

Hunting for antiharmonic stems in Erzya

Fejes, László

Research Institute for Linguistics

†Hungarian Academy of Sciences

fejes.laszlo@gmail.com

Abstract

This paper presents a research study aimed to clarify whether there are antiharmonic stems in Modern Standard Erzya. At least two types of Erzya stems more or less liable to antiharmony were identified using the material of an Erzya dictionary and Internet search.

A cikk egy olyan kutatást mutat be, mely azt kívánta tisztázni, hogy vannak-e antiharmonikus tövek a mai sztenderd erzában. Egy erza szótár anyaga és internetes keresés alapján két olyan tőtípust sikerült azonosítani, mely kisebb-nagyobb mértékben hajlamos az antiharmóniára.

1 Introduction

In the literature on Erzya phonology, we usually find that Erzya has vowel harmony (Бондарко and Полякова 1993, 94–95; Keresztes 1990, 37; Keresztes 2011, 22–23; Bartens 1999, 66–67). However, suffix alternations due to harmony rather suggest that Erzya has vowel-consonant harmony. To our knowledge, the question of antiharmony in Erzya has never been discussed in the literature.

In Section 1, the notion of harmony, disharmony and antiharmony will be defined. The basic regularities of Erzya¹ harmony will be presented as well. It will be determined what kind of stems must be considered antiharmonic in Erzya.

In Section 2, a test on the material of an Erzya dictionary will be presented. Using a Perl script, word forms which show the symptoms of antiharmony were collected. Their antiharmonic behavior

¹In the following, the term *Erzya* should be understood as Modern Standard Erzya. From the point of view of harmony, Erzya dialects can strongly differ from each other.

was certified also by tests via Google². It will be defined in phonological terms what kind of stems may be antiharmonic.

In Section 3, the result of another test via Google will be presented: some stems, the antiharmonicity of which could not be tested automatically in the dictionary material, will be tested with some forms expectedly occurring on the Internet.

2 Harmony, disharmony and antiharmony in Erzya

2.1 Harmony, disharmony and antiharmony in general

Harmony is a phenomenon according to which in a given language, phonemes belong to two groups the members of which cannot occur together in a given domain (typically inside a word). These groups are usually divided by some phonetical feature. For example, in Finnish, Hungarian, Hill Mari or Turkish, front and back vowels do not typically occur in the same word form. However, there can be phonemes which do not belong to either of these classes, i.e. they can occur together with the phonemes of both harmonic classes inside the domain: these are called neutrals.

The most evident sign of neutrality is the behaviour of the phoneme in suffixes: neutrals do not alternate due to harmony. For example, in Finnish and Hill Mari, /i/ and /e/³ never alternate with other vowels due to vowel harmony. However, their behavior is different: Finnish neutral vowels are transparent (the backness or frontness of the vowel after them is identical with the backness or frontness

²One of my reviewers disapproves of the use of Google and warns me that I ignore (Kilgarriff, 2007). Nonetheless, since I do not deal with statistics based on Google data, Kilgarriff's criticism does not apply to this study. I use Google solely to find a given form and do not deal with its frequency.

³Also /i:/ and /e:/; since the length is never relevant, phonemes differing only in length will not be differentiated.

of the vowel before them), while Hill Mari neutral vowels are opaque (there must be a front vowel after them, due to the fact that they are front vowels). Neutrality can sometimes be gradual, as in Hungarian. Usually /i/⁴, /e:/ and /ɛ/ are considered to be neutrals. High /i/, despite some special exceptions, does not alternate in suffixes due to harmony; mid /e:/ alternates with /a:/ in approximately half of the suffixes (but not in the other half); low /ɛ/ practically always alternates with /ɒ/ or /o/ (and, in the latter case, also with /ø/). In a similar way, /i/ is always transparent (although two /i/s following each other morpheme-internally can also be opaque, c. f. /ɒli-nɒk/ ‘Ali-DAT’ but /ɒlibi-nɒk/ ~ /ɒlibi-nɛk/ ‘alibi-DAT’), /e:/ and /ɛ/ can be both opaque and transparent, although /e:/ is always transparent in some stems (/kɒʃte:j-nɒk/, but */kɒʃte:j-nɛk/ ‘manor-DAT’), and /ɛ/ is always opaque in some stems (/okto:ber-nɛk/, but */okto:ber-nɒk/ ‘October-DAT’). Turkish has no neutral vowels.

Forms containing phonemes that belong to different harmonic classes (such as Finnish /amatɔ:ri/ or Hungarian /ɒmɒtɔ:r/ ‘amateur’) are called *disharmonic*. Sometimes even forms with both back harmonic and phonetically front neutral phonemes phonetically (such as Finnish /kone/ ‘machine’ or Hungarian /lo:ve:/ ‘money (coll.)’) are called disharmonic.

In languages with harmony, stems containing only neutral phonemes take suffixes according to their phonetic value (e. g. stems with phonetically front neutrals take the front allophones of harmonic suffixes), e. g. Finnish /vede-s:æ/ and not */vede-s:a/ ‘water-INE’, Hungarian /vi:z-ben/ and not */vi:z-bɒn/ ‘water-INE’. However, there can be exceptions.

