The Inflectional Base(s) of the Russian Imperative*

Andrei Antonenko
Stony Brook University

1. Introduction

Reference grammars of Russian and previous literature on the imperative differ in what they take to be the base form for imperative formation. The researchers postulate that the imperative is formed from the basic stem of the verb (Jakobson 1948, Townsend 1980), from the present (non-past) tense stem (Vinogradov 1972:464-5), from the 3PL form (Zaliznjak 1977:89), from the non-past stem together with the 1SG form (Švedova 1982:620-1) among others.

There are two main goals of this paper. The first is to test Albright's (2002) claim that inflectional paradigms may be derived from an existing paradigmatic form (inflectional base). I evaluate various verbal forms as potential bases for imperative formation in Russian in order to determine which form serves as the best base for the imperative. In order to do so, I apply the Minimal Generalization Learner (MGL) model (Albright and Hayes 1999) to derive the imperative form of the verb from other verbal forms, and compare the results based on number of characteristics, including the percentage of imperatives derived correctly.

Under the assumption that learning is facilitated by access to more information, the second goal is to propose and test an extension of the Albright and Hayes' learnability model which allows it to take into consideration multiple bases. I suggest a way in which multiple bases can be incorporated into the model, and test this scenario on the Russian data to see whether it significantly increases the number of correctly derived imperatives. I then consider the predictions made by the MGL for the acquisition of the imperative forms in Russian and look at some acquisition errors.

* I am grateful to Christina Bethin, Ellen Broselow, Bob Hoberman, as well as to the audience of FASL 19 and two anonymous reviewers, for their comments on this paper and/or previous versions of it. Needless to say, all errors are mine.

2. Imperative formation: basic patterns

In the Russian verbal conjugation the present tense is inflected in two dimensions: number (singular and plural), and person (1st, 2nd, and 3rd) and there are no gender differences. Because the present tense forms serve as the future tense in perfective verbs, the "present tense" paradigm is often referred to as NON-PAST.

Russian verbs consist of a root, followed by an optional verbal suffix and by inflectional affixes. The root is a minimal meaningful morpheme, incorporating the main semantic features of the verb; the stem is the root together with the verbal suffix. All other forms of the verb are derived by attaching inflectional endings to stems. Traditionally Russian is said to have two stems for each verb. The *non-past stem* (NON-PAST), is used in the formation of the present tense paradigm of imperfective verbs, the simple future of perfective verbs, the present, and imperatives among others. The *infinitive-past stem* (INF-PAST), is used in the formation of the past tense, and the infinitive, among others.

In order to form the imperative, a speaker needs to choose between two allomorphs: -∅ and *-i*, depending on prosody and the sonority sequencing of consonant clusters: *-i* is preferred if the stem ends in a consonantal cluster of rising sonority, e.g *krʲíknʲ-i* 'shout', and if the stem is not stressed in the 1SG non-past form, e.g. *pʲiš-í* 'write'. Timberlake (2004) distinguishes several different stress patterns in the NON-PAST tense: stress is either fixed in all forms of the paradigm (1a), or shifts between the 1SG inflectional affix and the syllable preceding the thematic vowel / inflectional ending in other forms (1b).

(1) a. Fixed stress pattern: *plákatʲ, pláč-u, pláč-eš* 'cry-inf, 1SG, 2SG'
b. Mobile stress pattern: *pʲisátʲ, pʲiš-ú, pʲíš-eš* 'write-inf, 1SG, 2SG'

The existence of the mobile stress pattern (1b) is crucial: the position of the stress in the 1SG non-past form and the infinitive in general cannot be predicted from the other forms in the paradigm. Therefore, the learner needs access to the 1SG form in order to choose the correct allomorph: if the 1SG form of the verb is stressed on the stem, then the imperative allomorph is -∅ if this choice does not result in a violation of the

Sonority Sequencing Principle (SSP), as in (2a). [1] Otherwise, the *-i* suffix is chosen, as in (2b).

