

Contents

Содержание

1. Introduction	VII
2. The contents of the columns	VIII
2.1. The first column: the word	VIII
2.2. The second column: language	X
2.3. The third column: word class	X
2.4. The fourth column: sources	XI
3. Statistics	XII
1. Введение	XIV
2. Содержание граф	XV
2.1. Первая графа: слово	XV
2.2. Вторая графа: язык	XVII
2.3. Третья графа: часть речи	XVII
2.4. Четвертая графа: ссылки на источники	XVIII
3. Статистические данные	XIX
4. The alphabetical order of Erzya and Moksha – Алфавитный порядок мордовских букв	XXI
5. Bibliography – Литература	XXI
6. The reverse-alphabetized word list – Список слов в обратном алфавитном порядке	1

1. Introduction

This dictionary is comprised of words from the two literary languages Erzya and Moksha, as well as from the Mordvin dialects. The vocabulary is presented in the form of a reverse-alphabetized word list. The main purpose of the dictionary is to serve as a source for the study of derivation and word structure. Reverse dictionaries have already been published of the following Uralic languages: Hungarian (1969), Mansi (1971), Finnish (1972), Estonian (1979), Ludian (1986), Ingrian (1986), Livonian (1988), Votian (1989), Udmurt (1992), Saami (2002) and Mari (2002, 2003).

The basic vocabulary sources for the reverse dictionary published at present are the largest dictionaries of the modern Mordvin languages *Эрзянско-русский словарь* (1993) and *Мокшанско-русский словарь* (1998), as well as the special dictionaries of compound words *Вейсэ, башка, теишкс вельде* (2002) and *Марса, башка, китькскя вельде* (2002). Additional vocabulary sources include:

Erzya:

- A. П. Феоктистов: *Русско-мордовский словарь* (1971; known as Damaskin's dictionary);
- H. Paasonens *Mordwinisches Wörterbuch* (1990–1996);
- A. Alhoniemi & N. Agafonova & M. Mosin: *Suomalais-ersäläinen sanakirja* (1999);
- К. Г. Абрамов: *Валонь ёвтнема валкс* (2002).

Moksha:

- A. Ahlquist: *Versuch einer mokscha-mordwinischen Grammatik* (1861), pp. 149–212;
- J. Juhász: *Moksa-mordvin szójegyzék* (1961);
- H. Paasonens *Mordwinisches Wörterbuch* (1990–1996).

The words from sources using Latin script (Ahlquist 1861, Juhász 1961, Paasonen 1990–1996) were transliterated into Cyrillic. During the editing process, the list of words combining the vocabulary of all the above mentioned sources was corrected and orthographically unified. For this reason, words in the present reverse list sometimes differ orthographically from the original words in the sources. Such changes in orthography alter the word's location in the alphabetic order, which is something to keep in mind when searching for a particular word in a source. For example, in the reverse-alphabetized list, we have the Erzya verb *мадстнемс*,

but in the vocabulary source the corresponding word is *мацтнемс*. For details, see section 2.1.

In selecting words for the reverse dictionary, the compilers have tried to avoid excessive lexical variation, which means that not all dialectal and other variants have been included. Some productive derivational forms are represented by just a few examples. These include, for instance, words formed by adding the Moksha elements *-начк* and *-шовор* (e.g. *ардомначк*, *ардомшовор* 'during travelling, in the process of travelling'; for more examples, see Kelin 1991), and Erzya and Moksha words ending in the suffix *-мезнек* (e.g. Moksha *прянек-мезнек* 'everything including the head; the head and all other things').

The user should note that some of the vocabulary sources, e.g. *Suomalais-ersäläinen sanakirja*, also contain neologisms.

The material is arranged in four columns: 1) word, 2) language, 3) word class, 4) sources. In the following, the contents of each column are explained.

2. The contents of the columns

2.1. The first column: the word

The first column contains the words, alphabetized beginning from the end of the word. The user should recall that the dictionary contains both literary and dialectal vocabulary. Words belonging to the same word family can sometimes be found in two or more places if some component of the lexeme appears in both literary and dialectal form. There also are dialectal variants of suffixes. For instance, in addition to verbs ending in the literary *-мс* there are also entries with the suffixes *-мкс*, *-мекс*, *-мокс*.

In the orthographical tradition of the Mordvin languages, we can observe two different principles that have been pursued: the morphological and the phonetic principles. According to the morphological principle, the phonetic assimilatory changes occurring in the process of deriving a word or combining the elements of a compound are not orthographically marked. In the opposite case, i.e., if such phonetic changes are reflected in the orthography, the word is said to have been formed by applying the phonetic principle. For instance, the Erzya word meaning 'mill' can be written using either the morphological principle, as *ведькев* (*ведь* 'water' + *кев* 'stone'), or by the phonetic principle, as *ведьгев*. A Moksha example: the word with the meaning 'river bank' could either be written according to the morphological principle as *ляйтърва* (*ляй* 'river' + *търва* 'bank'), or in accordance with the phonetic principle as *ляйдърва*. The vocabulary sources differ from each other with respect to these two principles,

and a single source also can be inconsistent in its application of the principles, e.g. *Марса, башка, китькскя вельде* (2002).