In Hungarian, there are also stems, which despite that they contain phonetically front neutral vowels, always take back variants of suffixes. These stems are called *antiharmonic*. Moreover, there is at least one class which is regularly antiharmonic: verbs containing /i:/, except for /tʃi:p/ ‘nip, peck, burn (food)’, always take the back allomorph of harmonic suffixes: /i:r-ok/ ‘write-1SG’ (c. f. /i:r-ɛk/ ‘Irish-PL’). Antiharmony is very rare with stems containing /e:/ (/tʃe:l-ok/, but */tʃe:l-ɛk/ ‘target-PL’; /he:j-ak/, but */he:j-ɛk/ ‘peel-PL’) or /ɛ/ (/ʃva(:)j-c-bɒn ~ /ʃvej-c-bɒn/ ~ /ʃvej-c-ben/ ‘Switzerland-INE’, /ʃpa(:)j-z-bɒn ~ /ʃpej-z-bɒn/ ~ /ʃpej-z-ben/ ‘pantry-

⁴Also /i:/; when the length is not relevant, the two phonemes will not be differentiated.

INE’). As the examples show, in the last cases there is an /a(:)/ : /ɛ/ alternation in the stem and that evokes vacillation in suffixation. Historically the case of /he:j/ ‘peel’ is similar, since it is etymologically identical with /hɔ:j/ ‘hair’.

Antiharmonic stems in Hungarian except for some marginal examples (/dere:k/ ‘waist’ : /dere:k-nɒk/ ‘waist-DAT’ : /derɛk-ɒt/ ‘waist-ACC’; /piʃil-ɛk/ ~ playfull (dialectal?) /piʃil-ok/ ‘pee-1SG’.

On the contrary, Finnish has no antiharmonic stems. In Hungarian, antiharmonic stems are always suffixed by the back variants of suffixes (and vacillating stems can be suffixed by both the back and the front allomorph of the suffix). There are no such stems in Finnish, but two stems are suffixed by an antiharmonic allomorph of one and the same inflectional suffix: /mer-ta/, but */mer-tæ/ ‘see-PART’, /ver-ta/, but */ver-tæ/ ‘see-PART’. The phenomenon is much more general in derivation, see e. g. /ki:t-os/ ‘thanks’ from /ki:t:æ-/ ‘to thank’, /itku/ ‘cry(ing)’ from /itke-/ ‘to cry, to weep’ etc. (c. f. Hakulinen et al. 2004, §16).

In general, we must conclude that antiharmonic stems are those stems which are not suffixed by the allomorphs we would expect by the regularities of harmony in the given language, but those allomorphs which are unexpected.

2.2 Harmony in Erzya

Erzya has five vowels: /i/, /e/, /a/, /o/ and /u/. Out of these only /e/ and /o/ alternate with each other due to harmony (see below), therefore, the other three must be considered neutrals. However, all the neutrals are opaque: if the vowel after them is a mid one, /i/ must be followed by /e/, /a/ and /u/ must be followed by /o/ – at least in suffixation. Nonetheless, these rules can be overridden by consonants.

In Erzya harmony, consonants also play a key role. Dental consonants can be arranged into non-palatalized vs. palatalized pairs: /t/ vs. /tʃ/, /d/ vs. /dʃ/, /s/ vs. /sʃ/, /z/ vs. /zʃ/, /tʃ/ vs. /tʃʃ/, /n/ vs. /nʃ/, /l/ vs. /lʃ/, and /r/ vs. /rʃ/. Stem-final palatalized consonants and the palatal /j/ trigger the use of front allomorphs of suffixes alternating due to harmony, independently of the quality of the last vowel in the stem. Therefore, Erzya “vowel harmony” should be considered vowel-consonant harmony.

With non-dental consonants, palatality plays no phonological role: however, phonetically they are palatalized in a palatal environment (before front vowels or palatalized dentals or /j/). According to

(Keresztes, 1990, 25) and (Keresztes, 2011, 18), all consonants other than dentals are alternated allophonically, although there are only labial and velar examples given. Barts (1999, 27) states that only labials and velars have palatalized allophones before front vowels. Имайкина (1996, 9) claims that alveolars /ʒ/, /ʃ/ and /tʃ/ are always “hard”, that is they are never palatalized.

Based on Keresztes (1990, 37) and Keresztes (2011, 22–23), the following suffixation types can be distinguished from the point of view of harmony:

- Both triggers and targets are vowels: /kudo-sonzo/ ‘house-INE-3SG’ : /vel^je-se-nze/ ‘village-S3-INE’;
- Both triggers and targets are consonants: /kal-t/ ‘fish-PL’ : /kal^j-t^j/ ‘willow-PL’;
- Triggers are vowels and targets are both vowels and consonants: /kudo-vtomo/ ‘house-ABE’ : /vel^je-vt^jeme/ ‘village-ABE’;
- Triggers are consonants and targets are both vowels and consonants: /kal-do/ ‘fish-ABL’ : /kal^j-d^je/ ‘willow-ABL’;
- Triggers are vowels but targets are consonants: /kudo-t/ ‘house-PL’ : /vel^je-t^j/ ‘village-PL’;
- Triggers are consonants but targets are vowels: /kal-so/ ‘fish-INE’ : /kal^j-se/ ‘willow-INE’.

It seems that sibilants and affricates (and consonant clusters including them) are never targets of harmony.

Since stem-internal (dis)harmony is not relevant from the point of view of antiharmony, it will be not discussed here. Let it be enough that there is no strict harmony inside stems.

2.3 Antiharmony in Erzya

The basic rules of Erzya suffixation due to harmony are quite simple: if the last vowel of the stem is front or the stem-final consonant is palatalized, the front (palatalized) variant of the harmonizing suffix must be chosen. In all other cases, the back (non-palatalized) variant of the harmonizing suffix must be chosen.

Since antiharmonic stems are those which do not choose the expected suffix allomorph, stems must be considered antiharmonic if:

- the last vowel of the stem is front and/or the stem-final dental is palatalized, but the stem is suffixed by the back (non-palatalized) variant of harmonizing suffixes;

- or the last vowel of the stem is back and the stem-final dental (if there is one) is not palatalized, but the stem is suffixed by the front (palatalized) variant of harmonizing suffixes.

