Overapplication of Yer vocalization in Russian

Ben Hermans
University of Tilburg / CLS

1. Introduction

Yearley (1995) made a significant contribution to our understanding of Yer Vocalization in Russian by claiming that a yer vocalizes to eliminate a complex coda. Although this idea is appealing, the problem remains why Yer Vocalization sometimes overapplies. Interestingly, this instance of opacity cannot easily be explained in current versions of OO-Correspondence.

This article proposes the following solution: in an inflectional paradigm *in principle* any instance of the stem can be the Base in an OO-Correspondence relation. *In actual fact,* however, just one instance is capable of entertaining such a relation. Which stem instance is selected as the 'Designated Stem' is determined by the constraint system, established on independent grounds.

In Section 2 I give Yearley's analysis of Yer Vocalization; Section 3 shows that Yer Vocalization sometimes overapplies, and Section 4 sketches my solution.

2. The transparent phonology of Yer Vocalization; Yearley's account

Russian has two 'yers', vowels that alternate with zero. A few examples are:

(1)

front yer		back yer	
veter	'wind, nom. sg.'	lasok	'weasel, gen. pl.'
vetra	'wind, gen. sg.'	laska	'weasel, nom. sg.'

Yearley suggests that, underlyingly, yers are unlinked to a mora, whereas stable vowels are linked. The constraints in (2) determine when a yer is realized.

(2) a. Dep/Max-m: A mora in the output/input corresponds to a mora in the input/output.

Linguistics in the Netherlands 2002, **85–95.**
ISSN 0929–7332 / E-ISSN 1569–9919

b. Dep-V/Max-V: A vowel in the output/input corresponds to a vowel in the input/output.
c. NoComCod/Ons: A complex coda/onset must be avoided.

Dep-m must dominate Max-V. The effect is that it is better to delete a vowel than to insert a mora (in tableaux yers are represented with capital letters):

(3)	vetEra	Dep-m	Max-V
	vetera	*!	
☞	vetra		*

Dep-m also dominates NoComOns (dots represent syllable structure):

(4)	vetEra	Dep-m	NoComOns
	.ve.te.ra.	*!	
☞	.ve.tra.		*

In its turn Dep-m must be dominated by NoComCod, because insertion of a mora is preferred over the creation of a complex coda:

(5)	lasOk	NoComCod	Dep-m
☞	lasok		*
	lask	*!	

On the other hand, NoComCod must be dominated by Dep-V, because in forms lacking an underlying yer a consonant cluster in coda position is not split up by an epenthetic vowel. In the following tableau this is illustrated with the form *lask* 'caress', the gen. pl. of *laska* 'caress, nom. sg.'

(6)	lask	Dep-V	NoComCod
	lasok	*!	
☞	lask		*

The hierarchy we have established so far is summarized in (7):

(7) *The grammar of (Transparent) Yer Vocalization*
Dep-V » NoComCod » Dep-m » Max-V, NoComOns

As a result of this hierarchy, a yer is deleted (as in (3) and (4)), unless deletion creates a complex coda (as in (5)). Furthermore, forms lacking a yer do not receive one, not even at the cost of a complex coda (as in (6)).

There are two more constraints, regulating the interface with morphology:

(8) a. ANCHOR: If a segment is located at the right edge of a stem, then its correspondent is located at the right edge of a syllable.
b. ONSET: A syllable must have an onset.

To understand why ANCHOR is necessary, consider the following forms:

(9)

bašn'+a	'tower, nom. sg.'	bašen	gen. pl.
golov+k+a	'head, dim. nom. sg.'	golov+ok	gen. pl.

The first row in (9) shows that the morpheme *bašn'* has a yer, because it contains an alternating vowel. The second row proves that the diminutive must also have a yer, and the reason is the same. Now, when *bašn'* is combined with the diminutive we get a sequence of two consecutive yers in a row. Interestingly, only the first yer can be vocalized, not the second. The hierarchy in (7) cannot explain this, as is shown by the following tableau:

(10) bašEnOka	DEP-m	MAX-V
.baš.nlo.ka.	*	*
.ba.š enl.ka.	*	*

ANCHOR explains why the second candidate in (10) is preferred over the first; only in this candidate is the stem of the diminutive aligned with a syllable edge.

