

- (a) *myti* ‘wash’ (Modern Russian *myt’*), *po-myti* ‘PFV-wash’, *vy-myti* ‘completely-wash’, *myti sja* ‘wash oneself’, *po-myti sja* ‘PFV-wash oneself’, *vy-myti sja* ‘completely-wash oneself’

'my-l-ŭ	'my-l-a	'my-l-o	'my-l-i
↓ --	↓ --	↓ --	↓ --
po-'my-l-ŭ	po-'my-l-a	po-'my-l-o	po-'my-l-i
- ↓ --	- ↓ --	- ↓ --	- ↓ --
'vy-my-l-ŭ	'vy-my-l-a	'vy-my-l-o	'vy-my-l-i
↓ ↓ --	↓ ↓ --	↓ ↓ --	↓ ↓ --
'my-l-ŭ sja	'my-l-a sja	'my-l-o sja	'my-l-i sja
↓ -- ↓	↓ -- ↓	↓ -- ↓	↓ -- ↓
po-'my-l-ŭ sja	po-'my-l-a sja	po-'my-l-o sja	po-'my-l-i sja
- ↓ -- ↓	- ↓ -- ↓	- ↓ -- ↓	- ↓ -- ↓
'vy-my-l-ŭ sja	'vy-my-l-a sja	'vy-my-l-o sja	'vy-my-l-i sja
↓ ↓ -- ↓	↓ ↓ -- ↓	↓ ↓ -- ↓	↓ ↓ -- ↓

- (b) *moči* ‘be able’

'mog-(l)-ŭ	mo'g-l-a	mo'g-l-o	mo'g-l-i
→ --	→ --	→ --	→ --

- (c) *dati* ‘give’, *pro-dati* ‘sell’, *vy-dati* ‘give away’, *dati sja* ‘give oneself’, *pro-dati sja* ‘sell oneself’, *vy-dati sja* ‘give oneself away’

'da-l-ŭ	da-'l-a	'da-l-o	'da-l-i
---	---	---	---
'pro-da-l-ŭ	pro-da-'l-a	'pro-da-l-o	'pro-da-l-i
- - - -	- - - ↓	- - - -	- - - -
'vy-da-l-ŭ	'vy-da-l-a	'vy-da-l-o	'vy-da-l-i
↓ - - -	↓ - - -	↓ - - -	↓ - - -
da-l-ŭ 'sja	da-'l-a sja	da-l-o 'sja	da-l-i 'sja
--- ↓	--- ↓	--- ↓	--- ↓
pro-da-l-ŭ 'sja	pro-da-'l-a sja	pro-da-l-o 'sja	pro-da-l-i 'sja
- - - ↓	- - - ↓	- - - ↓	- - - ↓
'vy-da-l-ŭ sja	'vy-da-l-a sja	'vy-da-l-o sja	'vy-da-l-i sja
↓ - - - ↓	↓ - - ↓	↓ - - - ↓	↓ - - - ↓

The stress in the verbs with ↓ roots is uniform (fixed left-bound).
 Suffixless verbs with → roots are extremely rare and need not be considered here.
 The stress in the verbs with – roots can be strikingly different in the forms of the same word and in closely related words ('*prodali* ‘they sold’ ~ '*dali* ‘they gave’ ~ '*proda lis*’ ‘they sold themselves’) ⇒ a likely target for analogy.

2.3. Changes in stress of the past tense forms:

- (i) stress definalization (retraction)—stress shifts from the open final syllable to the penult (especially in western dialects, cf. Zalizniak 1985: 182–8);
- (ii) levelling within the paradigm of a verb;
- (iii) levelling across paradigms of closely related verbs.

2.4. Dominant patterns in Modern Standard Russian (Zalizniak 1977: 80–1; Timberlake 2004: 100–1):

(a)	'my-l-Ø	'my-l-a	'my-l-o	'my-l-i
	wash-PST-M.SG	wash-PST-F.SG	wash-PST-N.SG	wash-PST-PL
(b)	'vë-l-Ø	ve-'l-a	ve-'l-o	ve-'l-i
	lead-PST-M.SG	lead-PST-F.SG	lead-PST-N.SG	lead-PST-PL
(c)	'da-l-Ø	da-'l-a	'da-l-o	'da-l-i
	give-PST-M.SG	give-PST-F.SG	give-PST-N.SG	give-PST-PL