(2)		***Infinitive***	***1SG***	***Imperative***	***Gloss***
	a.	ob^{j}íd^{j}et^{j}	ob^{j}íž-u	ob^{j}íd^{j} (*ob^{j}íd^{j}i)	'offend'
	b.	kr^{j}íkn-ut^{j}	kr^{j}íkn-u	kr^{j}íkn^{j}i (*kr^{j}íkn^{j})	'shout'

While the choice of the imperative allomorph can be made based on the 1SG form, this form is not sufficient to choose the segmental content of the base of the imperative form, (3).

(3)			***1SG***	***Imperative***	***2SG***	***Gloss***
	a.	i.	l^{j}ubl^{j}ú	l^{j}ub^{j}í	l^{j}úb^{j}iš	'love'
		ii.	skobl^{j}ú	skobl^{j}í	skobl^{j}íš	'scrape'
	b.	i.	l^{j}ečú	l^{j}et^{j}í	l^{j}et^{j}íš	'fly'
		ii.	l^{j}ečú	l^{j}ečí	l^{j}éčiš	'heal'

In the examples (3a-i) and (3b-i) the final stem consonants in the 1SG form and the imperative differ: *bl^{j}* vs. *b^{j}* in (3a-i) and *č* vs. *t^{j}* in (3b-i). Such consonantal alternations ("mutations") are not predictable from the 1SG form. Final stem consonants in the 1SG of the verbs in (3a-ii) and (3b-ii) are preserved in the imperative, and do not undergo alternations. For verbs in (3b) the necessary information about consonantal alternations is available in the 2SG form.

Instead of deriving the imperative form from some other existing forms in the verbal paradigm, Jakobson (1948) proposes that all forms of the verbal paradigm can be derived from a single verbal base. This presupposes that speakers can somehow establish a basic stem for every verb. The main problem with an analysis of imperatives based on the notion of a basic stem of the verb is the problem of learnability. Even though derivation of the imperative from the basic stem is straightforward, identifying the basic stem itself is not, and this task is not necessarily easier than deriving the imperative from the other basic forms of the verb directly.

---

[1] Sequences violating the sonority sequencing principle are often allowed in Russian, especially in the nominal paradigm, e.g *rúbl^{j}* 'ruble', *t^{j}ígr* 'tiger' but they are banned in imperatives.

3. Choice of the base for Russian imperatives

In this section I will test the Minimal Generalization Learner (MGL), originally developed by Albright and Hayes (1999) and later described in Albright 2005. This is a rule-based model of rule discovery, which analyzes dependencies between one form in the paradigm ("base") and another ("output"), and generates a set of morpho-phonological rules which can be used to derive the output from the base.

The model starts by generating word specific rules for each pair of input-output, which are subsequently generalized by comparing rules producing the same change, and constructing other, more general rules, referring not to particular lexical items, but to environments in which certain changes occur. Each of the rules has a numerical characteristic, *reliability*, associated with it. Reliability is calculated by dividing the number of the forms included in the rule's structural change by the number of forms included in the rule's structural description. Reliability ratios are adjusted using lower confidence limit statistics to yield *confidence values* (see Albright 2005 for details). When new forms are derived, the rule with the highest reliability is applied, and the output obtained by using this rule becomes an actual output.

I compiled a list of the 531 most frequent regular Russian verbs, together with all six present tense forms, the infinitive, and the imperative from the online frequency dictionary by Sharoff. Verbs used have a corpus frequency of more than 50 instances per million.

Several irregular verbs such as *jést*^{j} 'to eat', *dát*^{j} 'to give', *jéxat*^{j} 'to ride, to go', and all prefixed forms of them were excluded. Verbs with the prefix *vý-* were also excluded: *vý-* is the only stressed verbal prefix in Russian, and it affects the formation of imperatives. Since *vý-* is not in a local environment with the imperative ending, the model will not be able to take its presence into account.

The Java-based version of the MGL, available at (http://www.linguistics.ucla.edu/people/hayes/learning/), was executed on the data with the task of learning how to derive the imperative form based on the non-past tense forms and the infinitive. Since 2SG, 3SG, 1PL, and 2PL differ only in the inflectional affix, I considered the derivation of the imperative from 1SG, 2SG, 3PL, and the infinitive.