The problem is how to find these stems (if they exist at all).

3 Looking for the antiharmonic stems

Since native speakers use their language unconsciously, they cannot elicit antiharmonic stems. However, they can notice that some stems are suffixed in an unexpected way when hearing slips of the tongue or when there are dialectal differences in the harmonic and antiharmonic suffixation of the same stems.

L2 learners are more probable to notice these irregularities when they try to suffix the stems due to the learned rules but native speakers consider the given form incorrect. However, they can meet suffixed forms from which they can learn which set of suffixes must be used with the stem without noticing these are irregular in a way.

By automatic parsing, it seems to be practically impossible to find antiharmonic stems in morphologically unanalysed texts. On the contrary, antiharmonic stems must pop up during the development of automatic morphological analyzers, since the suffixed forms of antiharmonic stems cannot be analyzed in the regular way. Although Erzya morphological analyzers exist,⁵ to our knowledge, no findings were reported on antiharmonic stems.

The research to be presented here was conducted on data from an Erzya–Hungarian dictionary (Mészáros and Sirmankine, 2003). The choice of the research material was determined by the fact that it was the only available material in an electronic format for the author (unfortunately, it is not public).

The original file was in Microsoft Document format. It was opened and saved in a html format. Using Perl scripts, just the headwords were kept, and they were Latinized (in a Setälä-like system) for a phonemic analysis.

Since Erzya harmonic suffixation is considered to be regular and predictable from the phonological form of the stem, dictionaries do not mark the

⁵At least Jack Rueter’s (see Rueter (2010), <http://giellatekno.uit.no/cgi/d-myv.eng.html>) and Timofey Arkhangelskiy’s (see Arkhangelskiy (2019), <https://bitbucket.org/timarkh/uniparser-grammar-erzya/src/default/>) are worth mentioning.

harmonic class of the stem (i.e. which set of allomorphs they are suffixed with). However, verbs always occur in a suffixed form, namely in the infinitive, which offers at least a restricted possibility to look for antiharmonic stems.

3.1 Erzya infinitive forms

Erzya infinitive forms have three typical endings: /-ams/, /-ems/ and /-oms/. However, there are different analyses for the morpheme boundaries in verb forms. According to Mészáros (1998, 33) (and similarly Pall, 1996, 20), in forms ending in /-ams/ /a/ is always the part of the stem. There are two types of verbs of which infinitive forms end in /-ems/ or /-oms/. In some stems, /e/ or /o/ belongs to the stem, in others it belongs to the suffix. There are some forms which show whether the vowel is the part of the stem or not: see Table 1.

meaning	‘stand, be’	‘understand’
INF	/aft ^j e-ms/	/tʃarkod ^j -ems/
PST.3S	/aft ^j e-s ^j /	/tʃarkod ^j -s ^j /
PST.3P	/aft ^j e-s ^j t ^j /	/tʃarkod ^j -s ^j t ^j /
IMP.2S	/aft ^j e-k/	/tʃarkod ^j -t ^j /

1. Table: Different morpheme boundaries for similarly looking infinitive forms

However, paradigms in Mészáros (1998, 15–16) show that not even /a/ is present in all verb forms (c. f. /kortams/ ‘speak:INF’ : /korti/ ‘speak:PRS.3S’), and it is also true for stems in which /o/ or /e/ are analyzed to be the part of the stem (e. g. /aft^jems/ ‘stand:INF’ : /aft^ji/ ‘stand:PRS.3S’, moreover /aft^jan/ ‘stand:PRS.1S’ and /aft^jat/ ‘stand:PRS.2S’, c. f. /kortan/ ‘speak:PRS.1S’ and /kortat/ ‘speak:PRS.2S’).

Bartens (1999, 122) states that there are stems with stem-final /a/, and this vowel-final stem is used in all the forms of the verb. On the contrary, all verbs with final /o/ or /e/ also have a consonant-final stem (the identical form but without the /o/ or /e/), which is used in more forms when the final vowel is /e/ (no examples presented). Bartens also adds that verbs with final /o/ or /e/ also have an /a/-final stem, used in the first and second singular of the present tense (see above).

Keresztes (1990, 39) (and similarly Keresztes, 2011, 76) states that all Erzya verbs, despite what kind of vowel the verb stem contains at the end of the vowel-final stem, also have a consonant-final stem. However, he says nothing on the distribution

of the two stem forms.

In the examples shown by Серебренников et al. (1993, 19–20), the morpheme boundary is sometimes before, sometimes after /e/ and /o/. Although there is a case in the introduction, in which the morpheme boundary is before /a/ (Серебренников et al. 1993, 17, 599: *сокламс-иза|мс* ‘till the soil’), it must be a typo since the morpheme boundary is at the expected place in the dictionary (Серебренников et al. 1993, 599: *сока|мс* ‘plow’, *сока|мс-иза|мс* ‘till the earth’).

