Computational Linguistics and Intellectual Technologies:
Proceedings of the International Conference "Dialogue 2022"

Moscow, June 15–18, 2022

Historical Text Corpora and the Conclusiveness of Linguistic Analysis

Anton Zimmerling
Pushkin State Russian Language
Institute, Moscow, Russia;
Institute of Linguistics, Russian
Academy of Science, Moscow, Russia
fagraey64@hotmail.com

Abstract

I discuss the methodology and conclusiveness of the corpus-based historical linguistics and analyze two formal models predicting the language-internal variation in Early Old Russian syntax. Linguistic models claiming a rigid distribution of grammatical features like ± overt realization of agreement markers activate hidden corpus characteristics such as profiles of text genres, chronology, vector of change, ± impact of L2, ± presence of supra-dialect features. In this case they can be valued and checked on text samples, where genre features are stable, while location and time vary.

Keywords: historical text corpora; conclusiveness; predictive linguistic models; Old Russian language
DOI: 10.28995/2075-7182-2022-21-586-593

Исторические корпуса текстов и лингвистическая доказательность

Циммерлинг А. В.
Государственный институт русского языка
имени А. С. Пушкина, Москва, Россия
Институт языкознания РАН, Москва, Россия
fagraey64@hotmail.com

Аннотация

В статье обсуждается доказательность лингвистических исследований, опирающихся на исторически корпуса текстов? и проверяются две формальные модели, предсказывающие распределение показателей согласования в древнерусском языке. Формальные лингвистические модели проверяемы в той степени, в которой они имплицитно опираются на скрытые характеристики корпусов, в том числе на профили жанров текста. Модельное представление о наличии в древнерусском языке нулевой связки 3 л. подтвердилось, однако тезис о том, что употребление ненулевых связок перфекта 3л. в XII в. всегда отражает влияние второго языка автора, опровергается на выборке текстов, представляющих разные древнерусские диалекты.

Ключевые слова: исторические корпуса текстов; доказательность; предиктивные лингвистические модели; древнерусский язык

1 Variation in Historical Text Corpora

Variation of grammatical parameters is a normal state for all natural languages, including old languages as they are reflected in historical text corpora. Linguistic models aim at predicting the distribution of language features in corpora and text collections. Parametric models generally have a greater predictive force, since they explain a visible diversity of features by different settings of the same constraint. E.g., predicative instrumental case ($Instr_{PRED}$) in Modern Russian is a case form assigned to predicative nominals or adjectives by the clausal subject. Standard accounts of Russian grammar, cf. [Bailyn 2012] fail to explain sentences like ?*мне*$_{DAT}$ *стало*$_{PAST.3SG.N}$ ***известным***$_{INSTR.PRED}$, [$_{CP}$ *что Вы в городе*], 'I got to know that you are in the city,' lit. 'to-me became known, that…' accepted by a minor part of Russian speakers. Such sentences are puzzling, since they lack any nominative subject. A parametric approach explains this variation and helps to find the invariant. Both groups of Russian speakers have the subject control parameter. The feature that varies across speakers is not the correlation between $Instr_{PRED}$ and subjects, but the status of sentential arguments like [$_{CP}$ *что Вы в городе*]: the minority of speakers licenses them as assigners of $Instr_{PRED}$, while the majority does not [Циммерлинг 2018]. An extreme case of language-internal variation can be described in terms of dialectal divergences, if a split in parameter values between different groups of speakers is consistent [Henry 1998].