ANCHOR must dominate DEP-m. To see this, consider a word like *l'ubovnik* 'lover'. The suffix *-nik* is built on a stem containing a yer (*l'ubov'*, 'love, nom. sg.'; *l'ubvi,* gen. sg.). In principle the cluster *vn* is well formed in onset position, as shown by a form like *vnuk* 'grandson'. If *vn* is allowed in onset position, then we must ask why *l'u.bov.nik* is preferred over **l'ub.vnik*. This cannot be explained by NOCOMONS, because that constraint is dominated by DEP-m. This means that the correct result can only be obtained if ANCHOR dominates DEP-m, as is shown in (11):

(11) l'ubOvnik	ANCHOR	DEP-m	NOCOMONS
.l'u.bovl.nik.		*!	
.l'ub.vlnik.	*!		*

ANCHOR never takes effect before a vowel-initial suffix. A form showing this is *kogtistyj* 'sharp-clawed'. The root contains a yer, as is evident from the paradigm:

kogot' 'claw, nom. sg.'; *kogt'a* 'claw, gen. sg.'. It is followed by the derivational suffix *-ist* (which is followed by the case ending). That a yer never shows up before a vowel-initial suffix can be explained if Onset is ranked above Anchor. Due to this ranking a prevocalic consonant is always syllabified in the onset, making vocalization superfluous. This is illustrated in (12).

(12) kogOtistyj	Onset	Anchor	Dep-m	Max-V
.ko.go.tlis.tyj		*	*!	
☞ .kog.tlis.tyj		*		*
.ko.gotl.is.tyj	*!		*	

The constraint Anchor is thus ranked in the following way:

(13) *The grammar of the morphology-phonology interface*
Onset » Anchor » Dep-m

To summarize: as a result of the hierarchy in (7) a yer is realized to avoid a complex coda. We have also seen that there is a tendency to align a stem with the end of a syllable. This is explained by the hierarchy in (13). Let us now turn to the opacity problem.

3. The opacity problem

Sometimes a yer appears for reasons that are unclear from the perspective of the hierarchies in (7) and (13). The forms in (14) are representative. Opaquely vocalized yers are written in italics. Deleted yers are marked with '#'.

(14) a. basic nouns; a sequence of 1 underlying yer

veter	'wind, nom. sg.'	vet#ra	'wind, gen. sg.'
bugor	'knoll, nom. sg.'	bug#ra	'knoll, gen.sg.'

b. diminutivation; a sequence of 2 underlying yers

vet*e*rok	nom. sg.	veter#ka	gen. sg.
bug*o*rok	nom. sg.	bugor#ka	gen. sg.

c. iterative diminutivation; a sequence of 3 underlying yers

veter*o*ček	nom. sg.	vet*e*roč#ka	gen. sg.
bugor*o*ček	nom. sg.	bug*o*roč#ka	gen. sg.

The difference between transparent and opaque vocalization can be described in the following way: we expect that in a sequence of consecutive yers every vocalized yer is preceded by a deleted yer. What we get, however, is a non-alternating pattern: *all* yers preceding the vocalized one are vocalized. In the following three tableaux I show what exactly causes this difference.