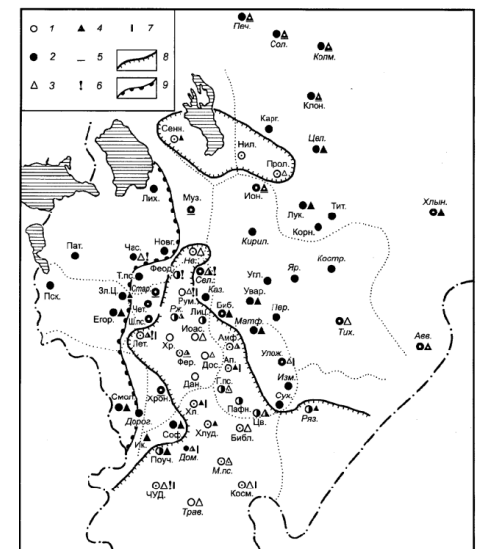
- (a): stress falls on the same syllable of the stem in all forms;
- (b): stress falls on the ending in F.SG, N.SG, and PL; in M.SG, where there is no overt ending, it falls on the stem;
- (c) stress falls on the ending in the feminine singular and on the stem in the other three forms.

The majority of type (b) verbs as well as (c) verbs originate from verbs with – roots.

3. Parameters of variation

3.1. Inter-speaker variation

Dialect: cf. map from (Zalizniak 2014: 97) where line 9 marks the border between initial stress in M.SG, N.SG and PL forms ('*prodal* ‘he sold’, '*prodalo* ‘it sold’, '*prodali* ‘they sold’ to the east of the line) and peninitial stress (*pro'dal* ‘he sold’, *pro'dalo* ‘it sold’, *pro'dali* ‘they sold’ to the west of the line) in the 16th century.



Generation: cf. *new* in “o'bnjalo and acceptable new obnja'lo, 'obnjali and acceptable o'bnjali” (Kalenčuk et al. 2012: 451) in 1.2.

Educational attainment: educated speakers are more likely to follow the advice of prescriptive dictionaries.

3.2. Intra-speaker variation:

Variation across different speech styles

Context-dependent variation: in some contexts, there is a statistical (but not absolute) preference for one variant form over the others.

4. Corpora as a (somewhat unsatisfying) resource for studying variation

4.1. Available resources:

4.1.1. Spoken corpora.ru: 34,000 tokens, spoken texts on pre-defined topics, obtained under experimental conditions.

Drawbacks:

- 1) a very small corpus;
- 2) limited vocabulary.

4.1.2. Russian National Corpus (www.ruscorpora.ru), Spoken subcorpus: 11.3m tokens from various sources.

Drawbacks:

- 1) a very heterogeneous corpus (transcripts from older books on spoken Russian, texts from collections of various Russian universities, movie scripts, ...);
- 2) no uniform metadata ⇒ difficult to study inter-speaker as well as intra-speaker variation;
- 3) stress is marked only in a small proportion of texts.

4.1.3. Russian National Corpus (www.ruscorpora.ru), Poetic subcorpus: 11m tokens, approx. 880 authors, 18th through 21st century.

Russian poetry was predominantly syllabo-tonic until the 2nd half of the 20th century ⇒ a valuable resource for studying the history of Russian stress as well as intra-speaker variation.

4.2. Checking the existence of variation:

4.2.1. The Spoken subcorpus of the Russian National Corpus can be used to check the existence of variation.

Corpus size without movie scripts: 7.8m tokens.

117 forms of 70 verbs (31 M.SG, 37 F.SG, 15 N.SG, 34 PL) exhibit at least some variation.

4.2.2. Examples

podo'bralas' × 1 (50%) ~ *podobra'las'* × 1 (50%) ‘she got nearer’;

'zanjalsja × 5 (33%) ~ *za'njalsja* × 3 (20%) ~ *zanjal'sja* × 7 (47%) ‘he occupied himself’

'prodal × 7 (50%) ~ *pro'dal* × 7 (50%) ‘he sold’

'obnjal × 5 (71%) ~ *ob'njal* × 2 (29%) ‘he hugged’

'pribyli × 28 (97%) ~ *pri'byli* × 1 (3%) ‘they arrived’

4.2.3. Quantitative analysis

These results are informative to different degrees (cf. *podobralas'* ‘she got nearer’ × 2 and *zanjalsja* ‘he occupied himself’ × 15).

