

Preface

Those who have attempted to explain the kinship of the Finnish and Chud without comparing the languages more closely have come to the conclusion that the Cheremis tribe is closely related to the Finnish or *suomi*³. One such person is the gifted, erudite and well-known miŝter Rask⁴, who divides the entire Chud category into three families: the Finnic, Ugric and Bjarmic and notes that the Cheremis, Mordvin, Finnish, Lappish and Estonian languages are only types of the same Finnish family.*⁵ But in my opinion Cheremis together with Mordvin and Chuvash⁵ form their own family by all merits. Considering the external and internal characteristics of the languages, the Finnish family shows much less kinship with the Cheremis than with the Bjarmian languages. As Rask wants to generally define the kinship of languages belonging to the Chudic family, the Finnish family is so far removed from the Bjarmian that not even the Finnish and Ugric families differ so much. According to Rask, peoples geographically very far apart are closely related through

*⁵) The Bjarmian family entails idioms that are slightly distinct from one another and are customarily called Permian, Zyrian and Votyak. The Ugric family includes the Hungarian, Vogul and Oŝtyak languages.

- 3 Castrén uses the expression (*lingua*) *sumica* to indicate Finnish and more extensively in reference to Finno-Ugric languages. Castrén uses the word *Finnish*, depending on the context, to indicate Finnish alone, or the entire Finno-Ugric language group.
- 4 Rasmus Rask (1787–1832), the well-known Danish linguist, visited Turku in 1818 and urged the Finns to study their kindred languages (Korhonen 1986: 24–25).
- 5 In the time of Castrén, the Chuvash language, which was spoken adjacent to Mari in the Volga region, was thought to be a relative of the Finnish language. It was not until 1857 that August Ahlqvist proved that Chuvash belonged to the Turkic language family (Korhonen 1986: 80).

❖ IV ❖

language. This is not the case. As the settlements of the Finns approach the borders of the Bjarmian tribe in the Arkhangel'sk region, so do the Finnish-speaking languages come very close to those of the Bjarmians. The Bjarmian-speaking tribes spread throughout the area between Karelia, inhabited by Finns, and ancient Bulgaria, where the Chuvash, Cheremis, and Mordvins have settled. So if you look at the location, the inhabitants of Bjarmia are between the Finnish and Bulgar tribes. The languages have adapted to the requirements of their location. Inevitably, therefore, the Bjarmian languages are more closely affiliated with the Finnish and Bulgar languages than the Finnish languages are with those of the Bulgars, or at least with the Cheremis. As for the Ugric family of languages, I would just like to mention here that I have encountered several instances where the Cheremis is consistent with the Ostyak and Samoyed, but I dare not say anything certain about this consistency so far, as it may be due to their mutual past. *)

*) In order to give clarity to the differences between the Chudic languages, we must observe that the Finnish family has borrowed from the Germanic, the Bjarmia from the Slavic and the Bulgars from the Tatar. Since the Voguls, Ostyaks and Samoyeds are neighbours of the Tatar, it is quite plausible that both the Ugric and Samoyed languages have been exposed to Tatar language influence.



The essence and plan of this booklet require me to say something about its origins. During my stay in Finland for a few months last year, I happened to learn that among the Russian soldiers gathered in Helsinki, there was a Cheremis who had been born in the Koz'modem'jansk Okrug of the Principality of Kazan'. With his help, I hoped to gain a more accurate knowledge of the language of the Cheremis than could be afforded by the completely worthless books written in the language. This Cheremis' lack of a competence in Russian and other languages was such that my use of his help was limited in the study of the grammar ^{*)} of the Cheremis language published in Kazan'. I concentrated on this study for a few weeks and wrote down my observations in

*) This was the only source I could use in Helsinki. Here, however, are all the books dealing with the Cheremis language:

Сочиненія принадлежащія къ Грамматикѣ Черемискаго языка. В Санктпетербургѣ 1775 года.

Мя Осподьнанъ Исусъ Христосанъ святой Евангелья Матфей-гыць, Марко-гыць, Лука-гыць, Иоаннъ-гыцать Марла сирмяшешъ вазактэма. Питеръ Алашта 1821 и годамъ.

Святой Апостолвянъ Пашавля, нынанать сирмяшвя, Святой Иоаннанъ Апокалиписъ-ге. Марла сирмяшешъ вазэктэмавля. Питеръ-Алашта 1827 и-годамъ.

Черемисск. Грамм. Казань 1837.