(15) a. vetErOk	Onset	Anchor	Dep-m	Max-V
☜ .ve.tr|ok.		*	*	*
☞ .ve.te.r|ok.		*	**!	
.ve.terl.ok.	*!		**	

b. vetErOkEk	Onset	Anchor	Dep-m	Max-V
☜ .ve.terl.č |ek.		*	**	*
☞ .ve.te.r|o.č |ek.		**!	***	
.ve.terl.očl.ek	*!*		***	

c. vetErOkEka	Onset	Anchor	Dep-m	Max-V
☜ .ve.tr|očl.ka.		*	*	**
☞ .ve.te.r|očl.ka.		*	**!	*
.ve.te.r|o.č |e.ka		**!	***	
.ve.terl.očl.ka	*!		**	*

(15a) shows that **vetrok* and *veterok* only differ with respect to Dep-m and Max-V. Since Dep-m dominates Max-V **vetrok* should be optimal. (15b) shows that the actual winner, *veteroček*, is worse than the expected winner, **veterček*, on two accounts: it contains more violations of Anchor and Dep-m. Finally, (15c) shows that the real winner, *veteročka*, is worse than the expected winner, because they differ only under Dep-m and Max-V. The one performing best under Dep-m should therefore be the winner. This prediction, however, is not borne out.

I would like to point out that an account in terms of cyclic constraint evaluation does not provide a solution. A cyclic account always runs into problems when a yer is followed by a stable vowel. Consider again our example *kogtistyj* 'sharp-clawed' (cf. (12)). On the first cycle the yer would be vocalized in order to eliminate a complex coda, deriving *kogot'*. On the second cycle, where *-ist* becomes available, the vocalized yer can no longer be distinguished from a stable vowel. Since stable vowels do not alternate with zero, incorrect **kogotistyj* is derived. The fundamental shortcoming of the cyclic account is that it cannot look ahead to the next cycle. Yet, that is exactly what is required; if on cycle n+1 there is a vowel, then the yer on cycle n is not vocalized.

To summarize: in a sequence of consecutive yers all yers but the last are vocalized. From the perspective of the hierarchy in (7) this is unexpected, because

it predicts an alternating vocalization pattern. Due to lack of space I have illustrated the problem with the diminutive only. I would like to point out, however, that the opacity problem appears whenever the morphology creates a sequence of consecutive yers. For a list of relevant affixes I refer the interested reader to Yearley's article. In the next section I will propose a solution to the opacity problem.

4. A synthesis of symmetric and asymmetric OO-Correspondence

OO-Correspondence (Benua 1997) postulates that an output form, the Base, can stand in a correspondence relation with another, morphologically related, form. This relation is *asymmetric* in the sense that the Base determines the structure of the related form, whereas the related form cannot determine the structure of the Base. In the spirit of this theory we can postulate an OO-version of Max-V:

(16) Max-V(OO):
A vowel in the Base corresponds to a vowel in the morphologically related form.

Let us assume that we have the following relations: an inflectional stem *followed by an overt ending* is the Base for the same stem without overt inflection (for instance the nom. sg. of the 1st declension). Also, it is the Base for the same stem followed by a derivational suffix. On these assumptions we get the following relations (corresponding strings are underlined):

(17)

Base	Related form
veterka	veterok
veterka	veteročka
veteročka	veteroček

If the stem of *veterka* is the Base of *veterok*, then it is easy to explain the vocalization of the first yer in *veterok*; Max-V(OO) is ranked above Dep-m:

(18) I(nput): vetErOk B(ase): veterk…	Max-V (OO)	Dep-m	Max-V
.ve.trok.	*!	*	*
☞ .ve.te.rok.		**	

Given this ranking a (transparently vocalized) yer in the Base is copied by the morphologically related form. The same effect is obtained in the other pairs in (17). Thus, the vocalized yer of *veterka* is copied by *veteročka*. In its turn *veteročka* determines the vocalization pattern of *veteroček*; both its yers are copied. Notice that this pair shows that Max-V(OO) dominates Anchor:

(19)	I: vetErOkek B: veteročk…	Max-V (OO)	Anchor	Dep-m	Max-V
	.ve.trlo.č lek.	*!	**	**	*
	.ve.terl.č lek.	*!	*	**	*
☞	.ve.te.rlo.č lek.		**	***	

If the reverse order would hold the second candidate in (19) would be optimal.