Historical corpora pose an extra problem: it is not always clear how to classify the author's idiom in a situation of language-internal variation. Assume that there are two expressions α and β that can be used in the same context, but only part of the speakers licenses them both. Let us also assume that the corpus has a system of genre tags and all texts representing genre G1, $T \in \{T_{G1}\}$ only display the option α, while all texts representing genre G2, $T \in \{T_{G2}\}$ display both α and β. This is exactly what one finds in Old Russian and historical corpora like RNC Old Russian corpus https://ruscorpora.ru/new/search-old_rus.html. The options α and β are values of the parameter that licensed either a zero form of the 3rd person BE-auxiliary in the Old Russian perfect construction ($\alpha = \varnothing^{3P}$), or a non-zero form ($\beta$ = *есть*$_{3SG}$, *суть*$_{3PL}$, *еста*$_{3DU}$) in this construction. G1 can be identified with Old Russian birch bark letters, graffiti and other non-bookish texts. At present moment, only birch bark letters are included in RNC and placed in a separate section of RNC Old Russian corpus https://ruscorpora.ru/new/search-birchbark.html, so that G1 can be associated in RNC with a single text genre. G2 are Old Russian bookish texts of the XI—XIII centuries. The texts from the RNC Old Russian corpus largely overlap with G2, with one exception: "Russkaja Pravda" has a characteristic G1 grammar and is a non-bookish text that should be classified with birch bark letters, if the demarcation of sub-corpora were solely based on the G1 vs G2 distinction. The prediction is that texts from G1 only have the option *пришелъ*$_{SG.M}$ $\varnothing^{3P}$ '<he> has come' with the zero copula, while the texts from G2 use both the zero copula and the overt 3rd person copulas *пришелъ*$_{SG.M}$ $\varnothing^{3P}$ ~ *пришелъ*$_{SG.M}$ *есть*$_{3SG}$. It is tempting to assume that the actual Old Russian usage was associated with G1, while overt 3rd person copulas in the perfect construction one finds in G2 were a tribute to the written tradition [Хабургаев 1978; Зализняк 2008: 236]. The idea that the speakers modified their grammar and added an extra option $<\alpha> \Rightarrow <\alpha, \beta>$, when they switched from colloquial speech (G1) to bookish texts (G2) is natural and almost trivial. A question however arises, whether this intuition can be checked empirically on corpora, since the absence of β is a definitional property of G1 and T_{G1} and the presence of β is a definitional property of G2 and T_{G2}.

The option β is not the only characteristics of G2. Semi-formally, G2 contains hybrid Old Russian texts, heavily influenced by a different, though closely related language, Old Church Slavonic. The notion of hybrid text can be formalized. The basic hypothesis is that the first language, L1 (= Old Russian, as reflected in G1) is suppressed in G2 by the second language, L2 (= Old Church Slavonic), while the texts from G1 are monolingual. Again, there is a methodological question, how one can prove that all speakers of Old Russian invariably switched to L2, when they used the option β and wrote down the 3rd person perfect form with overt copulas as *пришелъ есть*, *пришли суть*. A related methodological issue is whether the negative evidence from G1, where the option β is not displayed is conclusive and one can establish for sure that L1 Old Russian lacked any overt copulas in the 3rd person perfect form. For the sake of simplicity one can assume that the sets $\{T_{G1}\}$ and $\{T_{G2}\}$ are closed, since the possibilities of adding new texts to historical corpora are limited. The actual situation is different: recently discovered

birch bark letters are added to T_{G1}, while T_{G2} can be expanded by other bookish texts or alternative versions of the texts already included in Old Russian corpora. Regarding the illustration provided, the discovery of overt 3rd person perfect auxiliaries and sentences like *пришелъ есть* in a recently added birch bark letter will falsify the initial hypothesis that G1 completely lacks the option β, while the discovery of a bookish text, whose author generalized the use of zero copula in sentences like $\varnothing^{3P}$ *пришелъ*$_{SG.M}$ will challenge the initial hypothesis that the option β is visible in all texts from G2. The chosen model can rule out these exceptions as insignificant by claiming that the author of that birch letter reverted to G2/L2 grammar, while the author of that bookish text reverted to G1/L1 grammar, but such explanations are apparently circular and not conclusive.

I argue that historical corpus linguistics still has some tools to check similar hypotheses concerning the variation in a dead language, but one has to narrow the search and first establish a group of texts relevant for checking each grammatical feature.

2 Old Russian Grammar and Two Linguistic Models

Old Russian perfect is an analytical construction consisting of an active *l*-participle discriminating number and gender, and an auxiliary/copula (inflected present tense indicative BE-form) expressing the person-and-number agreement with a subject controller, if any. The grammar of Old Russian perfect is described by two linguistic models proposed by Andrej A. Zaliznjak (1993, 2008) and developed by other scholars.