The problem of the morpheme boundary can be crucial from the point of view of antiharmony. If we find an infinitive ending in /-ems/ where we expect /-oms/ (or vice versa), but the vowel belongs to the stem, it has nothing to do with antiharmony: at most we have to conclude that the stem is disharmonic. Unfortunately, grammars of Erzya do not help much in this case. They never state that stem types are completely lexicalized or (at least in some cases) they are determined by phonotactics (possible syllable structures). Nonetheless, based on the presented examples, it can be clear that stem types are at least partially lexicalized. Stem-final vowel is present in many cases when it is not necessary at all from the point of view of phonotactics: /vanoms/ ‘watch:INF’ : /vanos^j/ ‘watch:PRT.3S’ (Bartens, 1999, 129), /id^jems/ ‘save:INF’ : /id^jes^j/ ‘save:PRT.3S’ (Цыганкин, 1980, 277) etc. On the contrary, there are cases when epenthesis would be pertinent, however, the consonant-final stem is used: /kandoms/ ‘carry:INF’ : /kands^jt^j/ ‘carry:PRT.3P’ (Keresztes, 1990, 41), /t^jer^jd^jems/ ‘call:INF’ : /t^jer^jd^js^jt^j/ ‘call:PRT.3P’, /maksoms/ ‘give:INF’ : /makst/ ‘give:IMP:2S’ (Цыганкин, 1980, 277) etc. However, it still cannot be excluded that phonotactics can play some role at least in some cases. For example, Фейеш (2005) argues that although the epenthesis of /i/ in Komi-Zyryan verb forms depends on syllable structure in most of the cases, some borderline cases can be lexicalized, and other factors (e. g. the inner morphological structure of the stem) can also play some role.

3.2 Method and results

To find antiharmonic stems, we should check whether there are stems in which the last vowel before /-ems/ is back and the consonant immediately preceding it is not palatalized; or /-oms/ is preceded by a front vowel or a stem-final palatalized consonant. However, even in these cases we have to check

whether the /e/ or /o/ belongs to the suffix: if not, it does not say anything about the antiharmonicity of the stem; if yes, it is a clear sign of antiharmonicity. However, we cannot be sure that the stem will behave in an antiharmonic way with all the suffixes. However, the opposite is also true: if the infinitive shows a harmonic way of suffixation, it does not exclude that the stem is suffixed by the unexpected allomorph in some cases. Despite these restrictions, the method gives us possibility to find at least some antiharmonic stems, if they exist.

Based on these considerations, all the headwords ending in /oms/, /ems/ or /ams/ were collected from Mészáros and Sirmankine (2003) and they were counted according to the consonant before these endings by another Perl script.⁶ The results are presented in Table 2.

Consonant	/e/	/o/	/a/
Non-palatalized dental	18	713	216
Palatalized dental or /j/	778	0	357
Labial	102	153	108
Velar	19	19	50
Alveolar	10	6	47

2. Table: Infinitive endings with different stem final consonants

It is striking that palatalized dentals and /j/ are never followed by /-oms/. Moreover, non-palatalized dentals are overwhelmingly followed by /-oms/, and only exclusively by /-ems/. Although in all these cases the last vowel of the stem is /i/ or /e/, it seems that it is strongly predictable whether an infinitive ends in /-oms/ or /-ems/ just based on whether the consonant before it is palatalized or not. Most of the exceptions (where the last vowel is front but the consonant before /-ems/ is non-palatalized dental) can be divided into two groups:

- onomatopoetic words: /biznems/ ‘to bumble’, /vi3nems/ ‘to buzz, to whiz etc.’, /dirnems/ ‘to rattle, to crackle’ (~ /dʲirnʲems/ ‘to bumble, to bluster, to crackle’), /zeznems/ ‘to grizzle, to weep, to snivel; to moan, to grumble; to murmur, to mutter’, /irnems/ ‘to burr, to bellow’, /ki3nems/ ‘to rattle (human, animal)’, /kirnems/ ‘to snore, to rattle’⁸, /kitnems/ ‘to gig-

gle, to snicker’, /rʲiknems/ ‘to blub, to sob, to weep’, /ti3n-ems/ ‘to fizzle, to sizzle, to huss’;⁹

- ending in a sibilant or a cluster containing a sibilant: /pezems/ ‘to wash (head)’, /pivsems/ ‘to thresh out, to thrash out’, /rʲez-ems/ ‘to atrophy, to waste’, /rʲiznems/ ‘to grieve, to sorrow’.

On the one hand, onomatopoetic words tend to behave exceptionally in phonology cross-linguistically (c.f. Fudge 1970). On the other hand, as it has been shown above, dental sibilants tend to behave somewhat exceptionally in harmony: they do not undergo harmony in suffixes. The two phenomena seem to be related. The only case which cannot be categorized into either group is /pi3eldems/ ‘to be green’, which will be discussed below.

In the case of labials and velars, the situation is also simple. When the last vowel of the stem is front, the infinitive ends in /-ems/, when the last vowel is back, the infinitive ends in /-oms/. However, alveolars show a somewhat different picture. Most of the endings after alveolars with a last (and only) back vowel are /oms/: /vatʃoms/ ‘to get/be hungry’, /kuʃoms/ ‘to send’, /pan3oms/ ‘to open’, /pu3oms/ ‘to lose plant, to get dead (about flowers), to parch’, /uʃoms/ ‘to wait’, /tʃatʃoms/ ‘to be born’. However, there are four similar forms ending in /ems/: /lanʃems/ ‘to squat, to hunker’, /mantʃems/ ‘to fool, to cheat etc.’, /tokʃems/ ‘to touch, to poke’, /javʃems/ ‘to divide, to apportion, to distribute, to share out etc.’.

It seems that in all the last cases /e/ can belong to the stem. The vocabulary in Pall (1996), which contains information on whether the third person singular ending of the past tense form is attached to the vowel-final or to the consonant-final stem, contains only /mantʃems/ ‘to fool, to cheat etc.’. A Google search on the Internet gives forms like *ланчес(т)ь* ‘s/he hunkers (they hunker)’, *манчес(т)ь* ‘s/he fools (they fool)’, *токиес(т)ь* ‘s/he touches (they touch)’, *явиес(т)ь* ‘s/he shares out (they share out)’. On the contrary, forms like **ланчс(т)ь* ‘s/he hunkers (they hunker)’, **манчс(т)ь* ‘s/he fools (they fool)’, **токис(т)ь* ‘s/he touches (they

⁶NB! These are not always stem-final consonants, since the vowel before /ms/ may belong to the stem; in the case of /a/, according to some grammars cited above, it always belongs to the stem.

