



I. Litteræ.

Vocales.

1. Lingua Tscheremissa vocales, quæ propriæ vocari possunt, has octo habet: a, e, i, o, u, y, ä, ö, quæ similiter atque in Finnica, Russica et compluribus aliis linguis enuntiantur; præterea multas vocalium variationes. Ad has pertinent: 1:º Schevata seu semi-vocales; 2:º vocales emphaticæ. Illæ pronuntiantur sono celeri, aures paene præterlabente, confuso, harum vero sonus plenior, latior, gravior est, quam ipsarum earum, quæ proprio nomine vocales appellantur. Schevata nisi in brevibus et accentu carentibus syllabis non inveniuntur et ei in primis rei inserviunt, ut pronuciari possint plures consonantes concurrentes. Præterea syllabæ accentu carentes in fine verborum naturam schevatum frequenter assumunt, ac tum plerumque ejici solent et ante verba et ante terminationes, ex. gr. *kodaz'e*, qui, Dativus *kodaz'lan*, cui. Pronunciatio vocalium, quæ emphaticæ dicuntur, ex quibusdam et antecedentibus et sequentibus consonantibus pendere videtur, in primis e litteris *l, m, n* et *x*. Quando cum his concurrat, sonum suum in similitudinem litteræ *o*

I Phonetics

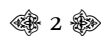
Vowels

§ 1 The Cheremis language has eight proper vowels: *a, e, i, o, u, y*³⁹, *ä, ö*, which are pronounced in the same manner as in Finnish, Russian and most other languages; furthermore it has many vowel variations. These include: 1. schwa or semi-vowels⁴⁰; 2. emphatic vowels. The former ones are pronounced fast, almost passing the ear, unclearly, whereas the sound of the latter vowels is fuller, stronger, more emphatic, and they are called proper vowels. The schwas are found only in short syllables lacking stress and they serve especially in making the pronunciation of consonant clusters possible. Moreover, syllables lacking stress adopt a schwa nature more often at the end of words, and then they are often apocopated both before words and suffixes, e.g., *kədəžə* ‘which’, dative *kədəžlan* ‘to which’. The pronunciation of the so-called emphatic vowels seems to be dependent both on the preceding and following consonants, especially *l, m, n* and *χ*. In contact with these, *a* starts to resemble *o, i* and becomes similar to the Russian *υ*⁴¹; others (*e, u*) do not differ in weight and sharpness from the corresponding vowels in German. Further, *a* has a strange variation where it nearly approaches *ä*, and often especially at the end of words whose last syllable lacks stress.⁴² The vowel *u* also has a smoother

- 39 Castrén used the letter *y* to indicate *ü*.
- 40 Hill Mari schwas are of shorter duration than the full vowels. They can, however, also occur in stressed syllables. Mari phoneticians consider the two schwas to be close-mid, but their quality varies and is dependent on the surrounding full vowels (Gruzov 1964: 31–34; Vikström – Zorina 2007: 41, 45, 79–82). In the first half of the 19th century, the orthography used for Mari was based on the Cyrillic script. In the two grammars and some biblical texts published before *Elementa*, Mari sounds not existing in Russian were written with vowel signs found in the Russian orthography or with their combinations. As such, the schwas could not be distinguished from full vowels in writing. Special signs for Mari sounds not existing in the Cyrillic orthography were not introduced until the second half of the 19th century by the Translation Committee working under St. Gurij’s brotherhood in Kazan’, and they are still in use today (Ivanov 1975: 27–28). Using Mari informants Castrén was able to distinguish the sounds *ä, ö* and *ü*. His description of the schwas shows that he, to some extent, understood the quality of these vowels, but he was not able to mark them in a distinct way. Instead, he used full vowels. In this edition, the schwas are indicated

with the signs *a* (back) and *ä* (front).

- 41 In the case of vowels Castrén calls *a* and *i*, he is actually describing the variation of the schwas.
- 42 This, again, concerns the schwas.
- 43 In the printed book erroneously *s* which is later corrected in the appendix *Corrigenda* and in its longer version attached to EGS.
- 44 Castrén uses the signs *s'*, *z'* and *c'* for *š*, *ž*, *č*.
- 45 I.e. *β*, *γ*, *δ* in the Finno-Ugric Transcription.
- 46 In the Swedish manuscript: Tenues have the same pronunciation as in all other languages but regarding liquids one has to notice that, after *k*, *l* turns into a strange, guttural sound influenced by *k*.

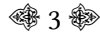


variant, which can sometimes be heard after *j*⁴³ and which, no doubt, has been borrowed from the Russian language.

