

On Linguistic Distance between Erzya and Moksha, Morphology

Abstract

The purpose of this article is to outline morphological facts about the two literary languages Erzya and Moksha, which can be used for estimating the distinctive character of these individual language forms. Whereas earlier morphological evaluations of the linguistic distance between Erzya and Moksha have placed them in the area of 90% cohesion, this one does not. This study evaluates the languages on the basis of non-ambiguity, parallel sets of ambiguity and divergent ambiguity. Non-ambiguity is found in combinatory function to morphological formant alignment, e.g. *молян* go+V+Ind+Prs+ScSg1. Parallel sets of ambiguity is found in combinatory-function set to morphological formant alignment where both languages share the same sets of ambiguous readings, e.g. *саиць* v s *сязовь* take+V+Ind+ScPl3+OcSg3, ScPl3+OcPl3. Divergent ambiguity is found in forms with non-symmetric alignments of combinatory functions, e.g. *саинек* take+V+Ind+Prt1+ScPl1, +Prt1+ScPl1+OcSg3, +Prt1+ScPl1+OcPl3 vs *сязоме* take+V+Ind+Prt1+ScPl1, *сязовьк* take+V+Ind+Prt1+ScPl1+OcSg3, +Prt1+ScPl1+OcPl3.

This morphological evaluation will establish the preparatory work in syntactic disambiguation necessary for facilitating Erzya↔Moksha machine translation, whereas machine translation will enhance the usage of mutual language resources.

Results show that the Erzya and Moksha languages, in the absence of loan words from the 20th century, share less than 50% of their vocabularies, 63% of their regular nominal declensions and 48% of their regular finite conjugations.

On ways to count mutual versus divergent dependent morphology

Two analyses will be given for the cohesion between Erzya and Moksha dependent morphology. One represents the approach of a language learner intent on maximizing the possible mnemonic means for learning two language forms at one time, while the other is interested in the prospects of automated morphological analysis, disambiguation and machine translation. Thus, the first approach features information from an article by Gabor Zaicz based on his presentation from CIFU 7: “*Beitrag zur Typologie und Statistik des erza- und mokscha-mordvinischen, (eine vergleichende Untersuchung)*”, whereas the second derives from the aspects of mutual non-ambiguity, parallel sets of ambiguity and divergent ambiguity within the scope of the two literary languages, Erzya and Moksha.

In his analysis of Erzya and Moksha dependent morphology, Gabor Zaicz (1990: 9-10) appears to approach Mordvin cohesion from the perspective of an interested language learner. Zaicz examines the morphological phenomena of verbal and nominal inflection from three perspectives: (1) verbs with modal/temporal, personal and derivational inflection; (2) nominals with case possessive marking and derivational inflection, as well as (3) infinitive, participle and verbal adverbs with derivational and case inflection. Each of the three perspectives is set off as a category

group for part-of-speech.

Zaicz examines the dependent morphology by distinguishing between verbal and nominal morphemes. In verbs a division is given between mood/tense, person and derivational formants, while in the declension of nouns and non-finite verbs the division falls between case, possessor indices and derivation.

In his examination of verbs, Zaicz maintains there are 17 mood/tense markers, 84 person markers and 21 derivational endings. The 17 mood/tense markers are divided into 4, which according to Zaicz, are readily identified as the same, 11 are recognized through etymological evaluation, and 2 are not identifiable. The 84 person formants are divided into the parallel values of 18 (readily identifiable), 64 (recognized through etymological evaluation), and 4 (not identifiable). Likewise the 21 deverbal derivational endings are divided into 3 (readily identifiable), 17 (recognized through etymological evaluation), and 1 (not identifiable). Thus the figures for cohesion between Erzya and Moksha verbal morphology are presented as 88.24% mood/tense, 95.24% person and 95.24% derivation for a mutual 92.91% of cohesion in Mordvin verbal morphology.

In his examination of nominal declension, Zaicz maintains there are 13 cases, 15 person markers and 22 derivational endings. The 13 case formants are divided: 7 (readily identifiable), 5 (recognized through etymological evaluation), and 1 (not identifiable). The 15 possessor indices: 3 (readily identifiable), 7 (recognized through etymological evaluation), and 5 (not identifiable). The 22 derivational endings: 13 (readily identifiable), 8 (recognized through etymological evaluation), and 1 (not identifiable). Thus the figures for cohesion between Erzya and Moksha nominal morphology are presented as 92.31% case, 66.67% possessor indices and 95.45% derivation for a mutual 84.81% of cohesion in Mordvin nominal morphology.

The only verbal forms (mood/tense, person or derivation) and nominal forms (case, possessor indices or derivation) identified in the article are the ones Zaicz deems to distinguish the two language forms. To come as close to examination of the same scope of phenomena, derivational suffixes will be left out of this work. Evaluation will be limited to the functions of mood, tense and person in verbs, as well as case and possessor indices in nouns, which for practical purposes should be readily identifiable.

There are discrepancies in figures presented by Zaicz that cannot be arrive at by enumerating phenomena by name. Whereas Zaicz speaks of 17 mood/tense markers, the enumeration of mood/tense functions is limited to 9 (cf. Tsyпкаikina, V. P. 2000: 187-190). In Erzya the number of mood/tense functions/markers can be enumerated as follows: (Ind+Prs) Indicative present,

(Ind+Prt1) indicative preterite 1, (Ind+Prt2) indicative preterite2, (Conj) conjunctive, (Des) desiderative, (Imprt) imperative and (Opt) optative, which is seven (7). The derivational conditional (Cond), which can be attested with three tenses (Cond+Prs, Cond+Prt1, Cond+Prt2), conjunctive (Cond+Conj) and subsequent derivation in “-кшномс”, for example, is not treated here, but its inclusion would raise the number to twelve. Therefore, it must be assumed that Zaicz's figures indicate allomorphic variation.