2.1 Zaliznjak's Model A

The first hypothesis of Zaliznjak, henceforth — 'Zaliznjak-A' claims that the agreement markers used in the perfect construction were clitics. The 1st -2nd person perfect auxiliaries were clausal-second enclitics, i.e. enclitics taking the second position in clause, while Old Russian lacked overt 3rd person perfect clitics and replaced them with a zero 3rd person auxiliary [Zaliznjak 1993: 285].

(i) <Zaliznjak-A> Old Russian agreement markers in the perfect construction were clitics with a fixed distribution. Such markers developed only in the 1st -2nd persons, but not in the 3rd person, where a zero auxiliary (copula) with feature '3rd person' was used instead. Therefore overt agreement markers in the 3rd person were redundant.

Although Zaliznjak himself did not postulate a zero 3rd person BE-copula for Old Russian, this step made in [Zimmerling 2020; Циммерлинг 2021: 26-27] seems inevitable, since the conclusion that colloquial Old Russian completely lacked overt 3rd person perfect auxiliaries is based on the assumption that the absence of an overt agreement marker signals the agreement feature '3rd person'[1]. The tag 'colloquial Old Russian' used by Zaliznjak or its synonym 'Old Russian 1, OR-1' introduced in [Zimmerling 2020] and the tag 'G1' used in this paper, refer to the same thing — the Old Russian idiom without overt 3rd person perfect auxiliaries. The difference is that OR-1 is an aprioristic model notion that is abstracted from with any collections of real texts, while G1 is a set of texts representing some specific genres, where (i) holds true.

2.2 Zaliznjak's Model B

The second hypothesis of Zaliznjak, henceforth — 'Zaliznjak-B' predicts the syntax of overt 3rd person BE-copulas, including 3rd person perfect auxiliaries in G2, i.e. hybrid bookish Old Russian texts. The

[1] A descriptive alternative — to postulate a defective morphological present tense paradigm for the Old Russian copula BE instead of recognizing the zero copula $\varnothing^{BE.PRES.3P}$ is not satisfactory, since Old Russian retained overt forms of the full verb BE.

prediction is that 3rd person BE-copulas form in G2 a complementary distribution with clausal subjects [Зализняк 2008: 240, 258-259].

(ii) <Zaliznjak-B> Old Russian agreement markers including 3rd person BE-auxiliaries tend to form a complementary distribution with the agreement controllers (grammatical subjects).

I henceforth call the idiom associated with G2 ‘Old Russian-2’, or ‘OR-2’.

2.3 Past Tenses in OR-1 and OR-2

There is a sharp split between OR-1 and OR-2 regarding the use of past tenses. Hybrid texts from G2 have the OR-2 grammar and use simple past forms (aorist and imperfect) along with perfect. The texts from G1 that have the OR-1 grammar either generalize the use of perfect as the only extant past tense form or use aorist and imperfect in insignificant quantities, cf. [Zaliznjak 2004: 142]. Thus, OR-2 is an idiom which 1) has aorist and imperfect, 2) licenses overt 3rd person perfect auxiliaries.

2.4 Old Russian Dialects

There are four major groups of Old Russian dialects: a) Southern; b) North-Western (Novgorod and Pskov), c) Western (Smolensk and Polotsk), d) Great Eastern (e.g., Suzdal). Dialectal markers can found both in texts from G1 and in texts from G2, though in the latter group they are often undesirable. That means that the presence/absence of overt 3rd person auxiliaries can also depend on the dialect, although this scenario has not been proven so far.

3 Analysis

I check models A and B on a sample of five non-anonymous G2 Old Russian texts of the XII century representing three dialects and measure three features: 1) the ratio of perfect and simple past tenses; 2) the presence of overt 3rd person perfect auxiliaries; 3) the correlation between overt 3rd person perfect auxiliaries and overt clausal subjects[2]. The control group of texts includes four G2 XIV-XV century texts of comparable genre.