I would like to stress that the correct results can only be obtained in the asymmetric model of OO-Correspondence. Only a yer in the Base is copied; the Base itself does not copy any yer. To give just one example: the nom. sg. copies the transparently vocalized yer of its Base; the Base, however, does not copy the transparently vocalized yer of the nom. sg. Although we seem to be on the right track, then, one fundamental problem remains to be solved. This concerns the definition of the notion Base.

Particularly suspect is the fact that the Base is defined as an inflectional stem followed by an overt ending. We must ask what is so special about a stem with an overt inflectional ending that it can function as Base. Not only is it conceptually suspect, it is also empirically wrong. In the third declension, for instance, there is an overt consonant-initial inflectional ending: *ju* (intr. case). When preceded by a yer the yer is vocalized. We thus get alternations like *.tser.kov.*', 'church, nom. sg.', *.tser.kvi.*, gen. sg., *.tser.kov'.ju.*, instr. sg. According to the definition alluded to above *tserkov'* qualifies as a Base, because it occurs before an overt inflectional ending, namely the instr. sg. Clearly, this leads to undesirable effects. Consider the word *.tser.kvuš.ka.* 'little church'. If *tserkov'* is a Base for the same stem followed by a derivational ending, then we predict that its vocalized yer is copied by the related form. We would thus predict **.tser.ko.vuš.ka*, rather than *.tser.kvuš.ka.* This, in fact, shows that an inflectional stem can only function as a Base if it is followed by a *vowel-initial* ending; neither the stem of the nom. sg. in the 1st declension, nor the stem of the instr. case in the 3st declension satisfies this condition.

What is so special about an inflectional stem followed by a vowel-initial ending that it functions as a Base? Quite clearly, it is not possible to build the notion 'followed by a vowel-initial suffix' directly into the definition of Base. We do not want phonological criteria ('vowel-initial') to be part of the definition of a morphological concept (Base), of course. That the facts seem to indicate that the Base acts like this should be derived in some indirect way.

To derive these effects we borrow certain ideas from symmetric OO-Correspondence, most explicitly developed in McCarthy (2001). This theory is symmetric in the sense that every instance of the stem of a given inflectional paradigm stands in a correspondence relation with every other instance of the same stem. Candidates

are pairs of corresponding stems. In every pair one member is the Base for the other member. Since any stem instance can be the Base for any other stem instance (of the same inflectional paradigm) the system is symmetric. Let us see how this would work in Russian.

Inflectional stems containing a yer in the final syllable have at least two surface realizations; one with a yer, occurring in a closed syllable, and one without a yer, occurring before a vowel-initial suffix. Two examples are given in (20). The first one is 'little wind', the by now familiar example of the 1st declension; the other example is *tserkov'* again, our 3st declension example.

(20)

Base			Related form	
a.	veterka	gen. sg.	veterok	nom. sg.
b.	veterok	nom. sg.	veterka	gen. sg.
c.	tserkvi	gen. sg.	tserkovju	instr. sg.
d.	tserkovju	instr. sg.	tserkvi	gen. sg.

The relation in (20a) gives the correct results. The yer of the Base, the stem *veterk,* is copied by the related form. In (20c) the Base does not have a yer. This is so because the cluster *kv* can be syllabified in the onset, due to the vowel-initial ending. Since the Base does not have a yer, there is nothing to be copied by the related stem. Notice, however, that the related stem has a transparent yer on its own right. The reason is that the ending starts with a consonant. Since the inflectional stem ends in a consonant Anchor can be satisfied. This constraint dominates Dep-m. Therefore, a yer is realized.