Three items will be presented: mutual lexicon; dependent nominal declension, and dependent finite conjugation. Mutual lexicon will be addressed with the largest known parallel corpus of Moksha and Erzya lexical material, the H. Paasonen Mordwinisches Wörterbuch (1990-1996). The nominal and verbal inflection, for work with automated analysis, will be divided into two sections: one dealing with combinatory functions for distinguishing unambiguous and ambiguous morphology within the individual languages. While nominal declension can be discussed utilizing a simple division: ambiguity versus non-ambiguity, finite conjugation involves a third division where there are parallel sets of combinatory functions, ambiguous in the same way for Erzya and Moksha. Here divergent ambiguity will mark the distinctive differences between the two languages. Combinatory functions will be set forth in the form of analysis tags, e.g. *стякшны* will be analyzed as *s_стякшномс_stand+V+Ind+Prs+ScSg3* and the finite verb combinatory function = *Ind+Prs+ScSg3*.

Rebuttal on lexicon

Depending on research practices the lexical cohesion of the two Mordvin language forms has been measured between 93% and 8% (1990-2015). Close affinity has been proposed with 92.42% for 116 words (Zaicz 1990 7-13), whereas the other extreme of the spectrum can be deduced from comparative lexical work as a mere 8.97% for 75,218 words (Luutonen & et al 2007). At a mid point between the 92% percent measured by Zaicz and the 9% attested in work by Luutonen, there comes the 27% – 45% span measured by Rueter (in press). Rueter provides a comparative analysis of the H. Paasonen dictionary of Mordvin dialects, attesting to a cohesion of 45% for 6952 mutual lexical roots and 27% for 21,754 derived word stems, see Table 1.

Affinity of the Erzya and Moksha vocabulary in the Mordwinisches Wörterbuch ⁱ					
Article sets	Erzya	Moksha	Union	Intersection	
				in numbers	Percent
Stem articles	14,395	13,122	21,573	5,944	26%
Root articles	5,302	4,807	6,952	3,158	45%

Table 1

(borrowed from Rueter: *Towards a systematic characterization of dialect variation in the Erzya speaking world, Isoglosses and their reflexes attested in and around the Dubyonki Raion.* in press)

The number of individual articles with attestation for both Erzya and Moksha representation of intersecting glosses falls short of six thousand; percentage-wise that means only a little over 26% intersecting Mordvinic vocabulary. A more feasible count might be elicited by tallying article sets, i.e. counting word groups with intersecting Mordvinic roots. In fact, there are 2,996 intersecting roots attested by the article structure of Heikki Paasonen's dictionary, which is 45% shared vocabulary. These figures, it will be noted, are based on the largest available data base for Mordvinic languages of over 21,500 lemmata, in a collection of 2703 pages of published dialect research.

Once the number of mutual lexical roots and stems has been outlined, the next step is to investigate the mutual distribution of word stems by part of speech. The examination of the H. Paasonen Dictionary materials is by no means complete, but the proportions of shared lexical material can be illustrated according to the results of a simple search. Here the 21,754 stem articles counted according to part-of-speech analysis attest to the highest cohesion in vocabularies at less than 80%, and the lowest at a little over 10%.

The distribution of cohesion between sets in the part-of-speech division of the vocabularies is not entirely without explanation. Whereas coordinate conjunctions, deemed as late loans in both languages, as well as numerals and determiners display a cohesion between 60% and 80%, proper nouns take us to a low of 11%. On the one hand, there are only 76 numerals, determiners and coordinate conjunctions total, while on the other hand, the 975 proper names can be seen as both toponyms and anthroponyms attesting to at least a diverse geographic habitation. The more extensively attested parts of speech, verbs, adjectives and common nouns attest to a mutual lexical cohesion of 26%, 27% and 28%, respectively, see Table 2 (Rueter, in print).

Affinity of Erzya and Moksha stem vocabulary in the Mordwinisches Wörterbuch by part-of-speech					
POS	Erzya	Moksha	\cap		U
A	1284	914	464	27%	1734
Adv	451	381	228	38%	604
CC	9	7	7	78%	9
CS	12	10	5	29%	17
Det	23	14	14	61%	23
Ger	16	20	6	20%	30
Interj	221	71	63	21%	299
N	5754	5037	2365	28%	8426

Num	37	35	28	64%	44
Pcle	61	42	20	24%	83
Po	109	82	71	59%	120
Prc	419	284	98	16%	605
Pron	86	55	42	42%	99
Prop	629	458	112	11%	975
Qnt	28	28	15	37%	41
V	5127	5371	2183	26%	8315
VGen	98	43	22	18%	119
Stem	112	75	24	15%	163
Misc.	48	18	18	37%	48
Total	14524	12945	5785	27%	21754

Table 2 (Rueter, in print)

Regular, dependent inflection

Examining the regular nominal and verbal inflection patterns of Erzya and Moksha, will provide an even more explicit illustration of mutual and divergent morphology. Since discussion with Professor Tsygankin has implied that derivational distinctions between Erzya and Moksha is quite extensive; Moksha tends to have a more extensive use for verbal derivation in both native and loan word stem, regular inflection here will be limited to case, number, definiteness and possessor indices strategies in nominals, on the one hand, and mood, tense, person and number categories in finite verbs, on the other.

Regular, dependent nominal declension

There is a distinct dichotomy in the dependent morphology of Erzya and Moksha in the nominal declension, i.e. Moksha attests to a symmetric system of function marking in the core cases of the indefinite, definite singular, definite plural and possessive paradigms, not attested in the Erzya counterpart. In the non-core cases, Moksha lacks the two definite declension categories for singular and plural. In Erzya, these categories are treated separately; due to the variation of case forms attested, whereas the definite singular can distinguish a maximum of 10 case, the definite plural may attest as many as 13 (see Rueter, 2010: 127-129).

In the core cases Erzya exhibits an inconsistency in the distinction for number of the possessa, and a virtual absence of dative marking for the first and second person possessors (see Table 3). The only dative formants for first and second person possessor are attested in the singular, and they are restricted to marking of kindred type terms (see Rueter 2010: 81-83, 112-127).

