Features of Common Slavic Ablaut Alternation

Ondřej Šefčík

1. Introduction

The issue of the my paper is, first, that the system of Old Church Slavonic ablaut could be described with the use of the set of ablaut values and features derived from them, and second, that such ablaut features could be taken for vectors; third, that such features form a vector and a metric space in which all differences between ablaut grades could be exactly stated.

The analytic approach to the features which is here presented is derived from that of Marcus (MARCUS 1967). There are some similarities to the methods of Dependency Phonology (see ANDERSON & EWEN 1987 or DURAND 1996) and closely to the method of phonemic description by Hubey (HUBEY 1999), who reworked a rather intuitive approach of DP using the algebraic form. The basic idea of Hubey's revisited DP-method is to found and to define basic orthogonal vectors and to describe the whole given system as a vector space. Third, our attention is more focused on metric spaces than vector spaces, which is the shift of focus, not a different approach at all.

The solutions of the method proposed here could be applied not only on the OCS ablaut system, but is relevant for the reconstructed Common Slavic ablaut system, too. The data from OCS are used purely as raw material for the presentation of the method, although the method itself is not determined by the data.

2. Ablaut grades

The OCS ablaut is such an alternation between morphs of one identified morpheme (most often a root-morpheme) which indicates grammatical information. Or in other words, the ablaut is the grammatical derivation of the morpheme. The OCS ablaut is then a non-linear alternation, i.e. such an alternation not triggered by the phonemic context of a given morph which is different from the linear alternation such as palatalization of velars neighboring palatal vowels and diphthongs.

(1): Ablaut grades of different verbal roots[1]:

√r_k: *rekǫ, rokъ, -rěkati, rьci, -ricati*

√t_k: *tečetь, tokъ, -těkati, -tačati, tьci*

√g_n: *ženǫ, goniti, -gańati, gъnati*

√sl_v: *slovo, slaviti,* (*slъšati*[2]), *slyšati*

√b_r: *berǫ, -borъ, bьrati, -birati*

√v_d: *vedǫ, voditi, věsъ*

√m_r: *mrěti, moriti, -mariti, -mьrǫ, -mirati*

√d_x: *duxъ, dъxnǫti, dyxati*

Each of the ablaut variants of the given morpheme is termed the **ablaut grade**. All ablaut grades form **the set of ablaut grades**.

In this phase of analysis the set of ablaut grades is an unordered set, i.e. there is no additional information on relations between different ablaut grades. Such an unordered set is not a system, because any system requires an organization, or better, the structure over it.

We will try to develop such a structure using sets of values of all ablaut grades to postulate ablaut features.

An ablaut value is then such a property which could be related to an ablaut grade and which makes the difference between grades in at least one example. For all ablaut grades we need to define at least a minimal set of such values.
This is equivalent to the relation between phonemes of any phonemic system and the sets of values of its phonemes, for both systems consist of a set of elements (phonemes, ablaut grades) and the organization itself (system of phonemic oppositions, system of morphonemic oppositions). Other similarities will be clear from the following lines.

3. Values of ablaut

Each ablaut grade has then its own set of such values, or in other words, any given ablaut grade will be attached to the one and only one set of values (or better – there is a bijective[3] relation between each ablaut grade and its set of values).

---

1 All prefixes are separated. All grades are not ordered due proportionality.

2 Probably such a form is only a variant of *slyšati*, but definitely of a different ablaut value.

Any value of a given ablaut grade is here symbolized $V\{x\}$ in general or in the case of a concrete value simply as {value}.
The set of values of a given ablaut grade is symbolized as $V\{x\}$ (i.e. $V\{x\} = v_1\{x\} \cup v_2\{x\} ... \cup v_i\{x\}$).

The set of values of the whole ablaut system is symbolized simply as V, which is the set union of particular values of concrete ablaut grades, hence $V = V\{a\} \cup V\{b\} ... \cup V\{z\}i$).