Although the relation in (20a) has nice results and the one in (20c) does not harm, the net result of symmetry is unsatisfactory. This becomes clear when we look at the two other pairs in (20). Maintaining the same rankings the effect would be that the transparently vocalized yer of the nom. sg. is copied by the stem of the gen. sg., leading to **veteroka.* Similarly, the yer in the stem of the instr. sg. would also be copied by the stem of the gen. sg., giving **tserkovi.*

Nonetheless, we wish to adopt the central idea of symmetric OO-Correspondence that any stem instance is relevant to any other stem instance of the same inflectional paradigm. However, strange though this might sound, we embed it in an asymmetric version of this theory. Central to our proposal is the following definition of Base:

(21) The Designated Stem
- a. Given a set of instances of an inflectional stem, *one and only one* instance is the Designated Stem (DS);
- b. DS is the Base in an OO-Correspondence relation;
- c. DS stands in an OO-Correspondence relation with any string whose morphology is a proper superset of DS's morphology.

The first clause declares that one particular instance of one and the same inflectional stem has a special status. The second clause declares that only this particular stem instance can be a Base in an OO-Correspondence relation. The third clause defines which forms are morphologically related to the Base; it says that all words are related that properly contain the Base (i.e. contain the Base plus some additional morphological information).

Notice now that nothing is said about which stem instance is the Base. In principle any stem can perform this function. However, in the end, one and only one stem instance (the Designated Stem) will be the actual Base. We thus get a kind of synthesis between symmetric and asymmetric OO-Correspondence. The idea that *any stem* instance can be the Base has a symmetric flavor. On the other hand, the fact that just *one stem* instance is actually selected as the Base is definitely asymmetric. Consider again the pairs in (20). *In principle*, the stem of *veterka* (*veterk*) can be the Base of the word *veterok*; the latter contains all the information contained in the former, but it also has more (it is specified as nom. sg.). The same also holds in the other direction. *In principle*, the stem of the word *veterok*, which is *veterok* as well (but without the specification 'nom. sg.'), can be the Base of the related word *veterka*; the latter contains all the information of the stem *veterok*, but it also has more (it is specified as gen. sg.). Similarly, *in principle* the stem of *tserkvi* (which is *tserkv'*) can be the Base of *tserkov'ju*, because the latter properly contains the former. Again, the relation also holds in the other direction; *in principle* the stem of *tserkov'ju* (which is *tserkov'*) can be the Base of the word *tserkvi*, because the latter properly contains the former.

Which principle decides that of all the instances of the stem of an inflectional paradigm just one particular instance will be the actual Base, the so called Designated Stem? There is only one non-arbitrary answer: it is decided by the system of ranked constraints. To this end, the constraint system, established on independent grounds, must be able to evaluate pairs of Base-Output forms, where, *in principle*, any stem instance of a given inflectional paradigm can be the Base. Let us go through the relevant cases to see how this works.

An inflectional stem which contains a yer in its underlying representation has two surface manifestations; one in which the yer is realized, and a second where the yer is not realized. A representative example can be found in the *tserkov-* paradigm. The stem instances are: *tserkv* and *tserkov'*. Both, then, can be the DS. But which one will actually succeed? Consider the tableau in (22). Notice that there are two forms in the upper box on the left. These are the two underlying forms corresponding to the two members of the pairs under evaluation. Assume, furthermore, that the first member of each pair is the Base for the related form on its right. It should also be noticed that the tableau keeps track of the violation marks of both members of a pair. In each cell the marks of the two members are separated by a comma.

(22)	tserkOvi (gen.sg.) tserkOvju (instr. sg.)	Max-V (OO)	Anchor	Dep-m	Max-V
☞	.tser.kv|(i) → .tser.kov|.ju		*,	,*	*,
	.tser.kov|.(ju) → .tser.ko.v|i		,*	*,*!	

In this tableau I have only considered pairs that satisfy Max-V(OO). This is to make explicit that the Designated Stem can (and should) be determined by means of the constraint system established on independent grounds (recall from (7) and (13) that Anchor » Dep-m » Max-V). The tableau shows that of the two stem instances *tserkv* and *tserkov'*, both appearing in the same inflectional paradigm, the former will be the Designated Stem. If this stem instance is chosen, it is possible to satisfy Max-V(OO), while at the same time maximally satisfying the rest of the constraint hierarchy. In this way we can account for the fact that of all the instances of an inflectional stem, whose final syllable contains a yer, only the yer-less instance can act as the Designated Stem. Only this instance, then, can be the Base in an OO-Correspondence relation. This is the result we want to have: only the stem instance occurring before a vowel-initial inflectional suffix can act as the Base.

