

Verb rhyme in Russian poetry: a quantitative analysis

**Olga Lyashevskaya, Kristina Litvintseva,
Ekaterina Vlasova & Eugenia Sechina**

Russian Linguistics

International Journal for the Study of
Russian and other Slavic Languages

ISSN 0304-3487

Russ Linguist

DOI 10.1007/s11185-020-09227-w



Your article is protected by copyright and all rights are held exclusively by Springer Nature B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".



Verb rhyme in Russian poetry: a quantitative analysis

Глагольная рифма в русской поэзии: количественный анализ

Olga Lyashevskaya^{1,2} · Kristina Litvintseva¹ ·
Ekaterina Vlasova¹ · Eugenia Sechina¹

© Springer Nature B.V. 2020

Abstract This article provides a quantitative corpus-based investigation of the Russian verb rhyme and its change in the Russian poetic tradition from the beginning of the 19th century to the 1960s. Versologists have studied the rhyme primarily as a phonetic entity, whereas morphology also contributes to the rhyme euphony due to the regularity of grammatical affixes. The research focuses on a micro-diachronic analysis of verb rhymes, summarises the identified historical trends, and defines acceptable and clearly avoided verb forms. The article also analyses the morphological patterns of verb rhymes including the most common lexical pairings and combinations of particular grammatical forms with different parts of speech. The study analyses data from the Corpus of Russian Poetry (a part of the Russian National Corpus) and introduces research methods and a corpus-based tool that were designed specifically for the statistical analysis and computational modelling of poetic features. The results show that authors experimented with word rhyme in various ways during different periods. Despite the idea of non-aesthetic verbal rhyme, which has existed since the time of A. Kantemir, its use in the historical perspective varies, there are periods of rise and fall. We distinguish two classes of rhyming pairs: combinations of two verb forms and morphologically dissimilar combinations of a verb form with a word of another part of speech. We conclude that restrictions on verbal rhyme apply mainly to combinations of past tense and infinitive forms.

The research was prepared within the framework of the Academic Fund Program at the National Research University Higher School of Economics (HSE) in 2018–2019 (grant #18-05-0047) and by the Russian Academic Excellence Project «5-100».

✉ O. Lyashevskaya
olesar@yandex.ru

K. Litvintseva
tinalitvina@gmail.com

E. Vlasova
evlasova@hse.ru

E. Sechina
grunwaldus@gmail.com

¹ National Research University Higher School of Economics, Moscow, Russia

² V.V. Vinogradov Russian Language Institute of the Russian Academy of Sciences, Moscow, Russia

Nevertheless, the rhyming of indicative forms and morphologically heterogeneous rhymes is used more freely.

Аннотация Статья представляет корпусное исследование русской глагольной рифмы и связанных с ней диахронических тенденций в русской поэтической традиции с 1800-х до 1960-х гг. До сих пор рифма исследовалась преимущественно с позиции фонетики, тогда как грамматические факторы учитывались недостаточно. Между тем морфология языка также влияет на эвфонию рифмы из-за регулярности аффиксов. Наше исследование проведено на материале Поэтического корпуса Национального корпуса русского языка. Для количественного анализа была создана база данных, содержащая дополненную разметку стиховедческих и лингвистических признаков, в том числе разметку рифмованных пар и цепочек, уточненные леммы и грамматические признаки. Результаты показывают, что авторы по-разному экспериментировали с глагольной рифмой в разные периоды. Несмотря на представление о неэстетичности глагольной рифмы, бытующее еще со времен А. Кантемира, ее употребительность в исторической перспективе носит волнообразный характер, выделяются периоды подъема и спада. Мы выделяем два класса рифмованных пар: сочетания двух глагольных форм и морфологически гетерогенные сочетания глагольной словоформы со словом другой части речи. Мы приходим к выводу, что внутри глагольной рифмы обнаруживаются ограничения в основном на сочетания форм прошедшего времени и инфинитива. Вместе с тем, рифмование индикативных форм и морфологически гетерогенные рифмы используются более свободно.