Начатки Христiанскаго Ученiя или краткая свѣщенная исторiя и краткiй Катихизисъ на Черемискомъ языкѣ. Горнаго нарѣчiя. Казань 1832.

- 6 The correct form of the word is *tər*, i.e. it contains a schwa vowel for which there was no character in use in Castrén's time. See § 1 and the corresponding marginal notes.
- 7 Castrén describes the vowel variation between the dialects in *a ~ o*: *a* appears in Hill Mari, and *o* in Meadow Mari. The majority of the text sources used by Castrén represented Hill Mari, but part of them also came from Meadow Mari or mixed dialects.

❖ VI ❖

Helsinki. *) Then I travelled through the territory of the Cheremis to Siberia and went through my work with the help of another Cheremis man and made necessary corrections. What I did not want or could not correct in the text itself, let me mention in the following.

1. Upon observing the more precise pronunciation of vowels, I have become convinced that the Cheremis language uses variations such as those in the first paragraph. The variations, or schwas, as they are called, each derive from a short and weightless vowel, and not only in multi-syllable words but also in single-syllable words, e.g. *tör*, or *tr* 'calm'⁶. But I am certain that the variation in stressed vowels is not necessarily dependent on specific consonants; a specific note must be made of variation in the letter *a*. Word-initial *a*, namely, is very much affected by this kind of variation, and, in some individual dialects, it is pronounced approximately as *o*, e.g. *ala* or *ola* 'city'; *alak* or *olak* 'meadow'; *altalem* or *oltalem* 'I cheat'; *amasa* or *omasa* (*omsa*) 'door'; *andžem* or *ondžem* 'I observe'.⁷

The same book: "Лѣснаго нарѣчя." Same place of publication.

Сельскій полицейскій уставъ для государственныхъ крестьянъ, переведенный на Черемискій языкъ горнаго нарѣчя, Михайломъ Кроковскомъ. Санктпетербургъ 1843.

*) It should not be left unmentioned that this grammar, which I wrote in Swedish, has been translated into Latin for the most part by others.

❖ VII ❖

It often occurs in the first syllable following a consonant, e.g. *jadam* or *jodam* ‘I ask’; *jal* or *jol* ‘foot; leg’; *kajam* or *kojam* ‘salient’⁸; but less frequently it is found in the second syllable, e.g. *kadalam* or *kodolam*⁹ ‘I see smb off’; *agam* or *ogom* ‘I do not want’¹⁰. The second variation of the letter *a* where it is realised as *ä* (*e*) is most often heard at the end of words (see § 1).

2. Paragraph four lists more vowel changes. Oftentimes, this includes not only the variations *a* and *o*, *a* and *ä*, *ä* and *e*, which are described elsewhere, but also the variation *a* and *i*, e.g., *maara*, *maari* ‘Mari’,¹¹ *e* and *i*, e.g. *šeder*, *šidir* ‘star’,¹² *kec*, *kic* (ablative ending),¹³ *o* and *u*, e.g. *koat*, *kuat*¹⁴ ‘power’, etc. Of the consonants, the liquids *l* and *r* readily alternate, e.g. *örtnjel*, *örtnjer* ‘saddle’, etc.¹⁵

3. In paragraph three, the vowels are divided into hard and smooth. Perhaps it would be better to divide them into hard (*a*, *o*, *u*), smooth (*ä*, *ö*, *y*) and light vowels (*e* and *i*),¹⁶ as in Finnish. Upon closer examination of the Cheremis language, I have found that the light vowels are in harmony with the hard and smooth. But the hard and smooth do not go together well. They never occur together in the root itself, but appear in the derived forms of words, which sometimes contain hard and smooth vowels (see § 3).

4. Although consonants usually harden at the ends of words (§ 6), there are some words that also end in soft vowels. In these, the last syllable becomes stressed and somehow lengthens. Lengthened syllables seem to require that subsequent consonants, especially *z* and *ž*, be smooth

- 8 The word in the example, *kajam*, is a verb, meaning ‘I look like smth’.
- 9 This word does not have the *a ~ o* variation between dialects in the first syllable, and the second syllable has a schwa. The correct form in both Hill Mari and Meadow Mari is *kodalam* ‘I see smb off’.
- 10 Both forms have a back schwa *ə* in the second syllable. The meaning is simply that of an auxiliary verb of negation for the first person singular without a main verb.
- 11 The word meaning ‘Mari’ ends in a back schwa, and the vowel of the first syllable is not long, but, stressed, and it may be slightly drawn out: *marə*.
- 12 Both syllables have a front schwa vowel *ä*: *šädär*.
- 13 The correct form is *käc*.
- 14 The Tatar loanword *kuat* is the same in both Hill and most Meadow Mari dialects, and a difference is only notable in stress; the stress falls on the first syllable in Hill Mari and on the second in Meadow Mari.
- 15 The variation *l ~ r* is limited to a few words.
- 16 Here Castrén uses the terms *majores*, *minores*, *mediae*, literally ‘big, small, middle ones’ found in older Finnish grammars. Here they have been given the same names *hard*, *smooth*, *light* which Castrén uses further in the text and which are also found in *Ostiacica* (p. 53), according to the original German edition.