Divergence in core-case marking strategies of Erzya and Moksha												
Language	Case	Num	Indef	Def	PxSg1	PxSg2	PxSg3	PxPl1	PxPl2	PxPl3		
Erzya	Nom	Sg	∅	-Ось	-Ом		-ОзО					
		Pl	-Т	-ТНе	-ОТ		-ОнОк				-Онк	-Ост
	Gen	Sg	-Онъ	-Онть	-Он		-ОнзО					
		Pl		-ТНень	(-Ом)							
	Dat	Sg	-Нень	-Онтьень	-Нень	-Тень	-Онстэнь			-Онстэнь		
		Pl		-Тненень	[+Kin] ⁱⁱ	[+Kin]						
Moksha	Nom	Sg	∅	-Сь	-Озе	-це	-Оц	-Онъке	-Онте	-Сна		
		Pl	-Т	-ТНе	-Оне	-тне	-Онза					
	Gen	Sg	-Онъ	-ть	-Озен	-цен	-Онс	-Онъконъ	-Онтьень	-Снонь		
		Pl		-ТНень	-Онень	-тнень	-Онзонъ					
	Dat	Sg		-Онди	-Озенди	-ценди	-Онцты	-Онъконъди				
		Pl	-Нень	-Тненди	-Оненди	-тненди	-Онзонды		-Онтенди	-Снонды		

Table 3

The core-case paradigm in Table 3 can be read to have 37 combinatory functions for a mutual Erzya-Moksha bench mark. Erzya and Moksha share 10 combinatory functions, 4 in the indefinite declension and 6 in the definite declension. A divergence, however, can be attested in the possessor indices, where $6+6+6+3+3+3 = 27$ Moksha formants attest to 27 distinct combinatory functions that are only partially addressed by 11 Erzya formants. Erzya displays nearly total ambiguity in nominative versus genitive distinction and lacks dative expression in dependent morphology for most possessor indices cells.

Non-core case distribution for Erzya and Moksha dependent morphology exhibits a large gap in mutual paradigms. Moksha has no dependent morphology to indicate definite non-core case, whereas Erzya attests to 9-10 cases in the plural definite category and 7 cases in the singular definite category (see Rueter 2010: 74-104). Hence, there are no mutual combinatory functions for the 16 definite non-core case cells in the mutual pool of functions.

It is often maintained that morphological distinction between the four spatial cases, illative, inessive, elative and prolativ, in Erzya and Moksha is simply one of dependent versus independent case, or that the Erzya system of both synthetic and analytic expression of the four cases is merely an issue of synonymy (cf. Keresztes, L. 2011: 68-70). In fact, non-core cases in Erzya and Moksha will require further study in their individual coding of functions, e.g. the Erzya functions inherent in the 5 constituents *кудосо* 'house_N+SP+Ine+Indef', *кудосонть* 'house_N+Sg+Ine+Def', *кудонть эйсэ* 'house_N+Sg+Gen+Def эйсэ_Ро+Ine', *кудотнесэ* 'house_N+Pl+Ine+Def' and *кудотнень эйсэ* 'house_N+Pl+Gen+Def эйсэ_Ро+Ine' cannot be directly aligned with the 3 Moksha

Caus		Онкса												
13	12	11	10	10	10	10	10	10	10	10	10	10	10	10

Table 4

From a total of 73 combinatory functions in the non-core cases, there are 70 mutual to Erzya and Moksha, with only 3 affording a split between the two language forms.

From Tables 3 and 4, above, and the adjoining text, we can enumerate a total of 126 combinatory functions for the mutual Erzya-Moksha bench mark of case, number, definiteness and possessor indices categories. Of these 126 combinatory functions, 80 are shared unambiguously by both languages. Erzya attests to 16-17 definite cases functions completely lacking from Moksha dependent morphology. Moksha attests to 27 unique possessor indices functions while Erzya only sports 11 formants to mark these functions. These figures indicate a mutual dependent morphology of the Erzya and Moksha languages at 63%.

Regular, dependent finite verbal conjugation

Unlike regular nominal declension, the types of combinatory functions considered here for finite verbs, tense, mood, person and number, might be considered to be semantically more concise. For this reason the formulation of verbal tables will be more regular and predictable.

Here we will enumerate the instances of ambiguous and non-ambiguous tense / mood / person function combinations of verbs in the Erzya and Moksha literary languages. For morphological purposes the present, first preterite and imperative will be dealt with as a single unit; this is due to the parallel sets of ambiguities in these three paradigms.

The second preterite, conjunctive, desiderative and optative, however, will be afforded separate evaluation, which due to their enhanced morphological uniqueness. Comparison of combinatory functions and ambiguity will be made on the basis of totals for: shared non-ambiguity; parallel ambiguity patterns, and divergent ambiguity.

In the contemplation of the Erzya and Moksha verbal paradigms, a bench mark for paradigm internal homonymy should be established. Interference from external word forms might be minimized by choosing a mood with exceptional morphology, such as that found in the desiderative. Assuming that the Erzya *-Ыксал-* and the Moksha *-Олексал-* formants are unique to their respective languages, we are ready to embark on our first examination of tense/mood/person function ambiguity.

The desiderative, as is the case with other mood/tense paradigms, has 34 functions expressing

person and number of the subject and object arguments. Each of the two language forms, Erzya and Moksha, have a divergent number of formants involved in the equation. Erzya has 11 functions expressed by unambiguous morphology while the remaining 23 functions are expressed by 7 formants. At the same time, the 34 functions are represented in Moksha by a constellation of 12 unambiguous formants and 8 ambiguous to express the remaining 20 functions. In simple numbers that would appear to be a mere difference of one formant.

From a perspective of Erzya↔Moksha machine translation, a mutual Mordvinic perspective, however, the paradigms are much more divergent. In fact, the two languages share nine parallel, non-ambiguous formants (Table 5.1) and three parallel sets of ambiguous formants (Table 5.2) to handle 19 of the total 34 functions. This means that the remaining fifteen functions are represented by 6 formants in Erzya, and 8 in Moksha (Table 5.3).