3.1 The Main Group

The main group includes five non-anonymous author texts representing three dialects and seven idioms. “Daniil’s Pilgrimage” and “Vladimir Monomakh’s Edification” are South Russian texts dated by the first quarter of the XII century. Three texts come from Novgorod: “Kirik’s Inquiries” are dated by ca. 1132-1156 A.D., “The Edification of Elija-John” is dated by 1166 A.D., while “Antonius’ Pilgrimage” is dated with ca. 1200. Two texts represent the genre of *khozhenie*, i.e. pilgrimage, two texts — the genre of *pouchenie*, i.e. edification. “Kirik’s Inquiries” are written in form a dialogue between church hierarchs and ordinary clerics who ask them concerning various issues in canonic law. Its main part is written by Kirik, while two supplements are written by Savva and Elija.

Four of these texts are currently added to RNC Old Russian corpus as separate entries, the data from “Monomakh’s Edification” was in March 2022 only available as part of the larger compilation

[2] An anonymous reviewer argues that from the corpus linguistics perspective, modeling based on one parameter may lead to overgeneralization. In this paper we however used three parameters — 1) ± overt 3rd person perfect auxiliaries, 2) ± complementary distribution of auxiliaries and agreement controllers, 3) ratio of the perfect clauses from the total amount of three past tense forms — for measuring the variation in XII century Old Russian grammar. The analysis confirmed that these three parameters are independent.

("Laurentius' Chronicle")[3]. For the remaining four texts, RNC tells the number of words and sentences, but not the number of clauses, which is the relevant parameter for the processing of perfect clauses. The texts from RNC Old Russian corpus are not annotated syntactically. "Monomakh's Edification" consists of several parts. I ignored the last one — "The Lenten Prayer", since it does not contain any perfect constructions. Other four texts from main group were processed and annotated in their entirety. I also checked an extra version of "Kirik's Inquiries", which differs from the version included in RNC.

3.2 The Control Group

The control group includes two Novgorod texts from the mid XIV century — "Stefan's Pilgrimage" and "The Discourse about Paradise" by Vasiliy Kalika — and two Suzdal texts from the mid XV century — "The Pilgrimage to the Florence Council" and "A Note about Rome". Three of them represent the genre of *khozhenie*, while Vasiliy Kalika's text is comparable to Elijah-John's "Edification".

3.3 Syntactic Annotation

The text fragments containing the perfect construction in the 3rd person were extracted and parsed in order to check model B. 3rd person perfect clauses were assigned the value '+SUBJ' if they contained an overt subject expressed by the accented noun/pronoun or NP, otherwise they were assigned the value '-SUBJ', even if the subject position is reconstructed in a larger syntactic context.[4]

3.4 Disambiaguation of Contexts and Semantic Annotation

In order to check model A, all contexts containing overt 3rd person perfect auxiliaries were annotated semantically. Standard perfect auxiliaries, including 3rd person auxiliaries *есмь*^{1}, *сумь*^{1}, *ecma*^{1} are forms lacking the phrasal accent: they are placed clause-internally and do not bring about existential or verificational meanings. Existential /verificational 3rd person auxiliaries *есмь*^{2}, *сумь*^{2}, *ecma*^{3} are accented words with a free linear distribution. They can take clause-initial position, adjoin to initial proclitics *и, а*, *нъ* or stand clause-internally. Sentences with standard non-existential & non-verificational forms *есмь*^{1}, *сумь*^{1}, *ecma*^{1} were assigned the value 'PERFECT I', sentences with existential or verificational *есмь*^{2}, *сумь*^{2}, *ecma*^{2} were assigned the value 'PERFECT II'. Some contexts with clause-internal 3rd person perfect auxiliaries are two-way ambiguous between PERFECT I and PERFECT II. PERFECT I is a standard construction with person-and-number agreement, where the alternation $\varnothing^{3P} \sim$ *есмь*^{1}, *сумь*^{1}, *ecma*^{1} is not semantically-driven. PERFECT II is possible only in the 3rd person and shows number, but not person agreement: *есмь*^{2}, *сумь*^{2}, *ecma*^{2} bring about the diagnostic existential or verificational semantics and do not alternate with $\varnothing^{3P}$.