N.B. As individual changes in vowels are probably dependant either on the following consonants or on the stress or perhaps on other matters, I neither want nor dare to introduce any new letters for them, especially as experience has taught me that in less cultivated languages the pronunciation of vowels by one individual cannot be fully trusted in every respect. One must, however, observe that the vowel variations noticed by me in the Cheremis language also occur in Samoyed.

Consonants

§ 2 The consonants in the Cheremis language are: *b*, *d*, *g*, *f*, *j*, *k*, *l*, *m*, *n*, *p*, *r*, *s*, *z*, *c*, *š*, *ž*, *č*, *t*, *χ*, *v*.⁴⁴ The faint sound *h* is not used at all, *f* rarely. The spirant *χ* sounds like the Russian *x*. When the sounds *b*, *g*, *d* occur between two vowels they are pronounced like spirants, thus *b* resembles *v* whereas *g* and *d* become fricatives like the ones in Lappish and other known languages.⁴⁵ Otherwise they are usually pronounced like half-voiced plosives and often appear after a preceding *m* or *n*. Tenues have the same pronunciation as in the German language with the exceptions of *k* and *l*; the former is pronounced like *q* in Tatar and the latter sometimes like *l* in Russian.⁴⁶ Of the sibilants, *s* sometimes gets



a smooth, slightly aspirated pronunciation close to the Finnish aspirated *t*. The other sibilants are pronounced almost in the same way as in Russian: *z* like *з*, *c* like *ц* (*ts*), *ž* like *ж*, *š* like *ш*, *č* like *ч*. The consonants *ž*, *š* and *č*, however, do not have the same sharp sound as the ones in Russian, instead, they come close to the composed sounds *zj*, *sj*, *cj*.^{*)} 47

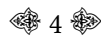
Vowel alternations

§ 3 As in the Finnish⁴⁸ languages in general, the vowels in the word stem do not vary, but in the suffixes, on the contrary, they vary and are dependant on the vowels in the stem; in the same way in Cheremis, the vowels in the suffixes adapt to the immutable ones of the stem. This law⁴⁹ so prevalent and widely used in the Finnish language has started to lose its power and meaning in Cheremis, therefore some suffixes have already acquired fixed vowels such as *ko* ‘wave’, pl. *kovälä* or *kovlä*⁵⁰; *pört* ‘room’, instr. *pörtän*⁵¹. In the Finnish language, *ä* always occurs in smooth suffixes, *a* in hard ones.⁵² Likewise, in Cheremis, the hard suffixes take the hard *a*, but in the smooth suffixes *ä* and *e* alternate. In Cheremis, a hard vowel in the suffix is always triggered by *a*, *o*, *u* in the word stem, while a smooth vowel is triggered by all other vowels. As it sometimes happens that the stem contains both hard and

*) The letter *c* compiled of the letters *t* and *s* is sometimes expressed in this book with the compound letters *ts* in order to make its formation clearer.

- 47 Castrén often writes *č* with *dč*, too. In this edition *dč* is used to represent instances of *dč*.
- 48 “Finnish” in this context is equivalent to the later term *Finno-Ugric*, which refers to both languages and peoples (cf. *Ostiacica* 2018: 41).
- 49 I.e. vowel harmony.
- 50 The plural ending (with its vowel) in Hill Mari goes back to an independent word of Turkic origin and it does not adapt to vowel harmony (cf. Luutonen 1997: 56–57).
- 51 Instead of the front schwa, *ä*, Castrén writes an *o* in *pörton* (see marginal note 40), where it, of course, would be counter to vowel harmony.
- 52 See marginal note 16.

- 53 Castrén writes the palatalized sounds *ń, l', t'* etc. with *nj, lj, tj*.
- 54 In the first or *am* conjugation the present tense forms 1SG and 2SG have an *a* in the personal suffix and an *e* in the 3SG; these vowels belong to the suffixes.
- 55 In the second or *em* conjugation the present tense forms 1SG and 2SG have an *e* in the personal suffix and *a, ä* in the personal suffixes of the third persons.
- 56 The change *a- > u-* (Castrén: *o- > u-*) occurs in this word alone.



smooth or medium vowels *e* and *i*, the suffix can have either of the two vowels (*a* or *ä*), though most often the choice depends on the last vowel of the stem; e.g. *loenäm*, less frequently *loenam* 'I caught'. The Cheremis language, in the same way as Samoyed, often joins adjacent words, if this is possible within the laws of euphonia. But in this conjunction or composition, it often happens that the vowels in the second word adapt according to the nature of the vowels in the first, e.g., *kogarnja* 'Friday' from *kogo* 'big' and *ärnjä* 'week'⁵³; *tagadčə* 'today' from *ta* 'this' and *kedčə* 'day'; *nällä* 'fourty' from *näl* 'four' and *lu* 'ten'; *agəl* 'is not' from *ak* 'not' and *əl* 'be'. But this law is as vague as the one concerning word suffixes.