Erzya-Moksha non-ambiguity, desiderative		
Compounded Functions	Formants	
	Erzya	Moksha
Des+ScSg1+OcSg2	- <i>Ыксэлитинь</i>	- <i>Олексолихтень</i>
Des+ScSg2+OcSg1	- <i>Ыксэлимик</i>	- <i>Олексолимайть</i>
Des+ScSg3+OcSg1	- <i>Ыксэлимим</i>	- <i>Олексолимань</i>
Des+ScSg3+OcSg2	- <i>Ыксэлинзеть</i>	- <i>Олексолензе</i>
Des+ScSg3+OcSg3	- <i>Ыксэлизе</i>	- <i>Олексолезе</i>
Des+ScSg3+OcPl3	- <i>Ыксэлинзе</i>	- <i>Олексолизень</i>
Des+ScSg3	- <i>Ыксэль</i>	- <i>Олексоль</i>
Des+ScPl2	- <i>Ыксэлиде</i>	- <i>Олексоледе</i>
Des+ScPl3	- <i>Ыксэльть</i>	- <i>Олексольхть</i>
9	9	9

Table 5.1

Shared non-ambiguity centers around the semantically non-ambiguous singular argument pairs and the subject conjugation for third and second persons plural. Instances of non-ambiguity that is not shared can readily be identified using the same criteria, i.e. Erzya has non-ambiguous forms for ScSg1+OcSg3 and ScSg2+OcSg3, while Moksha has non-ambiguous forms for ScPl1, ScSg1 and ScSg2, see Table 5.3, where non-ambiguous formants are indicated with the ⌘ sign.

Parallel sets of ambiguities involve the second person object where at least one of the arguments is in the plural. Thus when the subject argument of object conjugation is already in the plural as we observe with the second and third persons plural, we can only expect an ambiguity for number in the object, see Table 5.2. For an extensive discussion of homonymy in two argument agreement paradigms, see Trosterud (2007: 246-303).

Erzya-Moksha parallel sets of ambiguity, desiderative		
Compounded Functions	Formants	
	Erzya	Moksha

Des+ScP11+OcP12	<i>-Ыксэлидизь</i>	<i>-Олексоledязь</i>
Des+ScP11+OcSg2		
Des+ScSg1+OcP12		
Des+ScP13+OcP12		
Des+ScP13+OcSg2		
Des+ScSg3+OcP12		
Des+ScP12+OcP13	<i>-Ыксэлинк</i>	<i>-Олексолесть</i>
Des+ScP12+OcSg3		
Des+ScP13+OcP13	<i>-Ыксэлизь</i>	<i>-Олексолезь</i>
Des+ScP13+OcSg3		
10	3	3

Table 5.2

While Moksha has unique formants for all six of the subject conjugation slots, Erzya attests to non-ambiguity in object conjugation slots where both arguments are singular. Erzya subject conjugation formants for the first and second persons singular are homonymous to their object conjugation formants for third person plural object.

Erzya-Moksha divergent ambiguity, desiderative			
Compounded Functions	Formants		Ambiguity sets
	Erzya	Moksha	
Des+ScP12+OcP11	<i>-Ыксэлимизь</i>	<i>-Олексолемасть</i>	3
Des+ScP12+OcSg1			
Des+ScSg2+OcP11			
Des+ScP13+OcP11		<i>-Олексолемазь</i>	3
Des+ScP13+OcSg1			
Des+ScSg3+OcP11			
Des+ScP11	<i>-Ыксэлинек</i>	<i>-Олексолеме</i> ☞	1
Des+ScP11+OcP13		<i>-Олексолесък</i>	2
Des+ScP11+OcSg3			
Des+ScSg1	<i>-Ыксэлинь</i>	<i>-Олексоленъ</i> ☞	1
Des+ScSg1+OcP13		<i>-Олексолине</i>	1
Des+ScSg1+OcSg3			
Des+ScSg2	<i>-Ыксэлитъ</i>	<i>-Олексолетъ</i> ☞	1
Des+ScSg2+OcP13		<i>-Олексолитъ</i>	1
Des+ScSg2+OcSg3			
15	6	8	10

Table 5.3

In Table 5.3, Moksha has a more granular strategy for first person object marking. There also appears to be a retention of a distinction between subject conjugation first person plural and object conjugation first person plural subject with third person object. The priority of unique subject conjugation forms in Moksha, when aligned with their counter parts in Erzya, can be observed to

create checker of disalignment; due to the priority of singular argument patterns characteristic of the Erzya paradigm. In order to facilitate two-directional machine translation of verb forms in the desiderative paradigm, disambiguation will be required for 10 combinatory functions. Erzya requires a distinction for 8 combinatory functions: *ScSg2+OcPl1*, *ScPl2+OcSg1*, *ScPl2+OcPl1* vs *ScSg3+OcPl1*, *ScPl3+OcSg1*, *ScPl3+OcPl1* (second versus third person subject in first person object conjugation); *ScPl1* vs *ScPl1+OcSg3*, *ScPl1+OcPl3* (first person plural subject conjugation versus third person object conjugation with first person plural subject); *ScSg1* vs *ScSg1+OcPl3* (first person singular subject conjugation versus third person plural object conjugation with first person singular subject), and *ScSg2* vs *ScSg2+OcPl3* (second person singular subject conjugation versus third person plural object conjugation with second person singular subject). Moksha requires a distinction for 4 combinatory functions: *ScSg1+OcSg3* vs *ScSg1+OcPl3* (distinction for number in the third person object with first person singular subject argument), and *ScSg2+OcSg3* vs *ScSg2+OcPl3* (distinction for number in the third person object with second person singular subject argument).

If we count the mutual dependent non-ambiguity and parallel sets of ambiguity versus divergent ambiguity in the desiderative, we will arrive at a total of $9+3+10 = 22$ sets. By counting non-ambiguous and parallel matched ambiguity sets as 12 out of 22, we can arrive at a figure of 55% shared morphology in the desiderative, which for all practical purposes does not differ much from the 56% shared morphology for the desiderative when the corresponding total number of 34 functions, $9+10+15 = 34$, is applied.