3.5 The Ratio of Perfect vs Simple Past

All perfect clauses in 1-3 persons and all clauses containing aorist and imperfect forms were counted. The analytic plus perfect forms and other copular constructions with *l*-participles were not considered. The ratio of perfect clauses (P) from the total amount of three past tense forms (A & I & P), k_{PERF} = P/A + I +P was measured in order to estimate, whether P had a prominent status in the idiom of each author.

[3] By May 2022, the situation has changed: Monomakh's "Edification" and his "Lenten Prayer" were added to RNC Old Russian corpus as separate entries, while the entire "Laurentius' Chronicle" is no longer indexed as a part of RNC.

[4] At this stage, syntactic annotation was done manually. The assignment of the values '+SUBJ' and '-SUBJ' according to the principle outlined in the section **3.3.** requires a modification of the existing parsers.

3.6 Direct Quotations

The texts included several direct quotations from the Bible. In case these quotations included overt 3rd person perfect auxiliaries, the corresponding examples were considered borrowings from Old Church Slavonic and excluded from the sample.

4 Results and Discussion

4.1 Model A and Overt 3rd Person Auxiliaries

The results of this study partly falsify Zaliznjak's Model A. The zero BE-form $\varnothing^{3P}$ is the main option for all authors, but two South Russian authors use overt 3rd person perfect auxiliaries in non-existential and non-verificational contexts in the beginning of the XII century. Daniil and Monomakh have 21 unambiguous examples of PERFECT I with *есть*[1], *суть*[1], *есма*[1] in clause-internal positions. Three Novgorod authors from the second half of the XII century have only 9 examples with overt 3rd person auxiliaries: the contexts are ambiguous between PERFECT I and PERFECT II readings. In the body of "Kirik's Inquiries", overt 3rd person perfect auxiliaries are found only in the main part written by Kirik himself: his followers Savva and Elija do not use them in their supplements. XIV-XV century authors lack overt 3rd person perfect auxiliaries completely. The overall ratio of perfect clauses (k_{PERF}) in a text plays no role for the choice of overt 3rd person auxiliary. The figures are shown in tab. 1[5].

	DAN	MON	KIR_{KO}	KIR_{AU}	S & E_{KO}	S & E_{AU}	EL-JO	ANT	STEF	VAS	FLOR	ROME
Date	ca. 1104 – 1106	ca. 1097 – 1117	ca. 1132 – 1156				1166	ca. 1200	1347-1349	1347	1438-1439	1438
k_{PERF}	20. 3%	20.3%	13.2%	11.4%	17.4%	16.9 %	67.4 %	56.9 %	4.4%	22.6%	26.5%	71.4%
Dialect	Chernigov	Perejaslavl	Novgorod								Suzdal	
Overt 3rd person marker	18	3	5	3	0	0	2	2	0	0	0	0
$\varnothing^{3P}$	72	17	12	9	13	11	13	80	6	15	45	7
%	22.5%	15%	41.4 %	25%	0%	0%	13.3%	2.4%	0%	0%	0%	0%

Tab. 1. 3rd person perfect auxiliaries by XII century authors compared to XIV–XV century texts.

These figures suggest that overt 3rd person auxiliaries in PERFECT I were initially an optional feature of early South Russian dialects, i.e. an inherited feature by the G1 / OR-1 grammar of some XII century authors and not only a learned feature of G2/L2 grammar. In the beginning of the XII century Daniil still could select overt 3rd person auxiliaries in 22, 5% of PERFECT I clauses, but this variant of PERFECT I was rapidly declining.

4.2 PERFECT II as an L1 FEATURE

PERFECT II was a construction, where accented auxiliaries *есть*[2], *суть*[2], *есма*[2] did not alternate with $\varnothing^{3P}$. Daniil regularly used it in existential contexts and combined *есть*[2], *суть*[2] with the *l*-participle of the verb *быти* 'be' 31 times in sentences like (1): this peculiar combination of present and past BE-forms indicates a change of state and/or a canceled result.