In order to reduce the heterogeneity found in the sources and to show the structure of Mordvin words most clearly, the compilers of the present dictionary decided to prefer the morphological principle of orthography. In the reverse word list, this principle was applied in the spelling of Erzya and Moksha compound words (originating from any source) and in the writing of other Erzya words, as well. Exceptions to the morphological principle, i.e., applications of the phonetic principle, are as follows: 1) certain Moksha and Erzya compound words with fixed conventionalized spelling, e.g. Moksha *вержи, шинжарома, очижи; сизьгемонь, ведьгемонь*; Erzya *ниленьгемень, сисемзядт*, etc; 2) Moksha non-compounds from *Мокшанско-русский словарь* (1998) and the dialectological sources (Ahlquist 1961; Juhász 1961; Paasonen 1990–1996); 3) and Erzya verbs from the dialectological vocabularies (Феокистов 1971; Paasonen 1990–1996).

The orthographic representation of the Moksha reduced vowel is not straightforward. Earlier, the reduced vowel of the first syllable was not represented. According to the present orthographic norms (1995: 6), the reduced vowel can be rendered with one of four different letters – *ь, е, о* or *а* – depending on its location in the word. Only one of the vocabulary sources, viz. *Марса, башка, китькскя вельде* (2002), uses the letter *ь* for this purpose. The norm that officially introduced *ь* to represent the reduced vowel in the first syllable came into force on the 1st September 1999. The compilers of the present reverse dictionary decided to comply with this rule, which also increases the structural transparency of certain Moksha words, as all vowels now have an explicit orthographical representation. In the reverse word list, the letter *ь* is used to mark the reduced vowel in the first syllable also in words from sources not making use of *ь* for this purpose. For example, the words written in *Мокшанско-русский словарь* (1998) as *трва, плманжа, срхксемс* appear as follows in the reverse dictionary: *тьрва, пълманжа, сърхксемс*. The letter *ь* has been used to mark the reduced vowel in the non-first syllable only in cases where the reduced vowel appears in the first syllable of a compound word's second item, e.g. *валкърня* (*вал* + *кърня*), *вирьтьрва* (*вирь* + *тьрва*). The user should note that the letter *ь* also often functions purely as a separation mark which does not represent any phoneme, e.g. *вальёжа, вальюр, кудьётконь*.

Homonyms representing the same word class have not been entered separately. If homonyms differ in their word-class specifications, they are entered separately on different lines; see e.g. *эй*, which is presented in two entries: as a noun (n) and as an interjection (in).

2.2. The second column: language

The second column indicates the language to which the word belongs. The meanings of the labels are as follows:

- E** = Erzya word
- M** = Moksha word
- EM** = a graphemically identical word is found both in Erzya and Moksha

2.3. The third column: word class

The third column gives the word-class label for the word in the first column. The specifications used are: noun, numeral, pronoun, adjective, adverb, postposition, conjunction, interjection, particle, verb, participle, gerund, and descriptive word. The division into word classes in most cases reflects the classification of dictionary entries in *Эрзянско-русский словарь* (1993) and *Мокшанско-русский словарь* (1998). According to the system adopted in these dictionaries, pronominal words functioning as nouns and adjectives are labelled as pronouns (**pr**), whereas adverbial pronominals are classified as adverbs (**ad**).

If a vocabulary source provides a word with several word-class labels, these are given as alternatives in the reverse word list, e.g. **n/a** (noun or adjective), **pc/ge** (participle or gerund).

The abbreviations for word classes, in alphabetical order, are as follows:

- a** = adjective
- ad** = adverb
- co** = conjunction
- de** = descriptive word
- ge** = gerund
- in** = interjection (including words used in calling and tending animals)
- n** = noun
- nu** = numeral
- pa** = particle
- po** = postposition
- pr** = pronoun
- pc** = participle
- v** = verb

2.4. The fourth column: sources

The fourth column shows a list of sources in which the word appears. Page numbers are given in the case of *Suomalais-ersäläinen sanakirja* (e.g. **S365**), *Русско-мордовский словарь* (e.g. **D69**), and Ahlquist's vocabulary (e.g. **Q124**).