There were two reasons to include the infinitive form: 1) Acquisition errors often include imperative forms derived from the INF-PAST form as in *pʲisáj* (children's form) from *pʲisátʲ* (Inf) instead of expected *pʲišší* 'write', and 2) to numerically evaluate whether the infinitive serve as a second base needed to derive the imperative.

The sample data for the Learner is represented in (4) below:

(4)	***Inf***	***1SG***	***2SG***	***3PL***	***Imp***	***gloss***
a.	móč	mogú	móžeš	mógut	mogʲí	'can'
b.	skazátʲ	skažú	skážeš	skážut	skaží	'tell'
c.	čitátʲ	čitáju	čitáješ	čitájut	čitáj	'read'
d.	smotrʲétʲ	smotrʲú	smótrʲiš	smótrʲat	smotrʲí	'watch'
e.	žalʲétʲ	žalʲéju	žalʲéješ	žalʲéjut	žalʲéj	'have pity'
f.	rʲisovátʲ	rʲisúju	rʲisúješ	rʲisújut	rʲisúj	'draw'
g.	lʲubʲítʲ	lʲublʲú	lʲúbʲiš	lʲúbʲat	lʲubʲí	'love'
h.	otvʲétʲitʲ	otvʲéču	otvʲétʲiš	otvʲétʲat	otvʲétʲ	'answer'

There is a great deal of variability in consonant alternations and stress among different forms of the verb. The verb in (4a) has a stem-final consonant which alternates between four different possibilities: [č, g, ž, gʲ]. Similar, consonant alternations can be observed in (4b), (4g), and (4h). Further, the difference between (4b) and (4c) illustrates the problems which arise while trying to derive the imperative from the infinitive. Both of these verbs have a C-final root; however, the verb in (4c) has the suffix *-aj-* in the NON-PAST stem, and the verb in (4b) does not. Such a difference can only be captured by considering one of the non-past forms of a verb, since the suffix *-aj-* only surfaces there. A similar difference distinguishes (4d) and (4e), with the verb in (4e) having the suffix *-ej-* in the non-past forms of the verb. Example (4f) presents a suffix alternation between *-ova-* in the INF-PAST forms of the verb and *-uj-* in the NON-PAST forms. Also, the stress patterns are different for the verbs. The verbs in (4a), (4b), (4d), and (4g) have mobile stress, while the verbs in other examples have fixed.

Further, the data about the phonological features of Russian phonemes were provided to the learner[2] and list of phonotactically illegal sequences for imperative formation was fed into the model.[3]

There are two predictions: 1) Given the structure of the Russian imperative and its relation to the NON-PAST paradigm, the MGL should have a much lower success rate for the infinitive than for NON-PAST forms; 2) Given that the necessary information about stress and CC-sequences is found in the 1SG form, this form should be more successful in generating the correct imperative than other NON-PAST forms.

The algorithm generated a set of rules which can account for imperative formation from 1SG, 2SG, 3PL, and the infinitive forms. Wug-tests were also conducted in order to evaluate the correctness of the generated rules, and following the learning state, imperative forms were derived for all verbs used as the input to the MGL. A substantial number of the inputs yielded multiple possible forms for the imperative (with different confidence values).

(5) a. 3PL → Imperative; input *pr^{j}edlóžat* 'offer-3PL'
(i) **pr^{j}edlóž* (ii) **pr^{j}edlóži* (iii) *pr^{j}edloží*
b. 2SG → Imperative; input *uxód^{j}iš* 'leave-2SG'
(i) **uxód^{j}* (ii) **uxód^{j}i* (iii) *uxod^{j}í*
c. 1SG → Imperative; input *proisxožú* 'happen-1SG'
(i) *proisxod^{j}í* (ii) **proisxožÍ*
d. Infinitive → Imperative; input *nabl^{j}udat^{j}* 'observe-inf'
(i) **nabl^{j}ud^{j}í* (ii) *nabl^{j}udaj*

The examples in (5) provide several representative situations. The cases in (5a), (5b), and (5c) are representative of two common learning mistakes that arise if the base is chosen to be a NON-PAST forms of the verb. The choice of the imperative suffix depends on the stress in the

[2] The features included ±consonantal, sonority (rated on the scale from 0 to 5, with 5 corresponding to vowels, and 0 to stops), ±voice, ±LAB, ±COR, ±DORS, and ±palatalized (for consonants), height, and backness (for vowels).