[5] The abbreviations read: DAN = Daniil, MON = Monomakh, KIR = Kirik, S & I = Savva and Elija, EL-JO = Elija-John, ANT = Antonius, STEF = Stefan, VAS = Vasiliy Kalika, FLOR = "Pilgrimage to the Florence Council", ROME = "A Note about Rome", KO = "Novgorodskaja Kormchaja 1282 A.D.", AU = "Autonomous version of 'Kirik's Inquiries'".

(1) **И есть**^{2} 3SG **<u>был</u>** AUX.PAST.SG.M град NOM.SG.M Тивириада великъ велми (Daniil, LXXXII).
'And the city of T. **was once** very big <but **deteriorated** since that.>'

XII century Novgorod authors have 9 ambivalent sentences with overt 3rd person perfect auxiliaries in clause-internal positions: they should be rather interpreted as PERFECT II, than as PERFECT I. None of the Novgorod examples shows the combination *есть*^{2} + *былъ*, which is another relic preserved in the dialect of Daniil.

4.3 Model B and Clausal Subjects

The retrieved data does not fully support Zaliznjak's Model B. South Russian authors of the XII century (Daniil, Monomakh) do not have complementary distribution of overt 3rd person PERFECT I auxiliaries and grammatical subjects. A vast majority of PERFECT I clauses has the + SUBJ feature. In Daniil's text, a filled subject slot has no impact on the choice of the non-zero 3rd person marker: the expectancy is 20% both for + SUBJ and for - SUBJ clauses. The data from XII century Novgorod is scanty, but compatible with the prediction of Model B that the expectancy of overt 3rd person marker increases in – SUBJ perfect clauses.

	DANIIL		MONOMAKH		KIRIK$_{KO}$		KIRIK$_{AO}$		ELIJA-JOHN		ANTONIUS	
DATE	ca. 1104 – 1106		ca. 1097 – 1117		ca. 1132 – 1156		ca. 1132 – 1156		1166		ca. 1200	
DIALECT	Chernigov		Perejaslavl		Novgorod							
	+ SUBJ	-SUBJ	+ SUBJ	-SUBJ	+ SUBJ	-SUBJ	+ SUBJ	-SUBJ	+ SUBJ	-SUBJ	+ SUBJ	-SUBJ
OVERT 3rd PERSON MARKER	**15**	**3**	**2**	**1**	1	4	0	3	0	2	1	1
∅^{3P}	**60**	**12**	**15**	**2**	7	5	6	3	4	9	50	30
%	20%	20%	11.8 %	33.3%	12.5%	44.4%	0%	50%	0%	18.8%	2%	3.2%

Tab. 2. Grammatical subjects and overt 3rd person perfect auxiliaries in the XII century Russian

The contrast between South Russian and Novgorod authors of the XII century captured by tab. 2 can be explained by two factors. One of them is geographical. Different Old Russian dialects could have different settings of the parameter (ii) predicting the distribution of grammatical subjects in Old Russian idioms with overt 3rd person perfect auxiliaries. Another one is chronological. During the XII century some Russian dialects lost overt 3rd markers in PERFECT I. In both cases, Zaliznjak's Model B must be adjusted to specific data.

4.4 PERFECT I as an L1 FEATURE

Two relevant characteristics of G2/ OR-2 — the use of the simple past and overt 3rd person perfect auxiliaries — are displayed differently in the sample. Aorist and imperfect are chronologically stable and displayed uniformly in all dialects, since it is a learned L2 feature imported by all XII — XV centuries OR-2 texts. Meanwhile, PERFECT I with overt 3rd person perfect auxiliaries is a feature that was adapted only by those XII authors who had in their L1 grammar.

5 Conclusions

We discussed two linguistic models predicting language-internal variation in Old Russian syntax and checked them in a historical corpus consisting of several XII century author texts representing three dialects and seven different idioms. Linguistic models claiming a rigid distribution of language features or clustering of linguistic parameters activate many hidden text/corpus characteristics such as profiles of text genres, chronology and vector of change, ± impact of L2, ± presence of supra-dialect features,

etc. In this case they can be valued and checked on text samples, where the genre features are relatively stable, while location and time vary.