§ 4 Although it may be said that according to the law mentioned above the Finnish languages like to keep the vowels of the stem immutable, the Cheremis language has nevertheless, in this respect, deviated from its original nature. It allows not only the final vowels of the stem to change in many ways, but, even in the vowels of the stem itself, it tolerates certain transformations which do not depend on the vowels of the directly preceding word. The most common transformations are *a, ä* to *e*, e.g., *lədam* 'I read', *lədeš* 's/he reads'; *käčäläm* 'I search', *käčäleš* 's/he searches'⁵⁴; *e* to *a, ä*, e.g., *šalgem* 'I stand', *šalga* 's/he stands'; *särem* 'I turn', *särä* 's/he turns'⁵⁵; *a* to *u*, e.g., *əlam* 'I am', *ulə* '(there) is'⁵⁶. The word-final vowel is often dropped, especially if the next word begins with a vowel, e.g., *ves edem* 'another person' (*vesä* 'another').



Consonant alternations

§ 5 Most, if not all, languages of Finnish origin have traces of an important law according to which two consonants cannot follow each other in one syllable. In Cheremis, this law is modified in so far as liquids and among them first and foremost *l* and *r* and also sibilants can easily occur in the same syllable. Cheremis, as well as Zyrian, even place a pleonaſtic *d* before a sibilant, e.g., *pörtländžä* (actually *pörtlänžä*) ‘for his room’, *imnidondžə* (actually *imnidonžə*)⁵⁷ ‘with his horse’.⁵⁸ Sometimes even three consonants can occur in the same syllable, if two of them are liquids or sibilants, e.g., *šträš* ‘cloth’⁵⁹. – In order to avoid a futile sound as a result of two or more dissimilar consonants interfering with one another the language uses two ways; namely either 1) it omits one consonant, e.g., *šolštam* ‘I steal’, opt. *šolšnem*; or 2) adds a semi-vowel between them, e.g., *jalštem* ‘I bind’, ptcp. *jalštəšə* (actually *jalštšə* or *jalščə*).⁶⁰

§ 6 Like vowels, consonants can also be divided into hard and smooth. Hard ones are: *k, p, t, s, š, c, č*; smooth ones: *b, d, g, j, l, m, n, r, z, ž, v, x*.⁶¹ Most consonant alternations depend on the general tendency that the hard consonants match up with hard ones and the smooth with smooth ones. This law can be expressed simply: a hard consonant cannot follow a smooth consonant of the same speech organ. Therefore, they change: *k* into *g*, *p* into *b* (*v*), *t* into *d*, *s* into *z*, *š* into *ž*, *c* into *dz*; and these changes occur not only in

- 57 Castrén regards *don(o)* as a case suffix. In modern grammars it is classified as a post-position.
- 58 In the Swedish manuscript Castrén mentions that “The same *d* also occurs sometimes before *r* in the beginning of words, although otherwise neither *d*, *b* nor *g* can be pronounced in word-initial position”. He may be referring here to (rare) words starting with the sequence *t(ə/ä)r-*, to a few new Russian loanwords or to cases of progressive assimilation caused by the preceding word (as presented under § 6).
- 59 In Hill Mari dictionaries, this word is usually presented in the form *äštäräš*, but, in the dialects, forms without the schwa also exist.
- 60 According to the present-day orthography, the verbs *šolštam* and *jalštem* both have a schwa in the second syllable. It can be assumed that schwaless forms may exist in dialects and idiolects.
- 61 Here, Castrén uses the terms “hard” and “smooth” to indicate voiceless and voiced, respectively.