[3] The list included violations of the sonority sequence, (*bl*, *bn*, *br*, *bj*, etc.); sequences of a non-palatalized consonant followed by *i* (*bi*, *ni*, *ji*, etc.); sequences of palatalized labial or velar obstruents followed by *u* (*b^{j}u*, *k^{j}u*, *f^{j}u*, etc.); and word-final *k^{j}#*, *g^{j}#* and *x^{j}#*. Even though this list is far from being exhaustive, these sequences constitute the phonotactically illegal sequences to be avoided in imperatives.

1SG form of the verb; such information is not readily available in the 2SG and 3PL forms, and therefore the large number of 2SG and 3PL inputs yield at least two outputs, one with *-i* and one without, such as cases in (5a,b-i) and (5a,b-ii) (the 1SG form of 5a and 5b is stressed on the ending, and not on the root, like the 3PL and 2SG forms). Further, the Learner failed to predict whether stress in the imperative falls on the ending like in (5a,b-iii) or on the stem, as in (5a,b-ii), as both cases are possible in Russian. A problem of stem "mutations" can be seen in the derivation of the imperative from the 1SG form, (5c). The final stem consonant undergoes mutation $d^j \rightarrow ž$ in the 1SG form, and therefore the Learner generates two possibilities: one preserving the mutated consonant *ž*, and one having the non-mutated consonant d^j. A common problem identified by the Learner in the derivation of the imperative from the infinitive is demonstrated in (5d). The infinitive gives no indication of which suffixes occur in the NON-PAST stem, as suffixes *-aj-* and *-a-* both correspond to infinitives ending in *-at'*; only the finite forms of the verb provide information about them. Therefore in trying to derive the imperative from the infinitive, the Learner does not have valid information regarding the presence of such suffixes, and thus the two possible outputs cover both alternatives.

Other, less common problems will not be considered here.

4. Quantificational analysis

A quantificational analysis should be able to show whether any of the verbal forms considered above fares better than any other form for generating imperatives, based on a number of parameters. The three parameters considered in this analysis are the following.

1) The percentage of imperatives derived correctly. Here I assume that the surface form of the imperative generated by the Learner for each input is an output which is derived by the rule which has maximal confidence. The percentage of imperatives derived correctly was computed using wug-tests conducted after the learning stage.

2) The average number of outputs. As mentioned earlier, the Learner generates several outputs for each input. This characteristic provides the average number of outputs generated from each input. The smaller this number is, the less variability is observed in the outputs, and therefore the computational load required to derive the output is less. Thus, the

input form which has a lower number of outputs must be superior to the input form with a higher average number of outputs.

3) The average confidence of correct outputs. This characteristic is equal to the average confidence value of the rules which derive the outputs (imperative forms) occurring in the language. If the Learner failed to correctly generate an output at all, the confidence was considered to be equal to 0.

The results are given in the table below and in Figures 1-3.

	Percentage of imperatives derived correctly	*Average number of outputs*	*Average confidence value of the correct output*
1sg	95%	1.52	0.87
2sg	93%	1.60	0.85
3pl	93%	1.51	0.84
Inf.	87%	1.80	0.78

Figure 1 shows the percentage of imperative forms generated/predicted correctly out of a total of 531 verbs used in the test. Even though the absolute differences between the input forms are not large, the 2SG, the 3PL forms (each at about 93%) and the 1SG form (95%) fare better in comparison to the infinitive (87%). Logistic regression showed the effect of the choice of the input form on the percentage of the imperatives derived correctly (*Wald statistic* = 18.414; $p < 0.001$). Further, logistic regression post-hoc tests showed that the percentages of correct predictions from finite inputs (1SG, 2SG, 3PL) do not differ significantly from each other, while the number of correct

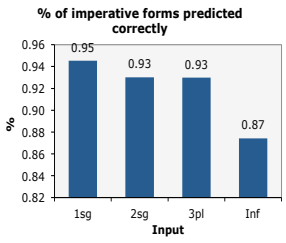


Figure 1. Percentage of imperative forms predicted correctly.

