марийских слов в словаре Палласа пока остаются неясными.