- 62 Cf. marginal note 57.
 63 This should be *säpton*. Stressed vowels are pronounced with a longer duration, which has led Castrén to occasionally write them with two letters.
 64 Sometimes but inconsistently, Castrén writes the fricative *v* with the letter *w*, which would actually express the phonetic value of this bilabial fricative better than *v*.
 65 This should be *žava*.



the same word but also in the beginning of the following word, e.g., *kəmbadčəš* ‘triple’ (*padčəš*), *pajan bört* ‘a rich house’ (*pört*), *iän dər* ‘frozen coast’ (*tər*), *iän zəkər* ‘frozen bread’ (*səkər*), *əšan žümbel* ‘a clever brother’ (*šümbel*), *kəm dzəbə* ‘three chickens’ (*cəbə*), *šəm džuda* ‘seven miracles’ (*čuda*). In addition, hard consonants often become smoother after a vowel, especially in word-final position, e.g., *kogəc* ‘out of the wave’, cf. *kitkəc* ‘from hand’, *kəzədən*⁶² ‘with a knife’, cf. *sääpton*⁶³ ‘with a rein’. In contrast, the smooth consonants following hard ones change into corresponding hard consonants, and in word-final position a smooth consonant often, though not always, transforms into a hard one, e.g., *kit* ‘hand’, illat. *kidəš*, *vəzət* ‘five’ or *vəc* ‘five’, etc. If a word-final hard consonant is preceded by a smooth consonant, this preceding consonant changes into a hard one, e.g., *imnižə* ‘his horse’, *imništ* (instead of *imnižt*) ‘their horse’. The smooth consonants seldom occur word-initially in a single word, but if a word is combined with others, the first sound of the second word is smoothed. According to this, a smooth consonant in word-initial position should also become hardened after a hard consonant at the end of the preceding word, which sometimes happens, but because of the strong preference of the language to smooth sounds, often after a hard ending of the preceding word a schwa or an *n* (*ən*) is added and thus the following consonant naturally stays unchanged, e.g., *wicə*⁶⁴ *žava*⁶⁵ ‘five frogs’ (instead of *wic* or



wit šaava)⁶⁶. For the same reason, the suffixes *c* and *cän*, *käc* (*gäc*) and *käcän* (*gäcän*) alternate, e.g., *saic* or *saicän*⁶⁷ ‘from behind’, *mänjgäc* or *mänjgäcän*⁶⁸ ‘from me’.⁶⁹ Presumably, increasing smoothness can also be seen in cases where the same or related consonant closes the preceding and starts the following word or syllable; here the word-final consonant is omitted, e.g., *vät sydä* ‘500’ (instead of *väc šydä*), *kändäx šydä* ‘800’ (instead of *kändäxš šydä*)⁷⁰, *ikänä* ‘once’ (instead of *ik känä*), *kəmätön* ‘in sixes’ (instead of *kəmät-ton*), etc.

§ 7 Apart from the consonant alternations that we presented in the paragraph above, there also occur various others, but they depend on laws that are even less common. So *m* followed by *g* often changes into *n*, e.g., *kənan gänjät* ‘sometime’ (instead of *kənam gänjät*). Furthermore, sometimes *p* changes into *v* and *v* preceded by *m* into *b*, e.g., *pad* ‘pot’, *toi vad* ‘a copper pot’, *kəm baž* ‘three roots’ (*važ* ‘root’). Additionally, one has to note the change of *č* to *rt*, which occasionally occurs, e.g. *kadčkam* ‘I eat’, opt. *kartnem*, *padčam* ‘I open’, opt. *partnem*, etc.⁷¹

N.B. What has been told about consonant alternations above is neither a full nor a definite list of them, but it gives a general idea of the laws which consonants follow, and they show astonishing similarities between Cheremis and Samoyed languages.

- 66 The word for ‘frog’ in Hill Mari is usually given in the form *žawa*. It is, however, more obvious that the word would, in some dialects, have a word-initial schwa (as, e.g., in the north-western dialect). In that case, it should read *wič əžawa*. In the Swedish manuscript, Castrén gives another example *kokə drämäs* ‘two women’, which should definitely read *kok ədrämäs*. Thus, in both cases, Castrén has segmented the words incorrectly.
- 67 In present-day Hill Mari *šajəc(ən)*.
- 68 Castrén writes *ń* with *nj* but not consistently.
- 69 Castrén regards the postposition *gäc(ən)* a case suffix of the elative/ablative. The word *gäc* has its origins in the noun *körgä* ‘inside’ and the Uralic separative case suffix **-tA > -c*. The unproductive case ending *c* is used with adverbial and postpositional stems. *(ə/ä)n* is an adverbial suffix, etymologically identical with the suffix *n* of the genitive-instructive case. Its use is optional. (Berezcki 2002: 40–43.)
- 70 In the Swedish manuscript, Castrén writes *kändänχ* and *kändänχš*, respectively. In dialects, forms like *kändäkčš* and *kändänχš* are also attested.
- 71 On rhotacism in Hill Mari, see Wichmann 1909: 114–119.