⁷C. f. /biznʲems/ ‘to go sour, to ferment (about milk)’.

⁸C. f. /kirnʲems/ ‘to contract, to shorten etc.’.

⁹One of my reviewers criticizes me for not distinguishing front [i] and centralized [ɨ]; they also state that in some cases the latter one occurs as a phoneme (not as an allophone of /i/ after non-palatalized dentals), even in some of the examples above. However, since centralization of front vowels, phonemic or allophonic, does not seem to affect harmony, I do not see the relevance of these facts here.

touch)', *явшс(т)ь 's/he shares out (they share out) etc.' are not attested. Nonetheless, it is easy to notice that in all these forms we would find such consonant clusters which are quite unusual even for Erzya. Moreover, if it was a case of inner disharmony in a stem, we would expect that it occurs more times and in more random phonological environments. Consequently, we cannot exclude that /e/ is epenthetic in these forms. If it is, then the stem must be antiharmonic. However, since we have no clear evidence for that, the question must be kept open.

It seems to be easy to find a phonetically based explanation for this ambiguous behaviour of alveolars. Alveolars are articulated near the palate, therefore they are near the palatal(ized) consonants. This could be a reason for them to behave as palatalized dentals in certain cases. However, it contradicts the claim by Имайкина (1996, 9) that alveolars are always "hard" (see Footnote 2.2).

3.3 Some other observations

The statistics presented in Table 2 show an unexpected tendency. As we have mentioned before, according to the literature, when the infinitive ends in /ams/, /a/ always belongs to the stem. This suggests that it has nothing to do with (morpho)phonology. However, statistics show that after harmonic consonants (dentals) harmonic vowels (/o/ and /e/) are much more frequent (more than two and a half times) than neutral (non-alternating) /a/. In the case of labials, harmonic vowels also prevail (it is less striking since numbers are divided between /o/ and /e/), although the range is a bit lower. However, in the other cases /a/ prevails: slightly in the case of velars and drastically in the case of alveolars (although the number of examples is almost negligible if we compare it to the number of cases with dentals). The question emerges, whether the choice between /a/ on the one side and /o/ or /e/ on the other side can be somehow conditioned. It may be worth comparing those cases when forms differ only in the vowel before /ms/.

As for /-ams/ vs. /-oms/, in two cases the two stem types are opposed: /kotʃkams/ 'to choose, to pick etc.' vs. /kofʃkoms/ 'to stub, to weed out', /palams/ 'to kiss' vs. /paloms/ 'to burn (up/down/away)' and in one case we have alternation: /tʃʲokordams/ ~ /tʃʲokordoms/ 'to sing (about nightingale)'.

In the case of /-ams/ vs. /-ems/, there is only one opposition: /puvorʲams/ 'to turn, to rotate,

to curl, to spin' vs. /puvorʲems/ 'to crust (over), to get knobby etc.'. Alternation is much more general: /ilʲtʲams/ ~ /ilʲtʲems/ 'to escort, to see off/out', /kemelʲdʲams/ ~ /kemelʲdʲems/ 'to sew ornaments etc.', /kengelʲams/ ~ /kengelʲems/ 'to lie', /ketʃkerʲams/ ~ /ketʃkerʲems/ 'to butt, to hurn, to stab etc.', /mendʲams/ ~ /mendʲems/ 'to bent, to bow, to crook etc.', /menstʲams/ ~ /menstʲems/ 'to make/set free, to rescue', /pupordʲams/ ~ /pupordʲems/ 'to bump, to falter, to stumble', /rʲeʒnʲams/ ~ /rʲeʒnʲems/ 'to go sour, to ferment (about milk)', /sanʲdʲams/ ~ /sanʲdʲems/ 'to grub up, to exterminate', /tolkanʲtʲa-ms/ ~ /tolkanʲtʲems/ 'to make oneself sweaty, to languish, to fatigue', /tʃemerʲdʲams/ ~ /tʃemerʲdʲems/ 'to press, to squeeze', /tʃirʲtʲams/ ~ /tʃirʲtʲems/ 'to twist, to warp etc.', /ezʲelʲdʲams/ ~ /ezʲelʲdʲems/ 'to lie'.

There is only one stem with an infinitive ending in /-oms/ and /-ems/: /piʒeldoms/ ~ /piʒeldems/ 'to be (vividly) green'. Since the last vowel of the stem, /piʒeldems/ is front, we would expect here the /-ems/ ending. Moreover, since the verb is derived from the adjective /piʒe/ 'green', we could also expect a palatalized form of the derivational suffix, something like */piʒelʲdʲ(-ems)/. However, although there are verbs ending in /lʲdʲ(-ems)/, there are no verbs derived from an adjective (or other word) with the suffix /-lʲdʲ-/.

On the contrary, there are some other verbs with the meaning 'to be coloured a certain colour' derived from the adjective meaning the colour by the suffix /ld/: /aʃoldoms/ 'to be (vividly) white' (← /aʃo/ 'white'), /oʒoldoms/ 'to be (vividly) yellow' (← /oʒo/ 'yellow'), /rauʒoldoms/ 'to be black' (← /rauʒo/ 'black'). In all these cases the basis of the stem contains back vowels.¹⁰

There are also some other, semantically farther derivations, all with back vowels: /nuzʲaldoms/ 'to be lazy, to laze, to idle' (←?, ~? /nuzʲaks/ 'lazy'), /kavtoldoms/ 'to doubt, to hesitate, to vacillate' (← /kavto/ 'two')¹¹, /gumboldoms/ ~ /kumboldoms/ 'to play in all colours of the rainbow' (←? /kumbo/ ~ /kumba/ 'carpet'). The last case, in all probability, is not a derivation, but speakers may feel a connection between colourful carpets and other things playing in all colours of the rainbow.