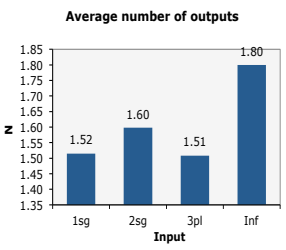


Figure 2. Average number of outputs

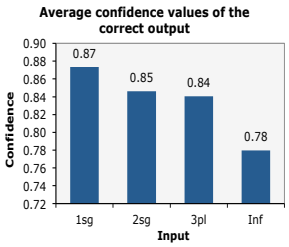


Figure 3. The average confidence of the correct output.

outputs from the infinitive is significantly different from the numbers obtained by using finite forms as inputs. Therefore, based on this parameter, no differences exist among non-past forms of the verb. The infinitive unsurprisingly shows significantly lower performance in deriving the imperative.

Figure 2 shows the average number of possible outputs generated from one input. The average number of outputs ranges from 1.51 when the 3SG form is used as a base for the imperative up to 1.80 when the infinitive form is used. Logistic regression results show that the effect of the input is significant. Post-hoc tests reveal that the differences in this parameter between infinitive and finite forms are significant, and the differences between the various finite forms are not.

With respect to the average confidence values for occurring output forms (Figure 3), the logistic regression Wald test results show that the effect of the input form is significant. Further, post-hoc tests reveal that the infinitive, with the average confidence of the correct output equal to 0.78, fares significantly worse than the other forms (which have average confidence values of correct forms ranging from 0.84 to 0.87), while there is no significant difference among the finite forms. It indicates that the occurring forms receive lower confidence values when derived from the infinitive as opposed to the finite forms.

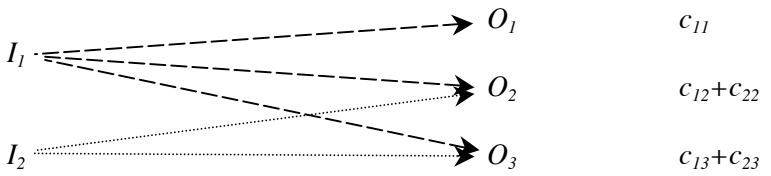
Were the statistical data computed based on the results of MGL surprising? Both yes and no. As predicted, infinitives do not fare as well in predicting imperatives in comparison to finite non-past forms of the verbs. Surprising however are two facts. First is the fact that even though the infinitive lags behind the finite forms, it is still fairly powerful in predicting the correct imperative form, performing successfully in 87% of the cases. Second, surprising is the fact that no statistically significant differences were found between 1SG, 2SG, and 3PL forms except with respect to the average confidence of the correct output parameter, and that one form, in spite of some traditional claims that two are needed, fares relatively well in predicting imperatives correctly for 93-95% of the cases.

5. Towards a multiple bases approach

The next question I will try to answer is what happens if the learner has access to several bases while trying to generate the imperative?

I will assume that the learner has already acquired the non-past tense forms of the verbs along with the infinitive. What that means is that the learner is able to apply the rules for the formation of the imperative not just from one base form, but from several base forms belonging to the same verbal paradigm. The confidences for obtaining the output form from different bases are then added together, and the sum of these confidences provides the final score for the given output form[4]. The winning output is the form for which the sum of the confidence values is maximal. Such a strategy is demonstrated schematically in (6), where *I* stands for input, *O* stands for output, and c_{ij} is a confidence value of obtaining O_j from I_i.