Any value could be classified with the help of three criteria, every one expressing some properties of the given value in a language system. Each criterion consists of a pair of possible incompatible properties.
The first criterion is **homogeneity**. Two values are **homogeneous** if they are members of the same subset $V_{i;}$ otherwise such values are **heterogeneous** (MARCUS 1967: 46-47, KORTLANDT 1972: 57; cf. ŠEFČÍK 2008: 6, ŠEFČÍK 2009: 186-187).

(2a) The values {*e*-grade} and {*o*-grade} belong to the same subset of values, hence they are homogeneous.

(2b) The values {reduced-grade} and {*o*-grade} do not belong to the same subset of values, hence they are heterogeneous.

The second criterion is **compatibility**. Two values are **compatible** if they are attached to the same ablaut grade {*x*} (members of the same set of ablaut values $V\{x\}$). If such values are not members of the same set of ablaut values (they are not attached to the one ablaut grade), they are **incompatible** values. It should be mentioned that all compatible values are heterogeneous, but not the other way round; thus, some heterogeneous values are not compatible (MARCUS 1967: 47-48, KORTLANDT 1972: 57 cf. ŠEFČÍK 2008: 6, ŠEFČÍK 2009: 187).

(3a) The values {*e*-grade} and {reduced-grade} are attached to the same ablaut grade; hence they are compatible.

(3b) The values {*e*-grade} and {*o*-grade} are not attached to the same ablaut grade; hence they are not compatible.

The third criterion is **contrastivity**. Any values v_i and v_j are **contrastive** if there are two ablaut grades such that $V\{x\} - V\{y\} = v_i$ and $V\{y\} - V\{x\} = v_j$. In other words, if replacement of one value of a given ablaut grade by another value results in another ablaut grade, the values are contrastive. Otherwise both values

---

3 Bijection is such a relation between sets *A* and *B* so that if every element of *A* is related to exactly one element in *B* and if then every element in *B* is related to only exactly one element of *A*.

are **incontrastive**. All contrastive values are homogeneous and therefore incompatible, but not the other way round (MARCUS 1967: 48-49, KORTLANDT 1972: 58 cf. ŠEFČÍK 2008: 6-7, ŠEFČÍK 2009: 187). Curiously, the ablaut system of OCS is in fact structured in a way, so that any pair of homogeneous values is necessary contrastive.

(4a) The values {*e*-grade} and {*o*-grade} are contrastive, because the absence of the value {*e*-grade} necessarily means that the grade is {*o*-grade}.

(4b) The values {*e*-grade} and {reduced-grade} are not contrastive, because the absence of one value does not necessarily mean that the second value is present.

Every value which is contrastive, homogeneous and incompatible we will term the **pertinent value** (MARCUS, 1967: 46-47).

For OCS (and CSl, too) we deal with the following set of ablaut pertinent values: {*e*-grade}, {*o*-grade}, {reduced-grade}, {non-reduced-grade}, {lengthened-grade}, {non-lengthened-grade}.
The above mentioned ablaut grades of different roots could hence be arranged in the following table[4]:

4 All prefixes are removed.

(5)	***e*G**	***o*G**	***ē*G**	***ō*G**	${}_{e}$**G**	${}_{o}$**G**	${}_{\bar{e}}$**G**	${}_{\bar{o}}$**G**
√g_b		*gubiti*				**gъnǫti* *gъnuti*[5]		*gybati*
√g_r	*žeravъ*	*gorěti*	**žarъ*[6]	*garati*		*gъrno*[7]		
√vr_t	*vrěmę*	*vratiti*			*vrъtitъ*			
√b_d	*bljusti*	*buditi*				*bъděti*		
√r_k	*rekǫ*	*rokъ*	*rěkati*	*račiti*[8]	*rьci*		*ricati*	
√t_k	*tečetь*	*tokъ*	*těkati*	*tačati*	*tьci*			
√g_n	*ženǫ*	*goniti*		*gańati*		*gъnati*		
√sl_v(s)		*sluxъ, slovo*		*slaviti*		*slъšati*[9]		*slyšati*
√b_r	*berǫ*	*borъ*			*bьrati*		*birati*	
√v_d	*vedǫ*	*voditi*	*věsъ*					
√m_r	*mrěti*	*moriti*		*mariti*	*mьrǫ*		*mirati*	
√d_x		*duxъ*				*dъxnǫti*		*dyxati*