¹⁰There are some colour names with front vowels, but verbs meaning 'being those colours' are derived in a different way: (/senʲ/ 'blue' → /senʲeʒdʲ(ems)/ 'to be (vividly) blue', /jaksʲtʲerʲe/ 'red' → /jaksʲtʲerʲdʲ(ems)/ 'to be (vividly) red'

¹¹C. f. Hungarian két 'two', kételkedik 'to doubt'.

The alternation /pizeldoms/ ~ /pizeldems/ can be explained as follows. Since the vowels in the stem are front, the front allomorphs of the harmonic suffixes should be attached. However, stems ending in non-palatalized dentals are typically suffixed with back allomorphs of the harmonic suffixes. These can take an analogical effect, which can be strengthened by the fact that all morphosemantically analogical derived verbs are suffixed with the back allomorphs. The harmonic pattern and the lexico-semantic patterns are in conflict in this case, which results in vacillation. Moreover, in this case an Internet search shows that the vowel before /ms/ belongs to the suffix, not the stem: *нижелдсь*, **нижелдэсь*, **нижелдось* ‘be.green:PST.3S’; *нижелдсть*, **нижелдось*, **нижелдэсть* ‘be.green:PST.3P’; *нижелдт*, **нижелдэк*, **нижелдок* ‘be.green:IMP.2S’.

There is another similar case: /mazildoms/ ‘to be beautiful’ (← /mazi(j)/ ‘beautiful’). In this case, the dictionary (Mészáros and Sirmankine, 2003) does not contain the infinitive form with the front allomorph (/mazildems/). One should argue the reason in this case is that the last vowel is neutral or that it stands after a non-palatal dental, therefore, it is somewhat retracted.¹² This retractedness can weaken the triggering of the front allomorph, and it can be a reason that there is no vacillation, but the back allomorphs of harmonic suffixes are used in all cases. Notwithstanding, an online search shows that the form /mazildems/ exists in the Erzya part of the MarlaMuter dictionary (<https://marlamuter.org/muter/Эрзя/>) (but the same source does not know the form /mazildoms/). The same source states that the vowel does not belong to the stem, and the Internet search can prove it: *мазылдсь*, **мазылдэсь*, **мазылдось* ‘be.beautiful:PST.3S’ – however, other forms (PST.3P, IMP) are not attested and Pall (1996) does not contain this verb either.

Apart from /pizeld-/ and /mazild-/, there were no stems found in which the last vowel is front but the stem-final consonant is non-palatalized. Since other forms show that in these verbs the last vowel of the infinitive belongs to the suffix, forms with back allomorphs of harmonic suffixes must be considered antiharmonic, although the stems are not unequivocally antiharmonic, since both can also be suffixed by the front allomorphs. We can conclude that these

¹²In Mordvinic languages, similarly to Russian /i/, front vowels following non-palatalized consonants having a palatalized counterpart are somewhat retracted.

stems are alternatively antiharmonic.

Stems /lantʃ(e?)-/ ‘to squat, to hunker’, /mantʃ(e?)-/ ‘to fool, to cheat etc.’, /tokʃ(e-)/ ‘to touch, to poke’, /javʃ(e-)/ ‘to divide, to apportion, to distribute, to share out etc.’ can be considered antiharmonic if we can raise the possibility that the /e/ in them does not belong to the stem.

Based on the verbs, we have a clue what kind of stems we have to look for among other parts of speech if we want to find antiharmonic ones. These are the stems in which the last vowel is front, but the stem-final consonant is non-palatalized and stems in which the last vowel is back and the stem-final consonant is alveolar.

4 Testing dubious stems

Having determined what kind of stems can be antiharmonic by a higher probability, first we have to collect them. Again, another Perl script was used on the same material. The number of headwords with a front last vowel and a stem-final non-palatalized dental consonant is somewhat over 450, the number of headwords with a back last vowel and a stem-final alveolar consonant is over 60. However, not all of these are suffixable stems (moreover, adjectives and numerals are suffixed rarely).

Since dictionaries do not provide information on the suffixation of these stems (except for the occasionally occurring sample sentences, in a fragment of which the stem is suffixed with a harmonizing suffix), we have to test the existence of the forms by a search on the Internet. Considering the facts that Erzya material on the Internet is quite restricted;¹³ some of the tested words are rare; suffixed forms are usually rarer than unsuffixed ones and not all the suffixes are harmonic, and – because of lack of time and resources – not all the possible forms showing (anti)harmony can be tested, it is expectable that we will not find enough material in all the cases. However, the following forms were tested:

- form with the ablative suffix -/do/ : -/to/ : -/de/ : -/te/ : -/dʲe/ : -/tʲe/;
- form with the inessive suffix -/so/ : -/se/;
- form with the relative suffix -/sto/ : -/ste/;
- form with the abessive suffix -/tomo/ : -/teme/ : -/tʲeme/;¹⁴

¹³Moreover, since material on other languages is much more, sometimes they produce irrelevant search results which mask the necessary information.

¹⁴The abessive case also has suffix allomorphs -/vtomo/ : -/vtʲeme/, which are used after a stem-final vowel.

- forms with the 3S possessive suffixes: *-/ozo/* : *-/eze/* (for singular) and *-/onzo/* : *-/enze/* (for plural).¹⁵

Unfortunately, local suffixes are rarely used with stems meaning humans, and that fact makes testing even more difficult.