The distribution of the word in the sources can only partially be inferred from the abbreviation string in the fourth column. This is because the words were extracted from the sources in two different ways. The bulk of the vocabulary was taken from computer files containing all the words of the source; these sources include *Эрзянско-русский словарь* (1993), *Мокшанско-русский словарь* (1998), *Вейсэ, башка, теишкс вельде* (2002), and *Марса, башка, китьксся вельде* (2002). The basic dictionary corpus thus obtained was then supplemented with manually selected items from other sources. These manually processed sources are *Русско-мордовский словарь* (1971), *Suomalais-ersäläinen sanakirja* (1999), *Валонь ёвтнема валкс* (2002), *H. Paasonens Mordwinisches Wörterbuch* (1990–1996), *Versuch einer mokscha-mordwinischen Grammatik* (1861) and *Moksa-mordvin szójegyzék* (1961). Those sources which were dealt with manually are indicated by upper-case abbreviations (**D, S, A, P, Q, J**); those sources from which practically all words were extracted automatically are indicated by lower-case abbreviations (**e, z, m, c**). The existence of a word in a manually processed source is indicated only if the word does not appear in any automatically extracted source. Thus the absence of an abbreviation referring to a manually processed source in the source column does not necessarily indicate the absence of the entry word in that source.

The abbreviations for the sources, in alphabetical order, are as follows:

- A** = Абрамов, К. Г.: *Валонь ёвтнема валкс*. Саранск: Мордовской книжной издательствась. 2002.
- c** = Алямкин, Н. С. & Ананьина, К. И. & Арискин, Н. И. et al. (eds.): *Марса, башка, китьксся вельде*. Саранск: Красный Октябрь. 2002.
- D** = Феоктистов, А. П.: *Русско-мордовский словарь. Из истории отечественной лексикографии*. Москва: Наука. 1971.
- e** = Серебренников, Б. А. & Бузакова, Р. Н. & Мосин, М. В. (eds.): *Эрзянско-русский словарь*. Москва: Русский язык, Дигора. 1993.
- J** = Juhász, J.: *Moksa-mordvin szójegyzék*. Budapest: Akadémiai. 1961.

- m** = Серебренников, Б. А. & Феоктистов, А. П. & Поляков, О. Е. (eds.): *Мокшанско-русский словарь*. Москва: Русский язык, Дигора. 1998.
- P** = Heikkilä, K. & Kahla, M. (eds.): *H. Paasonens Mordwinisches Wörterbuch*. Vols. I–IV. Helsinki: Société Finno-Ougrienne. 1990–1996.
- Q** = Ahlquist, A.: *Versuch einer mokscha-mordwinischen Grammatik nebst Texten und Wörterverzeichnis*. St. Petersburg: Eggers et Comp. 1861.
- S** = Alhoniemi, A. & Agafonova, N. & Mosin, M.: *Suomalais-ersäläinen sanakirja*. Publications of the Department of Finnish and General Linguistics of the University of Turku, 64. Turku. 1999.
- z** = Агафонова, Н. А. & Имайкина, М. Д. & Мосин, М. В. et al. (eds.): *Вейсэ, башка, тешкс вельде*. Саранск: Красный Октябрь. 2002.

3. Statistics

The number of words belonging to each language can be seen below:

E (Erzya)	41 968 words
M (Moksha)	39 996 words

The distribution of word classes is as follows:

Word class	E		M	
	Words	(%)	Words	(%)
noun (n)	23 401	(55.76 %)	19 095	(47.74 %)
verb (v)	10 068	(23.99 %)	13 296	(33.24 %)
adjective (a)	5 283	(12.59 %)	3 886	(9.72 %)
adverb (ad)	1 473	(3.51 %)	1 938	(4.85 %)
participle (pc)	1 497	(3.57 %)	1 482	(3.71 %)
descriptive word (de)	322	(0.77 %)	251	(0.63 %)
gerund (ge)	164	(0.39 %)	213	(0.53 %)
numeral (nu)	230	(0.55 %)	128	(0.32 %)
particle (pa)	158	(0.38 %)	221	(0.55 %)
pronoun (pr)	165	(0.39 %)	155	(0.39 %)
postposition (po)	140	(0.33 %)	142	(0.36 %)
interjection (in)	87	(0.21 %)	84	(0.21 %)
conjunction (co)	40	(0.09 %)	47	(0.12 %)
% total		102.53 %		102.37 %

The total percentage is more than 100 % because a number of words bear multiple word-class labels; e.g. n/a or a/ad.

The number of references to different sources:

m	(Мокшанско-русский словарь)	31 288
e	(Эрзянско-русский словарь)	26 835
P	(H. Raasonens Mordwinisches Wörterbuch)	11 258
z	(Вейсэ, башка, тешкс вельде)	6 183
c	(Марса, башка, китькскя вельде)	4 845
S	(Suomalais-ersäläinen sanakirja)	2 703
A	(Валонь ёвтнема валкс)	716
D	(Русско-мордовский словарь)	265
J	(Moksa-mordvin szójegyzék)	55
Q	(Versuch einer mokscha-mordw. Grammatik)	23

* * *

This dictionary is the result of collaboration between many people. The compilers want to express their gratitude to Elina Ahola, Alho Alhoniemi, Aleksandr Feoktistov, Eeva Herrala, Jack Rueter, Sirkka Saarinen, Milla Särmä, the graduates of the Philological Faculty of the Mordovian State University, as well as to the Finno-Ugrian Society, which has accepted the book for publication.