Note: Marking of ablaut grades:

{reduced non-lengthened *e*-grade} (represented by ь, marked ${}_{e}$G);

{reduced non-lengthened *o*-grade} (represented commonly by ъ, marked${}_{o}$G);

{non-reduced non-lengthened *e*-grade} (represented commonly by *e*, marked *e*G);

{non-reduced non-lengthened *o*-grade} (represented commonly by *o*, marked *o*G);

---

5 Only in Russian Church Slavonic, besides *gъbnuti*, see DERKSEN (2008: 197)

6 The Proto-Slavic form is reconstructed on Russian *žar,* Czech *žár*, Slovak *žiar*, P. *żar*, Serbo-Croatian, Slovene *žâr*, Bulgarian *žar*, cf. MACHEK (1971: 722), VASMER (1953: I: 410), DERKSEN (2008: 554).

7 Only in Russian Church Slavonic, see DERKSEN (2008: 199).

8 Cf. LIV (457–8).

9 Although *slъšati* we consider as a variant of *slyšati*, formally it is {reduced non-lengthened *o*-grade} = ${}_{o}$G**.**

{non-reduced lengthened *e*-grade} (represented commonly by *ě*, marked $\bar{e}$G);

{non-reduced lengthened *o*-grade} (represented commonly by *a*, marked $\bar{o}$G);

{reduced lengthened *e*-grade} (represented by *i*, marked ${}_{\bar{e}}$G);

{reduced lengthened *o*-grade} (represented by *y*, marked ${}_{\bar{o}}$G).

Two mutually contrastive pertinent values form an **ablaut feature**. Hence, any ablaut feature could have two values, one “negative”, the second marked “positive”. The values are attached arbitrarily (ŠEFČÍK 2008: 7, ŠEFČÍK 2009: 187).

Between both pertinent values of each given feature there is then a binary opposition. For simplicity we will mark each “positive” pertinent value in the features as 1, each “negative” pertinent value as 0, and hence each ablaut grade could be expressed in a binary code.

When organizing the above enumerated pertinent values of OCS, we face three ablaut features (AFS):

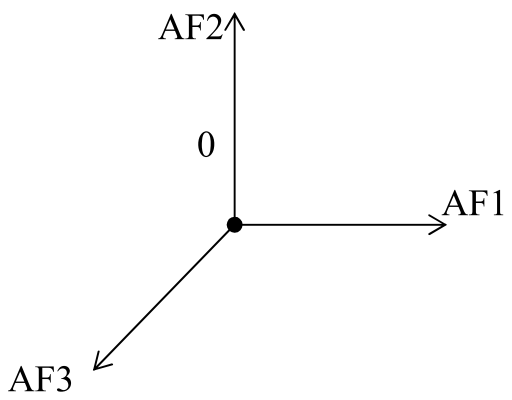
AF1: {non-reduced-grade} &{reduced-grade}

AF2: {non-lengthened-grade}& {lengthened-grade}

AF3: {*e*-grade} & {*o*-grade}

As we can see, all three AFs are mutually independent, i.e. no one is dependent on the other AF. Such relations between AFs can be expressed in the following graph.

(6) Systems of ablaut features of OCS:



In such a system there are basic vectors attached to basic ablaut grades: *e*G is in the zero position, ${}_{e}$G is the basic value on AF1 vector, $\bar{e}$G is basic on AF2 vector, *o*G is basic on AF3 vector.

4. Vector space

All three features then could be interpreted as vectors and the ablaut system as a vector space. Because AFs are mutually independent, and their dot product is zero, they form an **orthogonal** vector space.

For simplicity we will mark the AF1 vector as *R*, the AF2 vector as *L* and the AF3 vector as *O*.