Proceeding from the desiderative to the indicative second preterite, we can observe the aspect of paradigm external homonymy introduced by non-finite inflection.

The Indicative second preterite, like the desiderative, illustrated above, has 34 functions expressed by a divergent number of formants in the two Mordvinic literary languages of Erzya and Moksha. Erzya has 9 functions expressed with unambiguous morphology, but external interference sees the introduction of two new combinatory functions involving the short present participle in the singular and plural past tense of nominal predication. Hence the number of ambiguous combinatory functions rises to 27, which, in Erzya, are expressed with 9 ambiguous and 2 non-ambiguous formants. At the same time, Moksha addresses these 27 functions with 8 ambiguous and 7 non-ambiguous formants (see Tables 6.1-6.3).

Erzya-Moksha shared non-ambiguity, indicative second preterite		
Compounded Functions	Formants	
	Erzya	Moksha
Ind+Prt2+ScSg1+OcSg2	- <i>Блितинь</i>	- <i>Олихтень</i>

Ind+Prt2+ScSg2+OcSg1	- <i>Блимик</i>	- <i>Олимайть</i>
Ind+Prt2+ScSg3+OcP13	- <i>Блинзе</i>	- <i>Олизень</i>
Ind+Prt2+ScSg3+OcSg1	- <i>Блимим</i>	- <i>Олимань</i>
Ind+Prt2+ScSg3+OcSg2	- <i>Блинзеть</i>	- <i>Олензе</i>
Ind+Prt2+ScSg3+OcSg3	- <i>Близе</i>	- <i>Олезе</i>
Ind+Prt2+ScP12	- <i>Блиде</i>	- <i>Оледе</i>
7	7	7

Table 6.1

In Table 6.1 homonymy for Erzya third person subject conjugation reduces the number of shared non-ambiguous cells from 9 to 7. This leaves the singular argument object conjugation combinatory function prominent at the surface level.

Erzya-Moksha parallel ambiguity patterns, indicative second preterite		
Compounded Functions	Formants	
	Erzya	Moksha
Ind+Prt2+ScP11+OcP12	- <i>Блидизь</i>	- <i>Оледязь</i>
Ind+Prt2+ScP11+OcSg2		
Ind+Prt2+ScSg1+OcP12		
Ind+Prt2+ScP13+OcP12		
Ind+Prt2+ScP13+OcSg2		
Ind+Prt2+ScSg3+OcP12		
Ind+Prt2+ScP12+OcP13	- <i>Блинк</i>	- <i>Олесть</i>
Ind+Prt2+ScP12+OcSg3		
Ind+Prt2+ScP13+OcP13	- <i>Близь</i>	- <i>Олезь</i>
Ind+Prt2+ScP13+OcSg3		
10	3	3

Table 6.2

The same sets of parallel ambiguity can be attested in the indicative second preterite as were attested in the desiderative (see Tables 5.2 and 6.2).

Erzya-Moksha divergent ambiguity, indicative second preterite			
Compounded Functions	Formants		Ambiguity sets
	Erzya	Moksha	
Ind+Prt2+ScP12+OcP11	- <i>Блимизь</i>	- <i>Олемасть</i>	3
Ind+Prt2+ScP12+OcSg1			
Ind+Prt2+ScSg2+OcP11			
Ind+Prt2+ScP13+OcP11		- <i>Олемазь</i>	3
Ind+Prt2+ScP13+OcSg1			
Ind+Prt2+ScSg3+OcP11			
Ind+Prt2+ScP11	- <i>Блинек</i>	- <i>Олемеꞑ</i>	1
Ind+Prt2+ScP11+OcP13		- <i>Олесък</i>	2
Ind+Prt2+ScP11+OcSg3			

Ind+Prt2+ScSg1		-Олень☞	1
Ind+Prt2+ScSg1+OcPl3	-Блинь	-Олине	1
Ind+Prt2+ScSg1+OcSg3	-Блия☞		1
Ind+Prt2+ScSg2		-Олеть☞	1
Ind+Prt2+ScSg2+OcPl3	-Блить	-Олить	1
Ind+Prt2+ScSg2+OcSg3	-Блик☞		1
Ind+Prt2+ScSg3		-Оль☞	1
PrsPrc+Prt2+ScSg3	-Бль	-Йль☞	1
Ind+Prt2+ScPl3		-Ольхть☞	1
PrsPrc+Prt2+ScSg3	-Бльть	-Йльхть☞	1
19	8	12	14

Table 6.3

The external interference of the short present participle causes ambiguity in the Erzya side alone. Whereas the Moksha literary languages distinguishes four separate and unique combinatory functions, Erzya has two formants for coping with the functions Ind+Prt2+ScSg3, PrsPrc+Prt2+ScSg3, Ind+Prt2+ScPl3, PrsPrc+Prt2+ScPl3. It should be noted that the Erzya literary language recognizes three separate variants for the present participle (-Бл, -Блй, -Блия), and the Erzya short present participle -Бл has a low frequency in comparison with the longer and often contrastive present participle -Блия.

In the interest of two-directional compatibility for verb forms in the indicative second preterite paradigm, disambiguation will be required for 14 combinatory functions. Erzya requires a distinction for 12 combinatory functions: *ScSg2+OcPl1*, *ScPl2+OcSg1*, *ScPl2+OcPl1* vs *ScSg3+OcPl1*, *ScPl3+OcSg1*, *ScPl3+OcPl1* (second versus third person subject in first person object conjugation); *ScPl1* vs *ScPl1+OcSg3*, *ScPl1+OcPl3* (first person plural subject conjugation versus third person object conjugation with first person plural subject); *ScSg1* vs *ScSg1+OcPl3* (first person singular subject conjugation versus third person plural object conjugation with first person singular subject); *ScSg2* vs *ScSg2+OcPl3* (second person singular subject conjugation versus third person plural object conjugation with second person singular subject); *Ind+Prt2+ScSg3* vs *PrsPrc+Prt2+ScSg3* (finite verb conjugation versus participle with nominal conjugation), and *Ind+Prt2+ScSg3* versus *PrsPrc+Prt2+ScSg3* (finite verb conjugation versus participle with nominal conjugation). Moksha requires a distinction for 4 combinatory functions: *ScSg1+OcSg3* vs *ScSg1+OcPl3* (distinction for number in the third person object with first person singular subject argument), and *ScSg2+OcSg3* vs *ScSg2+OcPl3* (distinction for number in the third person object with second person singular subject argument).