Herfindahl–Hirschmann index (HHI): the probability that two tokens of the same word form randomly selected from a corpus have the same stress (Segura & Rodríguez Braun 2004: 111):

$$HHI = \sum_{i=1}^n p_i^2$$

(corpus frequency serves as an estimator of p)

4.2.4. 10 verbs with the lowest Herfindahl–Hirschmann index (i.e. most variable verbs) occurring at least 5 times:

Variant forms	Gloss	HHI
<i>'ožila</i> × 2 (40%) ~ <i>o'žila</i> × 1 (20%) ~ <i>oži'la</i> × 2 (40%)	‘she returned to life’	0.36
<i>'zanjalsja</i> × 5 (33%) ~ <i>za'njalsja</i> × 3 (20%) ~ <i>zanjal'sja</i> × 7 (47%)	‘he occupied himself’	0.369
<i>'načalsja</i> × 16 (22%) ~ <i>na'čalsja</i> × 16 (22%) ~ <i>načal'sja</i> × 40 (56%)	‘he started’	0.407
<i>'podnjal</i> × 16 (50%) ~ <i>po'dnjal</i> × 16 (50%)	‘he raised’	0.5
<i>'prodal</i> × 7 (50%) ~ <i>pro'dal</i> × 7 (50%)	‘he sold’	0.5
<i>'podnjali</i> × 4 (50%) ~ <i>po'dnjali</i> × 4 (50%)	‘they raised’	0.5
<i>'predali</i> × 4 (50%) ~ <i>pre'dali</i> × 4 (50%)	‘they betrayed’	0.5
<i>u'dalsja</i> × 3 (50%) ~ <i>udal'sja</i> × 3 (50%)	‘he succeeded’	0.5
<i>'sozdal</i> × 47 (49.5%) ~ <i>so'zdal</i> × 48 (50.5%)	‘he created’	0.5001
<i>'prožil</i> × 19 (53%) ~ <i>pro'žil</i> × 17 (47%)	‘he lived (for a period of time)’	0.502

The ranking should take into account the number of occurrences in a more sophisticated way.

The presence of variation in the corpus proves that the variation exists, but the absence of variation from the corpus does not prove anything, since many word forms are represented by one or two tokens only.

The number of examples is too small to study inter-speaker as well as intra-speaker variation.

4.3. Intra-speaker variation:

Poetic subcorpus of the Russian National Corpus can be used to study intra-speaker variation, because for some poets it contains very large bodies of text.

Poet	Tokens (total)	Past tense forms with variation		Frequency of past tense forms with variation
		Types	Tokens	
Vasily Zhukovsky (1783–1852)	249,051	20	166	1:1500
Nikolay Nekrasov (1821–1878)	172,322	26	89	1:1936
Valery Bryusov (1873–1924)	185,856	18	110	1:1690

Approx. 1 token in 1700 used by a speaker is an instance of possible variation.

The variation in individual word-forms can be persistent over time:

Zhukovsky:

'podnjal × 11 ~ *po* *'dnjal* × 7 'he raised'

po *dnjalsja* × 15 ~ *podnjal* *'sja* × 4 'he rose'

'otnjal × 2 ~ *ot* *njal* × 2 'he deprived'

ro *dilsja* × 4 ~ *rodil* *'sja* × 6 'he was born'

Drawbacks:

- 1) this might be individual language change rather than synchronic variation;
- 2) if a poet has two variant forms at his/her disposal, the choice is determined by neighbouring words—but the reason is that it is syllabotonic poetry rather than there is something in the language.

5. Rhythmic Hypothesis

5.1. We claim that for Standard and, maybe, other varieties of Modern Russian the following hypothesis holds true:

Rhythmic Hypothesis

Stress placement in Russian word forms with variable stress is at least partially determined by the immediate phonetic context. The optimal surface representation is in accord with the Principle of Rhythmic Alternation (for PRA see e.g. Schläuter 2015).

Russian is a stress-timed language so rhythm sensitivity should seem quite natural (Auer & Uhmann 1988).