Any concrete ablaut grade *V{x}* has hence its unique set of values v_1-v_2-v_3 of the features of the same ordering (*R–L–O*). The commas will be omitted in the following lines for simplicity. Values (and also features) for each ablaut grade are then ordered and both sets of ablaut grades and the sets of values could then be considered as an ordered sets due to the equivalence between both sets.

(7)

			AF vectors		
			R	L	O
Ablaut grades	{non-reduced non-lengthened *e*-grade}	= *e*G	0	0	0
	{reduced lengthened *e*-grade}	= $\bar{e}$G	0	1	0
	{reduced lengthened *e*-grade}	= ${}_{\bar{e}}$G	1	1	0
	{non-reduced non-lengthened *e*-grade}	= ${}_{e}$G	1	0	0
	{reduced non-lengthened *o*-grade}	= ${}_{o}$G	1	0	1
	{reduced lengthened *o*-grade}	= ${}_{\bar{o}}$G	1	1	1
	{non-reduced lengthened *o*-grade}	= $\bar{o}$G	0	1	1
	{non-reduced non-lengthened *o*-grade}	= *o*G	0	0	1

The AF vector space is also considered as **unit vector**, with the size of vector equal to 1, for the distance between two values of one feature is equal to 1 (more on distances see below).

The unit vectors, which are orthogonal, too, are termed **orthonormal vectors**. Hence our vector space of ablaut grades is an orthonormal vector space.

Any ablaut grade (i.e. *V{x}*) could be then described as a linear combination of orthonormal unit vectors which could be written as (v_x = value of a given ablaut feature, i.e. 0 for the unmarked value, 1 for the marked value):

$V\{x\} = v_1R + v_2L + v_3O$

As a consequences of arbitrary ordering, symbols for vector could be omitted and any set of values of a given ablaut grade V{x} could be expressed in a three digit code, as above in the table 8.

Any ablaut grade *V{x}* could then be taken for a vertex of a unit cube. The set of such vertices is then marked B^n, and that set is a set of ordered *n*-tuples, containing numbers 0 and 1, or in other words, as $B^n = \{0, 1\}^n$.

In that sense the elements of the set B^n form **dyadic** (or **Boolean**) **vectors** and the structure of the ablaut system of OCS is a **Boolean algebra**, consisting two elements 0 and 1 (i.e. binary two-valued Boolean algebra) (see KURATOWSKI 1977: 34–35).

The distance between any pair of ablaut grades is then given by the sum of differences between codes of given ablaut grades:

(8a) Example of minimal distance:		
V{non-reduced non-lengthened *e*-grade}	000	
V{non-reduced non-lengthened *o*-grade}	001	
V{non-reduced non-lengthened *e*-grade}– *V*{non-reduced non-lengthened *o*-grade}	001	Σ =1

(8b) Example of zero distance:		
V{non-reduced non-lengthened *e*-grade}	000	
V{non-reduced non-lengthened *e*-grade}	000	
V{non-reduced non-lengthened *e*-grade}– *V*{non-reduced non-lengthened *e*-grade}	000	Σ = 0

(8c) Example of more than minimal distance:		
V{non-reduced non-lengthened *e*-grade}	000	
V{reduced non-lengthened *o*-grade}	101	
V{non-reduced non-lengthened *e*-grade}– *V*{reduced non-lengthened *o*-grade}	101	Σ = 2

Such distance between codes is known as **Hamming distance**[10]. Hence, any difference between set of values, considered as codes, could be expressed as a Hamming distance.

The ablaut vector space which we now face could be considered as a **metric space**.

And indeed, we can consider the difference between two values of a same feature of two ablaut grades as a minimal distance sui generis in such a vector space and the identity between two values of the same feature of two ablaut grades as zero distance; then we can metricize such a vector space easily.