(6)



In order to check whether this strategy provides an increase in the percentage of correctly derived outputs, the sums of the confidence values were calculated for each pair of input bases (1SG+2SG, 1SG+3PL, etc.) and for all four input bases (1SG+2SG+3PL+Inf), and the output with the largest sum of confidence values was considered to be the winner for given combination of bases. The percentage of imperatives predicted correctly was calculated after that. The data is given in Figure 4. As one can see, when using two bases simultaneously, the correct output is obtained in 94-97% of the cases, while deriving the imperative form from all four bases leads to the correct result for 98% of verbs.

As previously, logistic regression was conducted in order to investigate the effect of the input form on the percentage of correct imperative derivations. The effect of number of bases used was found to be significant (*Wald statistic* = 56.325; *df* = 2; $p < 0.001$). Further, post-hoc tests were conducted in order to investigate whether multi-base model provides a significant improvement over the one-base model. Starting from the 1SG form, significant improvement was achieved by

[4] One can also use the average of the confidence values, but since the average is proportional to the sum, the results will be the same.

combining it with the 2SG form or with the infinitive. Further, in comparison to the 2SG form alone, only combining it with the 1SG form or with the infinitive produced significantly better results. Also, combining the 3PL with the 1SG or with the infinitive improved the imperative formation in comparison to considering the imperative derivation from the 3PL form alone. The infinitive benefitted from being combined with any of the finite forms. Lastly, there was no significant difference between considering two bases and four bases (*Wald statistic* = 3.561, *df* = 1; *p* = 0.059).

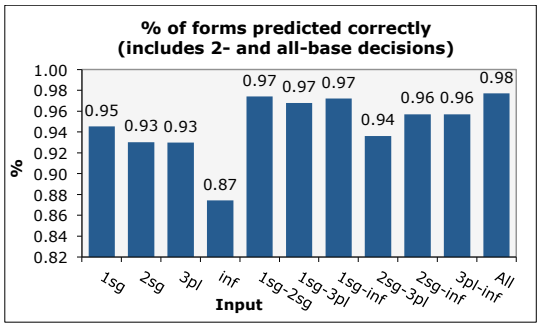


Figure 4. Percentage of imperatives predicted correctly.

The conclusion of this exercise is the following. Being able to use the rules deriving imperatives from two basic forms simultaneously and to combine the results afterwards produces significant improvement in the accuracy of imperative formation. Also, there is no need to resort to more than two forms in order to generate the imperative form. Finally, all pairs (except for 2SG + 3PL) fare equally well in the task of generating the imperative, and they produce results which do not differ from results obtained by combining all four forms.

6. Discussion

6.1. Multiple bases

It is expected that considering multiple bases of the verb would produce better results and increase the success rate of deriving any paradigmatic form, and quantificational analysis showed that this is indeed the case. The percentage of correctly predicted imperatives increases significantly when using two bases compared to one base. While the absolute increase in success rate ranged from 2% to 4%, entire classes of verbs were able to benefit from it.

The use of the 1SG form together with any other form resolved the problem with verbs exhibiting mutation only in the 1SG. Considering the

2SG or the 3PL form along with the 1SG form helped to alleviate problems with the class of verbs involving mobile stress. The infinitive clearly benefitted the most as information about the verbal class became available which allowed for the correct choice of the NON-PAST stem, providing access to the information about which suffix (*-a-* or *-aj-*, *-ova-* or *-uj-*) must surface in the imperative.

Lastly, since most problems are resolved by considering two forms, we would not expect the combination of all four forms to produce significantly better results than the combinations of two forms. This is exactly what we found: the combination of the four forms fares as well in deriving the imperative as the following combinations: 1SG+2SG, 1SG+3PL, 1SG+inf, 2SG+inf, 3SG+inf.

6.2. Acquisition errors in imperative formation

One of the important questions to answer is whether the MGL can account for acquisition errors in the formation of the imperatives, and can predict where errors are likely to be observed. The systematic data on children's acquisition of imperatives in Russian does not exist. However, a Google search revealed several weblogs/forums describing children's errors in imperative formation. Examination of the forms surfacing in children's speech provides evidence that here the infinitive is used as a base for imperative formation (7).