5.2. The Rhythmic Hypothesis (*RH*) in its general form is almost impossible to falsify ⇒ narrower formulations wanted.

Rhythmic alternations are usually assumed to occur more often “in closely-knit units” (Schlüter 2005: 30), so we arrived at *RH'*:

RH'

In V+DO sequences a configuration like ...σ'σ# 'σσ..., where the first word is one of the verbs in question, is less likely to occur than ...'σσ# 'σσ...; vice versa, ...σ'σ# σ'σ... is more likely than ...'σσ# σ'σ...

Other word classes, as well as other relevant types of context remain for further investigation (in progress):

- (i) DOs like σσ'σ..., σσσ'σ..., etc.;
- (ii) following words other than DOs inside one phonological phrase;
- (iii) following words over the border of phonological / intonational phrase;
- (iv) sentence-final position;
- (v) preceding words, i.e. left context.

6. Experiment

6.1. In our first experiment we aimed to test *RH'*.

6.2. Experimental design:

“Please read aloud the sentences below opening the brackets.

Example:

U Miti v komnate živut 2 (popugaj) →
At Mitya in room live 2 parrot

U Miti v komnate živut dva popugaja
At Mitya in room live two parrots

“There are 2 (parrot) living in Mitya’s room → There are 2 parrots living in Mitya’s room’.”

18 sentences:

1. *Dima vernulsja domoj, uvidev, čto na (ulica)idët dožd'.*
Dima returned home when he noticed that it was raining outside'
2. *Anton prodal braslet, kotoryj dostalsja emu v nasledstvo ot (babushka).*
Anton sold bracelet which found.itself to.him in inheritance from grandmother
'Anton sold the bracelet that he inherited from his grandmother'.
3. ...

4 sentences with the verb form *prodal* 'he sold', 2 sentences with the form *obnjaj* 'he hugged', and 12 filler sentences.

At the end, the respondents were asked whether they guessed the purpose of the experiment.

6.3. Results:

6.3.1. Number of respondents

$N_0 = 87$ respondents; 11 of them (13%) grasped the idea and were excluded from further consideration $\Rightarrow N = 76$.

6.3.2. Overall variability

	Stress on the 1 st syllable	Stress on the 2 nd syllable	
<i>prodal</i>	67 (44%)	85 (56%)	152 (100%)
<i>obnjaj</i>	167 (55%)	137 (45%)	304 (100%)
	234	222	456

\Rightarrow The two verbs are different from each other in respect to the preferred stressed syllable (two-sample proportion test, $\chi^2 = 4.355, p = 0.037$).

Cf. data from the Spoken subcorpus of the Russian National Corpus in 4.2.2:

	Stress on the 1 st syllable	Stress on the 2 nd syllable	
<i>prodal</i>	7 (50%)	7 (50%)	14 (100%)
<i>obnjaj</i>	5 (71%)	2 (29%)	7 (100%)
	12	9	21

Different patterns of stress assignment with different types of DO:

	Stress on the 1 st syllable	Stress on the 2 nd syllable	
' <i>daču</i> 'dacha', ' <i>knigu</i> 'book'	94 (62%)	58 (38%)	152 (100%)
<i>bra'slet</i> 'bracelet', <i>por'tret</i> 'portrait'	73 (48%)	79 (52%)	152 (100%)
	167	137	304

\Rightarrow Frequencies of '*prodal* and *pro'dal* depend on the stress of DO (two-sample proportion test, $\chi^2 = 5.315, p = 0.021$).

	Stress on the 1 st syllable	Stress on the 2 nd syllable	
' <i>Anju</i> 'Anya' <i>se'stru</i> 'sister'	39 (51%) 28 (37%)	37 (49%) 48 (63%)	76 (100%) 76 (100%)
	67	85	152

\Rightarrow Frequencies of '*obnjaj* and *ob'njal* seem to depend on the stress of DO, but the difference is not statistically significant (two-sample proportion test, $\chi^2 = 2.669, p = 0.102$).

prodal, obnjaj	Stress on the 1 st syllable	Stress on the 2 nd syllable	
' $\sigma\sigma$	133 (58%)	95 (42%)	228 (100%)
$\sigma'\sigma$	101 (44%)	127 (56%)	228 (100%)
	234	222	456

\Rightarrow When analyzed cumulatively, the verbs *prodal* and *obnjaj* show a definite tendency to adjust their stress to the stress of DO (two-sample proportion test, $\chi^2 = 8.436, p = 0.004$).

6.3.3. Intra-speaker variation

35 out of 76 (46%) respondents exhibit variation in at least one of the verbs.