- 17 The vowel in the first syllable is a back schwa ə.
- 18 Since this form of negation in the second preterite is a result of a suffixed verb 'to be' *alam*, the final syllable does not follow vowel harmony, and it should therefore be *altelam*.
- 19 After *l*, *t* seldom becomes voiced.
- 20 The last variant is Meadow Mari. The second syllable of the first form mentioned should have a schwa in the second syllable: *jaləštem*.
- 21 Henceforth schwas are marked as *ä* and *a*, so that the words would be recognizable from a perspective of the modern language.
- 22 *ža*, *žä* (in the original *ža*, *žä*) is not a nominal derivational ending; it is the third person singular possessive suffix. Adjectives in Mari (as in many other Finno-Ugric languages) can also appear as nouns, identifying qualities, at which time they are marked with a third person singular possessive suffix.
- 23 In the original *oza*, *özä*.
- 24 This derivative suffix has been borrowed from Chuvash (Galkin 1966: 51).
- 25 *Jykšä* is an irregular participle form of the verb *jyam* 'I drink', not the verb *jyktem* 'I give something to drink'.

❖ VIII ❖

word-finally. Instead, consonants that begin an accented syllable remain relatively hard, e.g., *tölté* (*toldé*)¹⁷ 'without fire', *oltéläm*¹⁸ 'I wasn't'.¹⁹

5. The double stressing which is dealt with in paragraph eight is not common. Thus, the stress of the two-syllable words falls on the first but more commonly on the second syllable. A first syllable that is accented very often lengthens, but a second syllable only lengthens in words that end in a consonant. A vowel that bears no stress shortens, becomes a schwa vowel, or disappears altogether. Even, the stress in three-syllable words often falls on the last syllable, and the vowel of the previous syllable is elided, e.g. *jalaštem* or *jalštem*, *jolštem*²⁰ 'I bind'. But three-syllable words which are formed from two-syllable words with stress on the second syllable, retain this stress in the second syllable. Roughly the same is true of multi-syllable words.

6. Nominal derivational endings with specific meaning are ones forming:²¹

- a) abstract nouns ending in *ža*, *žä*, e.g. *jažo* 'beautiful', *jažoža* 'beauty'; *kuža* 'long', *kužaža* 'length'; *kelgä* 'deep', *kelgäžä* 'depth'; *käžgä* 'fat', *käžgäžä* 'fatness'.²²
- b) actor names ending in *aza*, *äzä*²³, e.g. *loktem* 'I enchant', *loktəza* 'enchanter'; *kätem* 'I herd', *kätəzä* 'herder'.²⁴ But rather frequently the same meaning is indicated with the participle present ending *ša*, e.g. *loša* 'catcher', *jykšä* 'drunk', which are derived from the words *loem* 'I catch' and *jyktem*²⁵,

❖ IX ❖

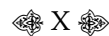
respectively. The same participle ending, added to the inessive form, is used in forming adjectives and especially nouns associated with place, e.g. *solaštəšə* ‘living in a village’ from the word *sola* ‘village’, iness. *solaštə*; *alaštəšə* ‘living in a city’, etc.

- c) adjectives ending in *n* (*an*, *än*, *ən*, *än*²⁶) that are in character possessives (§ 10, B), e.g. *lävrä* ‘filth’, *lävrän* ‘filthy’ or ‘of filth (gen.)’; *i* ‘ice’, *iän* ‘icy’ (gen. *in*); *koat*²⁷ ‘power’, *koatan* ‘powerful’, *naaməs* ‘shame’, *naaməsan* ‘shamed, of shame’;²⁸ *vət* ‘water’, *vädän* ‘watery’.
- d) adjectives ending in *də*, *də*²⁹, e.g. *šuldə* ‘inexpensive’, *nəgədə* ‘thick’, *paškudə* ‘nearby’³⁰, *pingädə* ‘voracious’³¹, *pəčkədə* ‘dusky’. This adjectival ending is common in Samoyedic languages, too.
- e) adjectives ending in *tämä*, *dämä*³², Fin. *toin*, *ton* (*ttoma*), Lapp. *taebme*, *tebme*, *teme*, Zyr. *täm*, *töm*, e.g. *naaməs* ‘shame’, *naaməstəmə*³³ ‘brazen, shameless’, *päləš* ‘ear’, *päləštämä* ‘deaf’, *jəlmä* ‘tongue’, *jəlmätämä* ‘mute’, *šur* ‘horn’, *šurdəmə* ‘polled, hornless’.