Acknowledgements

I am grateful to the anonymous reviewers for the valuable comments. The sole responsibility for the shortcomings is mine. This paper has been written with support from the project "Parametric description of the languages of Russian Federation" realized at Pushkin State Russian Language Institute.

References

Henry, Alison. Dialect variation, optionality and the learnability guarantee // Linguistica Atlantica. — 1998. — Vol. 20. — P. 51–71.

[1] Khaburgaev, Georgi A. The Fate of the Auxiliary Verb in Old Slavic Analytic Forms in Russian [Sud'ba vspomogatel'nogo glagola drevnih slavjanskih analiticheskih form v russkom jazyke] // The Bulletin of Moscow University. Series 9. Philology. — 1978. — N 2 [Vestnik Moscovskogo universiteta. Ser. 9. Filologija. —1978. — N 2].

[2] Zaliznjak, Andrej A. On the Language of Birch Bark Letters [K izucheniju jazyka berestjanyh gramot] /Janin, Valentin L., Zaliznjak, Andrej A. Novgorod Birch Bark Letters. From the Excavations 1984-1989 [Janin V.L., Zalizjak A.A. Novgorodskie gramoty na bereste. Iz raskopok 1984–1989 godov. Moscow: Nauka, 1993. — P. 191–319.

[3] Zaliznjak, Andrej A. Old Novgorod Dialect. 2nd ed. [Drevnenovgorodskij dialekt. 2-e izdanie. — Moscow: Jazyki slavjanskoj kul'tury, 2004.]

[4] Zaliznjak, Andrej A. Old Russian Enclitics [Drevnerusskie enklitiki]. — Moscow: Jazyki slavjanskoj kul'tury, 2008.

[5] Zimmerling, Anton V. (2018). Two Dialects of Russian Grammar: Corpus Grammars and Formal Models [Dva dialekta drevnerusskoj grammatiki: korpusnye grammatiki i model'] Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2018" [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii "Dialog 2018"], Moscow, pp.818–830.

[6] Zimmerling, Anton V. (2020). Zero forms in morphological paradigms: the verb "BE" in Russian // Computational linguistics and intellectual technologies, Issue 19 (26). Proceedings of the international conference "Dialogue 2020" [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii "Dialog 2020", Moscow, pp. 795 - 810].

[7] Zimmerling, Anton V. From Integral Frameworks to Aspective Description [Ot integral'nogo k aspektivnomu]. —Moscow-St.Petersburg: Nestor-Istoria, 2021.

References

[1] Henry, Alison. Dialect variation, optionality and the learnability guarantee // Linguistica Atlantica, 1998. — Vol. 20. — P. 51–71.

[2] Zimmerling, Anton V. Zero forms in morphological paradigms: the verb "BE" in Russian // Компьютерная лингвистики и интеллектуальные технологии. Вып. 19 (26). Труды международной конференции «Диалог 2020» [Computational linguistics and intellectual technologies, Issue 19 (26). Proceedings of the international conference "Dialogue 2020", pp. 795–810].

[3] Зализняк А. А. К изучению языка берестяных грамот /Янин В.Л., Зализняк А.А. Новгородские грамоты на бересте. Из раскопок 1984–1989 гг. — Москва: Наука, 1993. — Р. 191–319.

[4] Зализняк А. А. Древненовгородский диалект. 2-е изд. Москва: Языки славянской культуры, 2004.

[5] Зализняк А. А. Древнерусские энклитики. Москва: Языки славянской культуры, 2008.

[6] Хабургаев Г. А. Судьба вспомогательного глагола древних славянских аналитических форм в русском языке // Вестник Московского Университета. Сер. 9. Филология. —1978. — № 2.

[7] Циммерлинг А. В. Два диалекта русской грамматики: корпусные грамматики и модель // Компьютерная лингвистика и интеллектуальные технологии. Труды международной конференции «Диалог 2018», Москва, с. 818–830 [Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2018", Moscow, pp. 818–830].

[8] Циммерлинг А. В. От интегрального к аспективному. — Москва-Санкт-Петербург: Нестор-История, 2021.