Intra-speaker variation for *obnjaj*:

ob'njal se'stru ~ '*obnjaj* 'Anju 16 (73%)

'*obnjaj se'stru* ~ *ob'njal* 'Anju 6 (27%)

\Rightarrow Intra-speaker variation in case of *obnjaj* seems to conform to *RH'*, but the result is not statistically significant (exact binomial test, $p = 0.052$).

Intra-speaker variation for *prodal*:

'*prodal* \times 2 ('*daču*, '*knigu*) ~ *pro'dal* \times 2 (*bra'slet*, *por'tret*) 3

No other types of '*prodal* \times 2 ~ *pro'dal* \times 2 variation other than the perfect *RH'*-conforming distribution are attested.

'*prodal* \times 3 ~ *pro'dal* \times 110 (59%)

'*prodal* \times 1 ~ *pro'dal* \times 3 7 (41%)

If we assume that the respondents with a '*prodal* \times 3 ~ *pro'dal* \times 1 or '*prodal* \times 1 ~ *pro'dal* \times 3 variation have a dominant variant and a deviating one, we can check where the deviation occurs. In 16 out of 17 cases (94%), the deviation conforms to the *RH'* (exact binomial test, $p = 0.0003$).

6.4. Interpretation:

The experimental data confirms that there are some statistical tendencies conforming to the Rhythmic Hypothesis in its restricted form (*RH'*). It holds true at least for some verb forms and for direct objects immediately following the verb.³

7. Towards a phonological representation

7.1. If we assume a division of storage and computation, we are left with at least 4 logically possible ways to look at such stress variation:

- (a) No information about stress is stored in the lexicon; the assignment of stress takes place on later stages of spell-out.

³Two verb forms are clearly not representative of the whole Russian language, but we are in the process of running similar experiments with other verbs and different contexts.

(b) The information about stress is stored in the lexicon; a single form is marked for stress only once; special rules on later stages of spell-out reassign this mark (à la Liberman, Prince 1977 et seq.).

(c) The information about stress is stored in the lexicon; a single form is marked for stress more than once; on later stages of spell-out all surplus marks are deleted (à la Gussenhoven 1991 for English).

(d) The information about stress is stored in the lexicon; there are several competing forms, each carrying a single stress mark; on later stages of spell-out the most “appropriate” one is chosen.

	Number of forms	Number of stress marks	Operations with stress marks	Choosing the optimal form
(a)	single	none	adding	no
(b)	single	single	moving	no
(c)	single	multiple	deleting	no
(d)	multiple	single	none	yes

7.2. Most of these are problematic when applied to the data at hand:

(a) Russian has a morphologically dependent lexical stress system (see above), the vast majority of lexical items are marked for stress ⇒ it would be unusual to assume that a small group of verbs is not.

Still, cases like this have been evidenced: enclitomena in Old Russian had no stressed syllables and were assigned an “automatic” initial stress in the absence of stressed forms in the vicinity (more likely phrasal prominence along the lines of Gordon 2011).

There is no phonetic difference between our verbs and other words, and it could be expected if they were unmarked for stress.

(b) If we assume that there is only one form with one stress mark, then how do we determine the syllable that should bear it? By recourse to historical data? With the help of statistical data? Then the choice is more or less obvious with a 95% : 5% distribution but not at all obvious in the case of a 55% : 45% distribution; what counts as enough to choose the underlyingly stressed syllable?

Having established this, we’ll need to draw a line between clashes and lapses since some forms will receive stress on the 1st syllable, some—on the 2nd syllable, etc. Clashes and lapses are usually analyzed in different terms, so we’ll have to multiply posited rules, which we will not have to do in (a), (c) or (d) where we don’t get to change stress markings of a form in a given context.

(c) Firstly, in Gussenhoven’s (1991) analysis only feet heads can be marked with ‘accent’ (*). It is usually assumed after Halle and Vergnaud (1987) that in Russian only the accented and the pretonic syllable constitute the foot, all other syllables being unfooted. So under this

assumption a bisyllabic form can never comprise two feet: it is either (σ'σ) or ('σ)σ. Since only feet can attract accent, two accents are also impossible. It is also in words with secondary stresses that multiple stress marks are actually added, which is obviously not an option for forms like *prodal*.