The discussion of the indicative second preterite becomes problematic when we proceed to the conjunctive mood. While Erzya attests to a conjunctive mood formant in -вОл- and -вл-,

Moksha forms its conjunctive mood analytically with a combination of the indicative second preterite and the particle *ōa*. For this reason, the entire second preterite must be considered ambiguous on the Moksha side, whereas the ambiguity on the Erzya side will follow the pattern of the desiderative paradigm requiring distinction for 8 combinatory functions, see the discussion, above and Table 5.1-5.3.

If we were to assess the mutual dependent non-ambiguity and parallel sets of ambiguity versus divergent ambiguity in the indicative second preterite, we would arrive at a total of $7+3+14 = 24$ sets. By counting non-ambiguous and parallel ambiguity sets as 10 out of 24, we could arrive at a figure of 42% shared morphology in the indicative, which would be less than the 47% shared morphology for the same when the corresponding total number of 36 functions, $7+10+19 = 36$, is applied.

Despite the absence of a distinct synthetic conjunctive in Moksha, we are obliged to count the number of function sets requiring distinction in Erzya as 8 sets combinatory functions: *ScSg2+OcP11*, *ScP12+OcSg1*, *ScP12+OcP11* v s *ScSg3+OcP11*, *ScP13+OcSg1*, *ScP13+OcP11* (second versus third person subject in first person object conjugation); *ScP11* v s *ScP11+OcSg3*, *ScP11+OcP13* (first person plural subject conjugation versus third person object conjugation with first person plural subject); *ScSg1* v s *ScSg1+OcP13* (first person singular subject conjugation versus third person plural object conjugation with first person singular subject), and *ScSg2* v s *ScSg2+OcP13* (second person singular subject conjugation versus third person plural object conjugation with second person singular subject). Due to the absence of Moksha morphology, there is a 0% mutual morphology for 34 functions in the conjunctive.

The optative mood is a point of disparity between the Erzya and Moksha literary languages. In fact, even in the Erzya language where this mood might attest to a larger paradigm, little written evidence is available for its actual usage with all objects (cf. Tsyпкаikina, V.P. 2000: 168, 180-181). The Moksha literary norm only recognizes 4 forms. Here I have limited the optative to the most frequent 9 functions of the hypothetical 34 (see Table 7).

Erzya-Moksha ambiguity, optative		
Compounded Functions	Formants	
	Erzya	Moksha
Opt+ScSg1	- <i>Озан</i>	- <i>Озан</i>
Opt+ScSg2	- <i>Озат</i>	- <i>Озат</i>
Opt+ScSg3	- <i>Озо</i>	- <i>Озо</i>
Opt+ScP11	- <i>Остано</i>	
Opt+ScP12	- <i>Остадо</i>	
Opt+ScP13	- <i>Ост</i>	- <i>Ост</i>

Opt+ScSg3+OcSg3	- <i>ОссО</i>	
Opt+ScSg1+OcSg3	- <i>осса</i>	
Opt+ScSg3+OcSg2	- <i>Онзат</i>	
9	9	4

Table 7

On the basis of 9 distinct formant functions in Erzya, we can arrive at a total of 4 mutual functions, which indicates a 44% cohesion for Erzya and Moksha optative morphology. More work must be done on the documentation of this infrequent mood.

The indicative present, first preterite and imperative, as mentioned above, will be dealt with as one unit. This choice is due to the fact that combinatory function markers are attached directly to the verb stem without much detectable disambiguating morphology, on the one hand, and the fact that combinatory functions attest to ambiguous readings for formants of these categories. Therefore this unit will predictably consist of $34+34+10 = 78$ combinatory functions (see Table 8.1-8.3).

Erzya-Moksha non-ambiguity, indicative present, first preterite and imperative			
Non-ambiguous functions	Formants		
	Erzya	Moksha	
Ind+Prs+ScPl1	- <i>тАно</i>	- <i>тама</i>	
Ind+Prs+ScPl2	- <i>тАдо</i>	- <i>тада</i>	
Ind+Prs+ScSg1	- <i>Ан</i>	- <i>Ан</i>	
Ind+Prs+ScSg1+OcPl3	- <i>сынъ</i>	- <i>Сайне</i>	
Ind+Prs+ScSg1+OcSg2	- <i>тАн</i>	- <i>Хтя</i>	
Ind+Prs+ScSg1+OcSg3	- <i>са</i>	- <i>Са</i>	
Ind+Prs+ScSg2	- <i>Ат</i>	- <i>Ат</i>	
Ind+Prs+ScSg2+OcPl3	- <i>сыть</i>	- <i>Сайть</i>	
Ind+Prs+ScSg2+OcSg1	- <i>самак</i>	- <i>Самак</i>	
Ind+Prs+ScSg2+OcSg3	- <i>сак</i>	- <i>Сак</i>	
Ind+Prs+ScSg3+OcPl3	- <i>сынзе</i>	- <i>Сыне</i>	
Ind+Prs+ScSg3+OcSg1	- <i>самам</i>	- <i>Саманъ</i>	
Ind+Prs+ScSg3+OcSg2	- <i>тАнзат</i>	- <i>Хтянза</i>	
Ind+Prs+ScSg3+OcSg3	- <i>сы</i>	- <i>Сы</i>	
Ind+Prt1+ScPl3	- <i>сть</i>	- <i>сть</i>	
Ind+Prt1+ScSg1+OcSg2	- <i>Ытинъ</i>	- <i>ЙХтенъ</i>	
Ind+Prt1+ScSg2+OcSg1	- <i>Ымик</i>	- <i>Омайть</i>	
Ind+Prt1+ScSg3	- <i>сь</i>	- <i>Сь</i>	
Ind+Prt1+ScSg3+OcPl3	- <i>Ынзе</i>	- <i>Озенъ</i>	
Ind+Prt1+ScSg3+OcSg1	- <i>Ымим</i>	- <i>Оманъ</i>	
Ind+Prt1+ScSg3+OcSg2	- <i>Ынзеть</i>	- <i>Онзе</i>	
Ind+Prt1+ScSg3+OcSg3	- <i>Ызе</i>	- <i>Озе</i>	
Imprt+ScPl2	- <i>Одо</i>	- <i>Ода</i>	

Impprt+ScSg2	-ТЬ	-ТЬ	
Impprt+ScSg2+OcSg1	-Омак	-Омак	
25	25	25	

Table 8.1

There are 25 non-ambiguous functions mutual to Erzya and Moksha in the indicative present, first preterite and imperative.