4.1 Stems with a stem-final alveolar

Since if the last vowel in the infinitive form is analyzed as the part of the suffix, two of our possibly antiharmonic verb stems */lanʃ-/* ‘to squat, to hunker’ and */manʃ-/* ‘to fool, to cheat etc.’ end in the cluster */nʃ/*, it would be worth starting the test with nominals ending in the same cluster. However, we found no such nominal stems. The case with */javʃ-/* ‘to divide, to apportion, to distribute, to share out etc.’ is similar. However, the fourth verb stem, */tokʃ-/* ‘to touch, to poke’ ends in a cluster */kʃ/*, which is quite frequent among the stems ending in an alveolar. None of these words were found to be systematically suffixed with front allomorphes. The only form with a stem-final */kʃ/* and a front suffix allomorph was */pokʃ-te/* along with several instances of */pokʃ-to/* ‘big-ABL’. However, based on this example, it is questionable whether any similarity of alveolars to palatalized consonants plays any role in similar cases: after a palatalized consonant we would expect the *-/tʃe/* allomorph of the suffix, *-/te/* occurs after stems ending non-palatalized stem final consonants and with a front last vowel.

Moreover, some other forms were also found with a stem-final alveolar consonant and an antiharmonic suffix: */urʲaʒ-enze/* besides several */urʲaʒ-onzo/* ‘brother’s.wife-PL.3S’; */etaʒ-se/* besides several */etaʒ-so/* ‘floor-INE’; */of-se/* besides several */of-so/* ‘town-INE’¹⁶ However, it seems that antiharmonic suffixation is very periferic and it is not clear whether it has any connection with the alveolar stem-final consonant.

¹⁵According to (Keresztes, 1990, 58) or (Bartens, 1999, 72), in these cases the first vowel belongs to the stem. However, it seems that this kind of segmentation is based on language history, and no synchronic facts support this kind of analysis. Since the given vowel is not predictable from the phonological construction of the stem (and not a lexicalized property either), it should be rather analyzed as epenthetic or belonging to the suffix.

¹⁶This case is debatable because in an earlier version of the text containing the form */of-se/* we find */saranskoj-se/* ‘in Saransk’ here, and it was changed to */saransk of-se/* ‘in the town of Saransk’. It can be a simple typo emerged during the redaction.

4.2 Stems with a stem-final non-palatalized dental

Due to the high number of stems with a front last vowel and a stem-final non-palatalized dental consonant, we could not test all the possible stem-suffix combinations. However, we tried to test at least the more frequent ones of the most typical groups.

One of the most important groups is that of Russian words used in Erzya texts. Antiharmonic forms seem to occur sporadically: */kɔncʲert-sto/* besides several */kɔncʲert-ste/* ‘concert-ELA’; */alfavit-sto/* besides an */alfavit-ste/* ‘alphabet-ELA’;¹⁷ */dokument-onzo/* besides several */dokument-enze/* ‘document-PL.3S’; */dokument-so/* besides several */dokument-se/* ‘document-INE’; */internet-so/* besides several */internet-se/* ‘Internet-INE’; */pitʲer-so/* (even twice) besides several */pitʲer-se/* ‘Sankt-Petersburg-INE’ etc.

We can find similar cases with native Erzya words as well: */velʲks-so/* besides several */velʲks-se/* ‘top-INE, above; (sour.)cream-INE’; */lʲezks-ozo/* besides several */lʲezks-eze/* ‘help-3S’; */pelʲks-ozo/* besides several */pelʲks-eze/* ‘part-3S’; */pelʲks-so/* besides several */pelʲks-se/* ‘part-INE’; */tʲefks-ozo/* besides several */tʲefks-eze/* ‘mark-3S’; */tʲefks-so/* besides several */tʲefks-se/* ‘mark-INE’ etc.

These forms are very rare and must be considered exceptional. It is worth noticing that some of the examples are from newspapers from the first half of the 20th century. Therefore, this phenomenon cannot be related to language loss. Мосин (2015, 24) shows some other examples when we find */o/* instead of Modern Standard Erzya */e/* after non-palatalized dentals in the newspapers of the twenties and thirties. According to him, these cases reflect different dialectal forms from the time when the linguistic norms were not settled. It is possible that antiharmonic forms even today reflect dialectal forms. Nonetheless, this means that at least in dialects these antiharmonic forms are possible. The only problem can be if all these forms are reflections of dialects in which – similarly to Moksha – we find a reduced vowel (schwa) instead of */o/* and */e/*. In this case, we have only an allophonic alternation. Since Moksha orthography uses *o* for the back allomorph of

¹⁷However, */alfavit-osʲ/* instead of */alfavit-esʲ/* ‘alphabet-DEF.NOM’ on the same site. Surprisingly, on the same site we also find the sentence *Те алфавитось теевьсь 1862 иестэ ды формовсь 1874 иесто* ‘This alphabet was constructed in year 1862 and got its final form in year 1874’, in which we find both */ije-sto/* and */ije-ste/* ‘year-ELA’. In these cases, the use of the back allomorph cannot be explained either by the vowels or the consonants of the stem.

schwa, occurring always after non-palatalized dentals, it is impossible to differentiate whether the reflected dialectal form contains /o/ or /e/.

5 Conclusion

We could identify two verbs, /pizeld-/ ‘to be green’ and /mazild-/ ‘to be beautiful’, which can be suffixed in an antiharmonic way (but harmonic suffixation is also possible). This phenomenon can be explained by the fact that stems with a final non-palatalized dental usually contain back vowels and take back allomorphs of harmonizing suffixes (and it is particularly true for verbs with a deajectival derivational suffix /-ld-/) and they take analogical impact on the suffixation of these stems.