(d) The easiest solution given there are speakers with hardly any variation. Those who do show variation presumably store competing forms in the lexicon, whereas others store just one of the forms. The structure of these forms remains the same across speakers.

We have also seen that many speakers demonstrate a preference for a “basic” variant in experimental data, and the deviations from it tend to repair the rhythm. This could also be used in (b) to determine underlying stresses but is equally fitting here, with one form being more salient for a given speaker.

Note that this option could be implemented in OT terms save that we’ll probably have to postulate different constraint rankings for different speakers, which moreover would not be obligatory (to allow sequences like *pro'dal 'knigu* for speakers with variation). We would also have to tweak GEN to outlaw some surface forms which are not phonologically ill-formed but can never occur (like **proda'li* ‘they sold’, cf. *uve'li* ‘they led away’).

In this case, the details of how the speakers choose one of the forms must be elaborated.

To sum up, (b) and (d) with one stress in a form look more plausible than (a) or (c).

8. Concluding remarks

Many questions remain (and we hope to investigate them more closely):

- (i) How different verbs behave in different contexts (see 5.2)?
- (ii) Is there place in Russian data for effects similar to the findings of Kelly and Bock (1988), who contend that typical syntactic contexts predispose lexical classes—not single forms!—to certain stress patterns in English? Might this be the mechanism of congealing stress patterns in a paradigm?
- (iii) What is the sociolinguistic distribution of speakers with variable stress / different variants of fixed stress?

But the variation of stress in Russian past tense verbal forms is nonetheless instructive in many aspects:

- (i) we can actually see how accentual variation springs up;
- (ii) we can see how considerations of rhythm can influence lexical stress;
- (iii) it can prove a test case for phonological theories of variation.

References

- Auer, Peter and Susanne Uhmann. 1988. Silben- und akzentzählende Sprachen: Literaturüberblick und Diskussion. *Zeitschrift für Sprachwissenschaft* 7. 214–59.
- Avanesov, Ruben (ed.). 1988. *Orfoèpičeskij slovar' russkogo jazyka (Russian pronouncing dictionary)*. 4th ed. Moscow: Russkij jazyk.
- Gordon, Matthew. 2011. Metrical stress theory. In *The handbook of phonological theory*, ed. by John Goldsmith, Jason Riggle, and Alan Yu. Malden: Wiley-Blackwell. 141–63.
- Gussenhoven, Carlos. 1991. The English rhythm rule as an accent deletion rule. *Phonology* 8. 1–35.
- Halle, Morris and Jean-Roger Vergnaud. 1987. *An essay on stress*. Cambridge, Mass.: MIT Press.
- Kalenčuk, Maria, Leonid Kasatkin, and Rozalia Kasatkina. 2012. *Bol'šoj orfoèpičeskij slovar' russkogo jazyka (A comprehensive pronouncing dictionary of Russian)*. Moscow: AST-Press.
- Kelly, Michael H. and J. Kathryn Bock. 1988. Stress in time. *Journal of Experimental Psychology: Human Perception and Performance* 14. 389–403.
- Lieberman, Mark and Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8. 249–336.
- Schlüter, Julia. 2005. *Rhythmic grammar: The influence of rhythm on grammatical variation and change in English*. (Topics in English Linguistics 46.) Berlin: de Gruyter Mouton.
- Schlüter, Julia. 2015. Rhythmic influence on grammar: Scope and limitations. In *Rhythm in cognition and grammar: A Germanic perspective*, ed. by Ralf Vogel and Ruben van de Vijver. Berlin: de Gruyter Mouton. 179–206.
- Segura, Julio and Carlos Rodríguez Braun. 2004. *An eponymous dictionary of economics: A guide to laws and theorems named after economists*. Cheltenham: Edward Elgar.
- Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge, UK: Cambridge University Press.
- Zalizniak, Andrey. 1977. *Grammatičeskij slovar' russkogo jazyka (A grammatical dictionary of Russian)*. Moscow: Russkij jazyk.
- Zalizniak, Andrey. 1985. *Ot praslavjanskoj akcentuacii k russkoj (From Proto-Slavic to Russian accentuation)*. Moscow: Nauka.
- Zalizniak, Andrey. 2014. *Drevnerusskoe udarenie: Obščie svedenija i slovar' (Old Russian stress: General remarks and a dictionary)*. Moscow: Jazyki slavjanskoj kul'tury.