Erzya-Moksha parallel ambiguity patterns, indicative present, first preterite and imperative		
Compounded Functions	Formants	
	Erzya	Moksha
Ind+Prs+ScP11+OcP12	<i>-тАдизь</i>	<i>-тядязь</i>
Ind+Prs+ScP11+OcSg2		
Ind+Prs+ScSg1+OcP12		
Ind+Prs+ScP13+OcP12		
Ind+Prs+ScP13+OcSg2		
Ind+Prs+ScSg3+OcP12		
Ind+Prt1+ScP11+OcP12	<i>-Ыдизь</i>	<i>-Одязь</i>
Ind+Prt1+ScP11+OcSg2		
Ind+Prt1+ScSg1+OcP12		
Ind+Prt1+ScP13+OcP12		
Ind+Prt1+ScP13+OcSg2		
Ind+Prt1+ScSg3+OcP12		
Ind+Prt1+ScP11+OcP13	<i>-сынек</i>	<i>-саськ</i>
Ind+Prt1+ScP11+OcSg3		
Ind+Prt1+ScP12+OcP13	<i>-сынк</i>	<i>-састь</i>
Ind+Prt1+ScP12+OcSg3		
Ind+Prt1+ScP12+OcP13	<i>-Ынк</i>	<i>-Ость</i>
Ind+Prt1+ScP12+OcSg3		
Impprt+ScP12+OcP13		
Impprt+ScP12+OcSg3		
Ind+Prt1+ScP13+OcP13	<i>-сызь</i>	<i>-сазь</i>
Ind+Prt1+ScP13+OcSg3		
Ind+Prt1+ScP13+OcP13	<i>-Ызь</i>	<i>-Озь</i>
Ind+Prt1+ScP13+OcSg3		
24	7	7

Table 8.2

Table 8.2 attests to 24 combinatory functions sharing 7 mutual sets, which is a slightly different array, due to the first person plural subject with third person object.

Erzya-Moksha divergent ambiguity				
	Formants		Ambiguity sets	
	Erzya	Moksha		
Ind+Prs+ScPl2+OcPl1	<i>-самизь</i>	<i>самасть</i>	3	
Ind+Prs+ScPl2+OcSg1				
Ind+Prs+ScSg2+OcPl1				
Ind+Prs+ScPl3+OcPl1		<i>самазь</i>	3	
Ind+Prs+ScPl3+OcSg1				
Ind+Prs+ScSg3+OcPl1				
Imprt+ScPl2+OcPl1	<i>-Омизь</i>		3	
Imprt+ScPl2+OcSg1				
Imprt+ScPl2+OcPl1				
Ind+Prt1+ScPl2+OcPl1	<i>-Ымизь</i>	<i>-Омасть</i>	3	
Ind+Prt1+ScPl2+OcSg1				
Ind+Prt1+ScSg2+OcPl1				
Ind+Prt1+ScPl3+OcPl1		<i>-Омазь</i>	3	
Ind+Prt1+ScPl3+OcSg1				
Ind+Prt1+ScSg3+OcPl1				
Ind+Prs+ScPl3	<i>-Ыль</i>	<i>-Йхть</i>	2	
PrsPrc+Pl+Nom+Indef		<i>-Отъ</i> ☞	1	
Ind+Prt1+ScSg2			<i>-к^v</i>	1
Imprt+ScSg2+OcPl3				1
Ind+Prt1+ScSg2+OcPl3		<i>-Ык</i>	<i>-йтъ</i>	1
Ind+Prt1+ScSg2+OcSg3			<i>-к</i>	1
Imprt+ScSg2+OcSg3			1	
Ind+Prt1+ScPl1	<i>-Ылек</i>	<i>-Оме</i> ☞	1	
Ind+Prt1+ScPl1+OcSg3		<i>-Оськ</i>	2	
Ind+Prt1+ScPl1+OcPl3				
Gen	<i>-Онъ</i> ☞	<i>-Онъ</i>	1	
Ind+Prt1+ScSg1	<i>-Ынь</i>		1	
Ind+Prt1+ScSg1+OcPl3				1
Ind+Prt1+ScSg1+OcSg3	<i>-Ыя</i> ☞	<i>-Йне</i>	1	
Ind+Prt1+ScPl2	<i>-Ыде</i>	<i>-Оде</i> ☞	1	
PrsPrc+SP+Abl+Indef		<i>-Йда</i> ☞	1	
Ind+Prs+ScSg3	<i>-Ы</i>	<i>-Й</i>	2	
PrsPrc+Sg+Nom+Indef				
PrsPrc+Sg+Nom+Indef+Contrast			<i>-Ыця</i> ☞	1
34	12	15	20	

Table 8.3

The divergent ambiguity in Table 8.3 is distorted due to 5 external combinatory functions including four forms of present participle and the verbal genitive or *-Онъ* participle. The final enumeration for indicative present, first preterite and imperative is 83 combinatory functions. Of

these 83 functions 49 are mutual, thus establishing a 59% value for mutual non-ambiguity.