5. Metric space

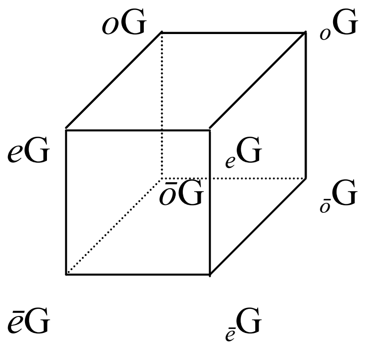
Any metric space is generally a tuple (A, ρ) where A is a set of elements and ρ is the distance between them given by the mapping of the set A on itself (i.e as a Cartesian product $A \times A$).
The metric space complies with the properties:

1) if two elements have null distance between each other, then we are dealing with one element (axiom of **identity**, i.e. $\rho(x, y) = 0,$ if $x = y$);
2) the distance between an element x and an element y is equal to the distance between y and x (axiom of **symmetry**, i.e. $\rho(x, y) = \rho(y, x)$);
3) the distance between x and z is equal or smaller than the sum of the distances between elements x and y and y and z (axiom of the **triangle inequality**, i. e. $\rho(x, y) + \rho(y, z) \geq \rho(x, z)$). See (MARCUS 1967: 34-35, KURATOWSKI 1977: 115).

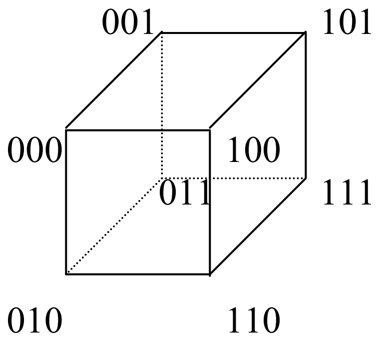
For a more obvious image of the relations inside the system of the OCS ablaut we can draw the graph, expressing the system of OCS, in the form of 3-cube (dyadic cube), with the length of each edge equal to 1. All grades are expressed by their sets of values, as described in the table 7. In the following examples the metric space of OCS ablaut is presented both with codes due to the Hamming distance, both with enumerated ablaut grades, in the form of two mutually proportional 3-D cubes, compare with AF vectors above:

10 First described by R. W. Hamming in HAMMING (1950).

(9a)



(9b)



The distance between any pair of vertices is then equal to the number of edges between them. Beside the null distance between a given ablaut grade and the same ablaut grade (axiom of identity) we face the following examples of distances, see example 10:

(10a) $e\mathrm{G} - o\mathrm{G} = 000 - 100 = e\mathrm{G} - {}_{e}\mathrm{G} = 000 - 010 = e\mathrm{G} - \bar{e}\mathrm{G} = 000 - 001 = 1$ etc.

(10b) $e\mathrm{G} - \bar{o}\mathrm{G} = 000 - 101 = e\mathrm{G} - {}_{o}\mathrm{G} = 000 - 110 = o\mathrm{G} - {}_{\bar{e}}\mathrm{G} = 000 - 011 = 2$ etc.

(10c) $e\mathrm{G} - {}_{\bar{o}}\mathrm{G} = 000 - 111 = o\mathrm{G} - {}_{\bar{e}}\mathrm{G} = 100 - 011 = 3$ etc.

The distance between any pair of ablaut grades is the same without regard to the order of vertices (axiom of symmetry), see example 11:

(11a) $e\mathrm{G} - o\mathrm{G} = 000 - 100 = o\mathrm{G} - e\mathrm{G} = 1$ etc.

(11b) $e\mathrm{G} - \bar{o}\mathrm{G} = 000 - 101 = \bar{o}\mathrm{G} - e\mathrm{G} = 2$ etc.

Such a metric space we term the **fine** metric space, because the distances could be precisely stated and enumerated.

6. Conclusion

In the present paper we have demonstrated that the Old Church Slavonic (and supposedly the Proto-Slavic) ablaut system could be expressed in the terms of vector and/or metric space. Such spaces offer strictly formal description of the whole system and the precise description of the relations between concrete ablaut grades. The proposed description using vector or metric spaces is not a completely new approach to the ablaut, but rather an improvement of the standard knowledge, using formal instruments.