Let us now see how we can explain the opaque cases. We start with *veterok*.

(23)	vetErOk (nom.sg.) vetErOka (gen. sg.)	Max-V (OO)	Anchor	Dep-m	Max-V
☞	.ve.ter|.k(a) → .ve.te.r|ok.		,*	*,**	*,
	.ve.tr|ok. → .ve.tr|o.ka		*,*!	*,*	*,*

We have seen before (tableau (15a)) that in a sequence of two consecutive yers in word final position we expect vocalization of the second yer; we thus expect **vetrok*, which is phonologically transparent, rather than the attested *veterok*. Another transparent instance of the same inflectional stem (i.e. a stem participating in the same inflectional paradigm) is *veterka*, where the inflectional stem is underlined (cf. tableau (10) for the argument that this stem instance is transparent). Consequently, the stem instances **vetrok* and *veterk* compete for DS-ship. In the tableau in (23) I show that the latter is a better Designated Stem. This follows from the ranking established on independent grounds. Here it is crucial that Anchor dominates Dep-m. We have established this independently on the basis of *l'ubovnik* (tableau (11)).

Since the stem instance *veterk* is a DS, the instance **vetrok* cannot be a DS. This is a consequence of the fact that among all the stem instances of the same inflectional paradigm only one instance can function as DS. The fact that *veterk* is the DS entails that its yer is copied by the nom. sg. This explains why we get *veterok*. Since

the stem **vetrok* is not a DS its transparently vocalized yer is not copied. This explains why **veteroka* or **vetroka* are ill formed.

Since the stem instance v*eterk* is a DS it can act as a Base for further morphological derivation. This explains why *veteročka* (which contains an additional diminutive) copies its yer. We thus get *veteročka* rather than **vetročka* (cf. (15c)). In its turn *veteročk* competes with **veterček* for DS-ship, because both are stem instances of the same inflectional paradigm: **veterček* is the maximally transparent stem of the nom. sg. and *veteročk* is the maximally transparent stem of the gen. sg. For reasons identical to the ones given in (23) *veteročk* is the DS, not **veterček*. This explains why *veteročk* forces **veterček* to copy its yers, so that we get *veteroček*.

The effect of all this is that in a series of consecutive yers all yers are vocalized, with the possible exception of the last, which is deleted or vocalized, depending on what follows. This, then, is our explanation of overapplication.

An informal characterization of this explanation could be as follows. We can think of a series of consecutive yers as a series of overlapping *pairs* of yers. The first yer of each pair is vocalized by transparent phonology and carried over to the next pair by OO-Correspondence.

5. Conclusion

In this article I have tried solve the problem of overapplication of Yer Vocalization. I have proposed a theory of OO-Correspondence that is neither entirely symmetric, nor asymmetric. From asymmetric theory it borrows the idea that there is a unique Base determining the phonology of related forms. From symmetric theory it borrows the idea that in principle any instance of the stem of an inflectional ending can be a Base. Whether a potential Base will be an actual Base is determined by the system of ranked constraints of the language at hand.

References

Benua, Laura (1997) Transderivational identity: phonological relations between words. PhD Dissertation, University of Massachusetts, Amherst.

McCarthy, John (2001) 'Optimal Paradigms'. Ms. University of Massachusetts, Amherst. ROA 485–1201.

Yearley, Jennifer (1995) 'Jer vowels in Russian'. In J.N. Beckman, L. Walsh Dickey and S. Urbanczyk, eds. *Papers in Optimality Theory*. Occasional Papers in Linguistics 18. GLSA, University of Massachusetts, Amherst, 533–71.