An overview of finite verbal conjugation ambiguity for the desiderative, conjunctive, optative, indicative present, first preterite, second preterite and imperative speaks of 196 combinatory functions. Of these 196 combinatory synthetic functions, 94 are mutually attested for both Erzya and Moksha regular verbal finite conjugation, i.e. 48% mutual non-ambiguity or parallel ambiguity.

Regular dependent inflection of finite verbs and nouns establishes 322 combinatory functions with a total of 174 mutually non-ambiguous or of parallel ambiguity. This would establish the cohesion of regular nominal and verbal inflection for the Erzya and Moksha languages as 54%.

Work required in Erzya syntax include disambiguation in: core-case possessor indices declension; first preterite subject conjugation vs 3rd person object conjugation; present participle versus third person, and imperative versus indicative first preterite.

Work required in Moksha syntax include disambiguation in: definite non-core-case analytic versus analytic-synthetic strategies; number in 3rd person object conjugation; contrastive participle, and imperative versus indicative first preterite.

With an established point of departure in place for syntactic disambiguation, the Erzya and Moksha languages can now be addressed by long term collaboration in the newly established Erzya and Moksha language department. Many more years of productive research! Коштось свал паро улезэ!

Literature

Ananyina, K. I. 2000 = Ананьина, К. И. 2000:

Падеженди малады форматне. Мокшень кяль, Морфология. 73-75. *Вузонь мокшень и финно-угрань отделениянь тонафнихненди учебник. /Сёрматф-тиф Н.С. Алямкинен кядяла. – Саранск: Тип. «Крас. Окт.», 2000.–236 с.–Модов.-мокша яз.*

Keresztes, László 2011: *Bezetés a mordvin nyelvészetbe.* Debreceni Egyetemi Kiadó.

Luutonen, Jorma et al 2007: *Lexica Societatis Fenno-Ugricae XXXI:I. Electronic Word Lists: Mari, Mordvin and Udmurt with SFOu WordListTool 1.3 for PC and Mac.* Finno-Ugrian Society 2007.

Paasonen, Heikki 1990-1996 = *Lexica Societatis Fenno-Ugricae XXIII,1-4: H. Paasonens Mordwinisches Wörterbuch* zusammengestellt von Kaino Heikkilä, Bände I-IV, unter mitarbeit von Hans-Hermann Bartens, Aleksandr Feoktistow und Grigori Jermuschkin, bearbeitet und herausgegeben von Martti Kahla. (XXXI-LXXXVI). Suomalais-Ugrilainen Seura.

Rueter, Jack 2015: *Towards a systematic characterization of dialect variation in the Erzya speaking world, Isoglosses and their reflexes attested in and around the Dubyonki Raion.* (In press).

Tsyпкаikina, V. P. 2000 = Цыпкайкина, В. П. 2000:

Глаголось. *Эрзянь кель, морфология*, 187–190. *Эрзянь кель, морфемика, валонь теевема ды морфология: Вузонь эрзянь ды финнень ождеделениянь тонавтицятнень туртов.*

/Редколлегиясь: Д.В. Цыганкин (отв. ред., Н. А. Агафонова, М. Д. Имайкина ды лият. — Саранск: Тип. «Крас. Окт.». 2000. — 280 с. — Мордов.-эрзя яз.

Zaicz Gábor 1990: *Beitrag zur Typologie und Statistik des erza- und mokscha-mordvinischen*, (eine vergleichende Untersuchung). In A/5. *Volgai és permi nyelvészet.* Budapest/Uppsala (H/S). 7-13.

Abbreviations and tags:

A = adjective; Abe = abessive; Abl = ablative, Com = comitative; Comp = comparative; Dat = dative; Des = desiderative; Ela = elative; Gen = genitive; Ill = illative; Imprt = imperative; Ind = indicative; Ine = inessive; Lat = lative; Loc = locative; N = noun; Nom = nominative; OcPl1 = object conjugation first person plural; OcPl2 = object conjugation second person plural; OcPl3 = object conjugation third person plural; OcSg1 = object conjugation first person singular; OcSg2 = object conjugation second person singular; OcSg3 = object conjugation third person singular; Pl = plural; Prl = prolative; Prs = present; Prt1 = first preterite; Prt2 = second preterite; PXP11 = possessor index first person plural; PXP12 = possessor index second person plural; PXP13 = possessor index third person plural; PXSg1 = possessor index first person singular; PXSg2 = possessor index second person singular; PXSg3 = possessor index third person singular; ScPl1 = subject conjugation first person plural; ScPl2 = subject conjugation second person plural; ScPl3 = subject conjugation third person plural; ScSg1 = subject conjugation first person singular; ScSg2 = subject conjugation second person singular; ScSg3 = subject conjugation third person singular; Sg = singular; SP = singular or plural; Temp = temporalis; Tra = translative

- i The nature of the H. Paasonen Dictionary is etymological, therefore statistics presented are related to the etymology and not the semantics of the individual words. It should also be noted that the bulk of the dictionary is based on the collections organized by one man, which means that there are numerous words and word forms not attested. Of course, this dictionary is a dialect dictionary attesting to spoken language forms and not necessarily the written literary languages.
- ii Despite one attestation of a dative-case possessor index formant for the second person plural in a translation by Vasili Dyomin (2008, Кузька эрзянь паз), Erzya cannot be considered to have first or second person possessor indices for anything other than kindred type terminology. Therefore the Erzya combinatory functions *Dat+PxPl1* and *Dat+PxPl2* should be considered blank. Furthermore, the nature of semantics of words marked with the *Dat+PxSg1* functions would readily be inferred by the use of *Dat+Indef* if it were not for the *Dat+PxSg2* combinatory function.
- iii In order to minimize the amount of space taken by the table entire combinatory functions including case and possessor index are give without a hyphen preceding the segment, otherwise, it should be assumed that the archimorpheme case given in the indefinite category serves as the stem for the subsequent possessor index, indicated with a “+” sign.
- iv The comitative in Moksha does not occur in nouns with modifiers (cf. Ananyina, K. I. 2000: 74-75).
- v The *Imprt+ScSg2+OcSg3* and *Imprt+ScSg2+OcPl3* combinatory functions share the same form.