The main difference between our (and Hubey's) approach and the method of DP is that DP "vectors" are not commutative, i.e. the "linear combination" of such vectors differs according to the order of "vectors". That is the reason why we use the term "vector" for DP in the quotations marks, because vectors are generally considered as commutative. On the contrary, our ablaut features form commutable vectors in the full sense of the word.

It should be kept in mind that our solution offers a model not for one solitary root, but for the complex model of the system.

The reader could see in table 5 that the distribution of reduced grades is uneven. Some roots have only *e*-colored grades ({reduced non-lengthened} and {reduced lengthened}), some roots only *o*-colored grades (again {reduced non-lengthened} and {reduced lengthened}). It is in a strong contrast to distribution of full grades that are present for both possibilities of vowel color, either *e*-colored or *o*-colored.

Hence, any individual root has only six grades: {non-reduced non-lengthened *e*-grade}, {non-reduced non-lengthened *o*-grade}, {non-reduced lengthened *e*-grade}, {non-reduced lengthened *o*-grade}, {reduced non-lengthened} and {reduced lengthened}. This model fits to any given root (although many roots are not attested for all those grades), but has its disadvantages: first – four grades are described using three values but two with two values, second – reduced grades of different roots could easily differ in the vocalism (compare roots √*n_s*- and √*g_n*- in the table 5).

The set of all roots could then be split in two disjoint (on the level of reduced grades) subsets: ***e*-roots** and ***o*-roots**. Each subset has its own ablaut pattern which is a sub-pattern of the general ablaut system, as described in the present paper. This opens new possibilities for the describing of root-morphology in Old Church Slavonic.

References

ANDERSON, J. M. & C. J. EWEN (1987) *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

ARUMAA, P. (1964) *Urslavische Grammatik. I. Band – Einleitung, Lautlehre*. Heidelberg: Carl Winter.

BRAINERD, B. (1971) *Introduction to the Mathematics of Language Study*. New York: American Elsevier.

DERKSEN, R. (2008) *Etymological Dictionary of the Slavic Inherited Lexicon.* Leiden: Brill.

DURAND, J. (1996) *Generative and Non-linear Phonology.* London – New York: Longman.

HAMMING, R.. W. (1950) Error Detecting and Error Correcting Codes. *Bell System Technical Journal* 26(2): 147–160.

HUBEY, H. M. (1999) *Mathematical and Computational Linguistics*. München: Lincom Europa 1999.

KORTLANDT, F. H. H. (1972) *Modelling the Phoneme – New Trends in East European Phonemic Theory.* The Hague: Mouton.

KOSMÁK, L. & R. POTŮČEK (2004) *Metrické prostory (Metric Spaces).* Praha: Academia.

KURATOWSKI, K. (1977) *Introduction to Set Theory and Topology.* Warszawa: Pergamon Press.

LIV = RIX, H. (ed.) (1998) *Lexikon der indogermanischen Verben.* Wiesbaden: Dr. Ludwig Reichert Verlag.

MACHEK, V. (1971) *Etymologický slovník jazyka českého.* Praha: Academia.

MARCUS, S. (1967) *Introduction mathématique à la linguistique structurale*. Paris: Dunod.

MEILLET, A. (1934) *Le Slave Commune.* Paris: Libraire ancienne Honoré Champion.

ŠEFČÍK, O. (2008) Values, Features, Fine Metrics and Oppositions. *Linguistica Brunensia* 56: 5–14.

ŠEFČÍK, O. (2009) Preliminary Description of the Czech Phonemic System Using Feature Geometry. In: DOČEKAL, M. & M. ZIKOVÁ (eds.) *Czech in Formal Grammar* 183-196. Muenchen: Lincom Europa.

VAILLANT, A. (1950) *Grammaire comparée des langues slaves. Tome I. Phonétique.* Lyon – Paris: IAC.

VASMER, M. (1953–1958) *Russisches etymologisches Wörterbuch I-III*. Heidelberg: Carl Winter.

Ondřej Šefčík

sefcik@phil.muni.cz

Masaryk University, Brno