There are also some adjectives ending in *la*, e.g. *totlə* ‘sweet’, *Soasla Maarə*³⁴ ‘Chuvash’. In Zyrian this ending is a case ending, and it is not beyond belief that the ending in Cheremis is of the same origin, which is why it is often used adverbially, e.g. *Rušla* ‘in a Russian way’ (*no Pycckiü*), *Marla* ‘in a Cheremis way’.³⁵

- 26 In EGT *en*, *in*.
- 27 In modern Hill Mari *kuat*.
- 28 This should be *naməs* and *naməsan*.
- 29 In the original *da*, *dä*.
- 30 The basic meaning of the word is ‘neighbour’, and it is a Chuvash loanword (Räsänen 1920: 181).
- 31 The basic meaning of the word is ‘hard, firm’ and figuratively ‘stingy’.
- 32 In the original *temä*, *demä*.
- 33 See marginal note 28.
- 34 In modern Hill Mari *suasla marə*.
- 35 Here Castrén assumes a single suffix where two Chuvash loan suffixes are at issue, one with a reduced vowel and the other with *a ~ ä*. *Totlə* is a Chuvash loan in Mari (Räsänen 1920: 228). The forms *rušla* etc. are translated as both ‘Russian, in a Russian way’ and ‘in Russian’. The Chuvash suffix *la* is defined as an adverb derivational suffix. Its usages are identical in both Mari and Chuvash, but in Mari it is called the modal-comparative case. (Fedotov 1965: 34, 40; Saarinen 1993: 148–150.)

- 36 In the original *tidä*, cf. § 24, marginal note 110.
- 37 The second and third person singular pronouns in Finnish, Saami, Mordvin and Mari have not developed from the same stems. In fact, the Finnish *se* and Mari *tädä* derive from two distinct demonstrative pronoun stems. (UEW 1988: 33–34, 453, 526–527, 539.)



In addition, there are numerous nominal endings whose meanings still evade me, e.g. *ž*, *š*, *k*, *kə*, *gə*, *ngə*, etc. I have not observed diminutive or augmentative nouns in this language.

7. The third person pronoun *tädä*³⁶ is derived from the root *tä*, which is also present in the language as such, this can be compared with the derivation of *puda* ‘that’ from the root *pu*, found in the Samoyed language. It should also be noted that the second and third person pronouns (*tänj* and *tädä*) do not come from one and the same root in Cheremis, but this distinction is also found in the Finnish language (*sinä*, earlier *tenä*, *te*, *tu*, *se* ‘it’), Lappish and Mordvin (*ton* ‘you (sg.)’, *son* ‘he/she/it’), Samoyedic (*pudar* ‘you (sg.)’, *pu* ‘that’) etc.³⁷

8. Regarding suffixes, I will merely point out that, after re-examining the Cheremis language, I have not found all the suffix forms I mentioned in the grammar under the authority of the soldier I mentioned above.

9. The vowel *e* of the present participle ending (*še*) and the ordinal ending (*mše*) is close to *a*, almost *ä*, as the author of the Kazan’ grammar writes here *a* instead of its smooth variant. This spelling, however, is incorrect, because *a* does not follow smooth vowels in the root. This suffix, however, can follow any vowels. But when the sound is so vague that it can hardly be expressed consistently with a single character, I have opted

❖ XI ❖

to use the letter *ä* in this instance and similar ones, and I do not think it is far from the correct pronunciation.³⁸

It is sufficiently clear from the points raised that my book, which only deals with the elementary grammar, is lame in that respect, too. I have not had the opportunity to study the matter more closely. If all goes according to my wishes, I will present the elementary grammar of Cheremis in more depth the second time around.

Given in Kazan', 1st May MDCCCXLV.

38 This is again an issue of schwa vowels in the suffixes mentioned that did not have their own symbols and whose quality was influenced by the surrounding vowels. In reality, both suffixes have two variants with a front schwa *ä* or back schwa *ə* depending on vowel harmony (cf. § 1 and its marginal notes).