Another group of verbs, /lantʃ(e)-/ ‘to squat, to hunker’, /mantʃ(e)-/ ‘to fool, to cheat etc.’, /tokʃ(e)-/ ‘to touch, to poke’, /javʃ(e)-/ ‘to share out etc.’ can be considered antiharmonic or (stem-internally) disharmonic depending on the morphological analysis. In this case, supposed antiharmony can be explained by the resemblance of alveolars to palatalized dentals (proximity of place of articulation). If these cases are treated as antiharmonic ones, these stems are the only consistently antiharmonic stems known.

We could not find any antiharmonic stems among nominals, but we found that stems with a last front vowel and a stem-final non-palatalized dental and stems with a last back vowel and a stem-final alveolar are suffixed in an antiharmonic way at least marginally. However, since no other kind of stems were tested, we cannot state that the degree this kind of antiharmonic behavior is higher than among stems with other kind of phonological structure (c. f. Footnote 17).

In any case, it must be emphasized that this research offers just a basic introspection to the problem of antiharmony in Erzya. A further study of the dialectal background is required and experiments should be made on the degree of acceptability of such forms.

In addition, although only in a marginal case, the results has relevance for the development of Erzya morphological analyzers. For example, it seems that the Giellatekno analyzer is based on the assumption that harmony is completely regular in Erzya. **The analysis of the forms *нижелдэмс*, *нижелдомс*, *мазылдэмс* and *мазылдомс*** suggests that the infinitive suffix is simply cut off and the harmonic class is generated based on the phonological structure

of the stem. Therefore, although the the *мазылд-* stem is generated based on the form *мазылдомс*, the analyzer does not recognize it. On the contrary, it can analyze the form *мазылдэмс*, but it is analyzed as a form of *мазылдомс*. Since its dictionary contains both *нижелдэмс* and *нижелдомс*, the form *нижелдэмс* is analyzed once as a form of *нижелдэмс*, once as a form of *нижелдомс*, but the form *нижелдомс* remains unanalyzed. These wierd results could be avoided if the determination of the harmonic class of the stem was based on the form of the infinitive suffix. However, it must be emphasized, that this anomaly exhibits only with these two (four?) verbs.

Acknowledgments

The research was financed by the NKFI 119863 Experimental and theoretical investigation of vowel harmony patterns. I would like to thank Péter Rebrus. I am particularly grateful for the assistance given by Nóra Wenszky and for Boglárka Janurik for the consultation on Erzya data. I am obliged to the reviewers for drawing my attention to certain important facts I left out of account. My special thanks are extended to György Soros for his help of founding the Department of Theoretical Linguistics of the Eötvös Loránd University.

References

- Timofey Arkhangelskiy. 2019. *Corpora of social media in minority uralic languages*. In *Proceedings of the fifth Workshop on Computational Linguistics for Uralic Languages, Tartu, Estonia, January 7 - January 8, 2019*, page 125–140. Association for Computational Linguistics.
- Raija Bartens. 1999. *Mordvalaiskielten rakenne ja kehitys*. Suomalais-Ugrilainen Seura, Helsinki.
- Erik Fudge. 1970. Phonological structure and ‘expressiveness’. *Journal of Linguistics*, 6(2):161–188.
- Auli Hakulinen, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen, and Irja Alho. 2004. *Iso suomen kielioppi*. Suomalaisen Kirjallisuuden Seura, Hämeenlinna.
- László Keresztes. 1990. *Chrestomathia morduinica*. Tankönyvkiadó, Budapest.
- László Keresztes. 2011. *Bevezetés a mordvin nyelvészetbe*. Debreceni Egyetemi Kiadó, Debrecen.
- Adam Kilgarriff. 2007. *Googleology is bad science*. *Computational Linguistics*, 33(1):147–151.

- Edit Mészáros. 1998. *Erza-mordvin nyelvkönyv kezdőknek és haladóknak*. JATEPress, Szeged.
- Edit Mészáros and Raisza Sirmankine. 2003. *Erza-mordvin–magyar szótár*. Savaria University Press, Szombathely.
- Valdek Pall. 1996. *Ersa keel. Õpiku konspekt ja sõnaloend*. [Valdek Pall], Tallinn.
- Jack Rueter. 2010. *Adnominal person in the morphological system of Erzya*. Suomalais-Ugrilainen Seura, Helsinki.
- Л. В. Бондарко and О. Е. Полякова. 1993. *Современные мордовские языки. Фонетика*. Мордовское книжное издательство, Саранск.
- М. Д. Имайкина. 1996. *Эрзянский язык. Учебное пособие для русскоязычных студентов. В 2 частях*. Издательство Мордовского Университета, Саранск.
- М. В. Мосин. 2015. [Отражение гласных и согласных фонем в текстах эрзянских газет 1920–1938 гг.](#) *Финно-угорский мир*, 23(2):22–30.
- Б. А. Серебренников, Р. Н. Бузакова, and М. В. Мосин. 1993. *Эрзянь-рузонь валкс*. <Русский язык>, Москва.
- Ласло Фейеш. 2005. [Эпентетическое ы в глагольных формах коми языка](#). In *История, современное состояние, перспективы развития языков и культур финно-угорских народов. Материалы III научной конференции финно-угроведов (1-4 июля 2004. г. Сыктывкар)*, pages 215–218. Институт языка, литературы и истории Коми научного центра УрО РАН.
- Д. В. Цыганкин, editor. 1980. *Грамматика мордовских языков. фонетика, графика, орфография, морфология*. Мордовский государственный университет имени Н. П. Огарева, Саранск.