(7)	***Infinitive***	***Imperative***	***Children's form***	***Gloss***
a.	pʲisátʲ	pʲiší	pʲisáj	'write'
b.	prʲátatʲ(sʲa)	prʲáč(sʲa)	prʲátaj(sʲa)	'hide-REFL'
c.	rʲisovátʲ	rʲisúj	rʲisováj	'draw'
d.	tancevátʲ	tancúj	tancováj	'dance'

No errors of other types were recorded, though given the paucity of the acquisition data, no real conclusions can be drawn. What is interesting, however, is that all of the errors in (7) are the types of errors also produced by the Learner, and they all involve the *-at^{j}* infinitive type. The infinitive in Russian does not show a distinction between *-a-* and *-aj-* verbs, and the Learner generates the imperatives of *-a-* verbs as if they belong to the *-aj-* class. Also, the Learner tends not to replace the *-ova-* suffix with *-uj-*. Examples (7a-b) present the cases when children postulate the existence of the suffix *-aj-* for verbs of *-a-* class, and there

are similar errors generated by the Learner. Examples (7c-d) show that children also fail to replace *-ova-* with *-uj-*: these forms are also similar to the imperative forms generated by the Learner.

These preliminary results show that children in fact use the infinitive as a base for the imperative, and the Learner models a stage in language acquisition correctly. The possibility of deriving the imperative from the infinitive correctly for 87% of the verbs makes learners think that the infinitive is a possible base, since it can account for a majority of the imperatives they hear.

6.3. Frequency and markedness of the base

The fact that Russian learners use the infinitive to derive the imperative of some verbs must be justified based on frequencies of the forms of the verbal paradigm in the input children hear. Vakar (1966) presents data about the frequencies of various verbal forms in spoken Russian. However, frequencies of particular persons and numbers are given separately, and not combined. I approximate the frequencies of the individual members of the verbal paradigm by multiplying the frequency of the particular person by the frequency of particular gender by the frequency of the mode. For instance, to find the frequency of the 1SG non-past form, I use the approximation *freq*(1SG non-past) = *freq*(1st) × *freq*(sg) × *freq*(indicative). The results are given in (8):

(8)

	singular	***plural***	***infinitive***
1st	14.30%	4.74%	15.30%
2nd	22.44%	7.44%	
3rd	15.08%	5.00%	

Because of comparatively high frequency of the infinitive form in the spoken Russian, it is plausible that learners use it as a base[5]. This leads to the question of whether the most frequent form of the paradigm

[5] It is unclear why given the high frequency of the 2SG form it is not used by children for imperative formation. One possibility is that children prefer to use non-finite forms (infinitives) in order to generate other non-finite forms, such as imperatives. I leave this question for further research.

always serves as its base (Bybee 2001). Our findings show that high frequency of a verbal form does not necessarily mean than it serves as a better base. While being the most frequent, the 2SG does not fare any better than other, less frequent forms such as the 3PL and 1SG. And the 3PL, being much less frequent than the 1SG form, is not significantly worse in predicting the imperative.

Similarly, the least marked forms are also not necessarily better bases (cf. Benua 1997). If relative markedness of the forms is evaluated in terms of mutations, we would expect the 3pl form to be the least marked, since there are no mutation patterns where the 3pl form has a mutated consonant while other forms do not. Further, if the mutation occurs in the 3pl form, it must occur in all other non-past forms. The 3pl form, however, is not the best base in predicting the imperative, as its success rate in deriving the imperative is statistically the same as the success rate of any other non-past form.

7. Conclusion

In this paper I tried to achieve two main goals. First, I evaluated the status of different verbal forms in Russian as potential predictors for the imperative. Based on my study, the debate in the literature about which form serves as the base for the imperative is justified, as all finite forms perform equally well at this task. Also, my study showed that the infinitive provides the worst results, though it turned out to be more successful than anticipated.

Albright's claim that there is no need to resort to the underlying representation and that a free-standing form can be used as a base for the paradigm was confirmed. While none of the forms of the verbal paradigm achieves 100% accuracy in predicting the imperative, one form is enough to achieve up to 95% success rate. Such a result is initially unpredictable, given the complexity of the Russian verbal paradigm. Further, since all the non-past forms were found to be statistically the same in terms of predicting the imperative, the task of deriving the imperative is simplified for language learners as there is no need to select a single base; this selection task is one of the most computationally complex parts of Albright's model of base discovery, and in the case of Russian imperative derivation it can be avoided; also, no matter which

finite form dominates the input to the learner, the imperatives will be acquired with the same success rate.

Equal similarity of the imperative to all non-past forms allows it to be treated as a certain point of equilibrium: connections of it to any other non-past forms are on average statistically the same. For learners acquiring Russian, at the initial stage the imperative exhibits the strongest connection with the infinitive. Later, it shifts away from the infinitive and its connection with the infinitive weakens, while the connection with the non-past forms strengthens until all the connections with non-past forms become equal.

Second, I proposed an extension to the MGL model to deal with multiple bases and tested it on Russian imperatives. I demonstrated that this extension provides significant improvement over the single-base model. I further examined the errors in imperative formation in the speech of Russian learners, and demonstrated that the MGL predictions about the problems in imperative formation are borne out. Based on the infinitive's frequency its fairly high success rate, it is unsurprising that it is often chosen as a base for imperative by language learners.

This extension to the MGL is relevant for the Optimal Paradigms (OP) approach (McCarthy 2005). I demonstrated that using two bases produces statistically significant increase in the number of imperatives derived correctly as compared to using only one base. The consequence of it is that having access to the entire verbal paradigm, as in the OP model, is preferential to using just one base, as in the original Albright model.

While in this paper I demonstrated that it is possible to derive the imperative from a free-standing form of the verb, it was not shown whether this approach is superior to the Jakobsonian approach of deriving the imperative from the abstract stem. I leave this question for future research.

References

Albright, Adam, 2002. The Identification of Bases in Morphological Paradigms. Doctoral dissertation, UCLA, Los Angeles, CA.

Albright, Adam. 2005. The morphological basis of paradigm leveling. In *Paradigms in Phonological Theory*, ed. by Laura Downing et al., 17-43. Oxford: Oxford University Press.

Albright, Adam and Bruce Hayes. 1999. An Automated Learner for Phonology and Morphology. ms. Available at http://www.linguistics.ucla.edu/people/hayes/learning/

Benua, Laura. 1997. Transderivational Identity: Phonological Relations Between Words. Doctoral dissertation, UMass Amherst.

Bybee, Joan. 2001. *Phonology and language use.* Cambridge: Cambridge University Press.

Jakobson, Roman. 1948. Russian Conjugation. *Word* 4:155-67.

McCarthy, John. 2005. Optimal Paradigms. In *Paradigms in Phonological Theory*, ed. by Laura Downing, Tracy Alan Hall, Renate Raffelsiefen, 170-210. Oxford: Oxford University Press.

Sharoff, Serge. The frequency dictionary for Russian. http://www.comp.leeds.ac.uk/ssharoff/frqlist/frqlist-en.html

Švedova, N. Yu. 1982. *Russkaja Grammatika* [Russian Grammar]. Moscow: Nauka.

Timberlake, Alan. 2004. *A Reference Grammar of Russian.* Cambridge: Cambridge University Press.

Townsend, Charles E. 1980. *Russian Word-Formation: corrected reprint.* Columbus: Slavica Publishers.

Vakar, N. P. 1966. *A Word Count of Spoken Russian: The Soviet Usage.* Columbus: The Ohio State University Press.

Vinogradov, V. V. 1972. *Russkij Jazyk.* [Russian Language]. Moscow: Vysshaja shkola.

Zaliznjak, A. A. 1977. *Grammaticheskij slovar' russkogo jazyka*. [Grammatical Dictionary of Russian]. Moscow: Russkij Jazyk.