1 The dispute about Russian rhyme: between phonetics and grammar

Rhyme has been an essential aesthetic element in Russian versification from the 18th century to the present day. In Russian, rhyming counterparts demand at least an adjacent stressed vowel and consonant, e.g. *vodój / gustój* ‘water / dense’, to create a euphony in the end of the verse line. Rhyming elements do not necessarily contain identical letter combinations, since the Russian spelling system has ten letters for five basic vowel sounds: *a / ja* [a], *u / ju* [u], *o / ě* [o], *y / i* [i], *é / e* [e]. Letters from the same ‘vowel pair’ still constitute an identical sound combination, see the example *dóm – ogném* ‘home – fire’ (rhyming element [om]) from Anna Akhmatova’s poem *Pesnja poslednej vstreči* written in 1911:

- (1) *Éto pesnja poslednej vstreči.*
Ja vzgljanula na temnyj dom.
Tol’ko v spal’ne goreli sveči
Ravnodušno-želtym ogném.
 ‘This is the song of the final meeting.
 I looked up at your house, — all dark inside.
 Just the bedroom candles burned with a fleeting,
 Indifferent and yellowish light.’ (Translation by Andrey Kneller)¹

The paired consonant letters *d / t*, *z / s*, *b / p*, *v / f*, *g / k*, *ž / š* also phonetically coincide with each other, as long as they are both devoiced, for example *šutjat – sad* ‘noise – garden’ (rhyming element [at]). Despite the alternations of the vowel and consonants letters, these rhyming elements introduce an exact rhyme, since they are phonetically identical.

¹Final Meeting: Selected Poetry of Anna Akhmatova. Translated by Andrey Kneller. Boston 2011.

Exactitude of rhyme, however, imposes severe constraints on word choice and poetic language in general, since the number of possible rhyming words in language is finite, and frequently used rhyming pairs become a cliché and lose their aesthetic power over the course of time. Russian grammar eventually offers a solution to this poetic challenge, as many grammatical forms have the same endings, which are natural exact rhymes. This feature of Russian grammar allows poets to create rhymes using a wide range of lexemes with the same or similar endings. On the onset of Russian versification, poets regularly employed grammatical rhyming elements, see the following examples from Lomonosov's translation of the first psalm, which scholars view as a typical rhyming pattern of his time (Wachtel 2004, pp. 28–29): *xódit – privódit* 'walk – lead', *stupát' – zasedát'* 'step – sit', *pospéšny – gréšny* 'rash – sinful'. Due to grammatical affixes, these rhymes coincide completely in all letters that follow the rhyming syllable: verbs rhyme with verbs and nouns rhyme with nouns.

A simple matching of morphological endings does not require great ingenuity, and this transparent, yet predictable rhyming technique resulted in a long-term discussion among Russian poets about whether rhymes of the same grammatical form are appropriate and aesthetically valuable or not. It is notable that poets mostly rejected verb forms and did not put severe limits on other parts of speech. The dispute was started in the 18th century by Antiochus Kantemir, the 'father of Russian poetry'. He condemned the infinitive forms ending in *-ati* with each other as being 'vile', but allowed them to rhyme with other parts of speech, for example, *mati* 'mother' – *spati* 'to sleep' (Gasparov 2000, p. 53). Later on, poets who used verb rhymes were accused of having poorly mastered the art of poetry (Samojlov 2005, p. 341), and there were authors, which were known to be consistently avoiding this type of rhyming (for example, in Vladimir Majakovskij, the verb rhymes are found in only 1% of poems with female endings and 2% with male endings (Gasparov 2000, p. 321)). However, in the 20th century, some authors intentionally played with homonymous and tautological rhymes, and, among them, with the verb rhymes.

This dispute highlighted one prominent quality of rhyme in Russian: Phonetic rules are not the only factor that determines euphony at the end of a verse line, rhyming practices and their development also rely to a wide extent on the grammatical features of word-forms. Although Russian versologists have previously described the boundaries of Russian rhyme patterns (Žirmunskij 1923; Tomaševskij 1959; Gasparov and Skulačeva 2004), we still do not know, how rhyming practices have changed over time and to what extent grammatical factors have contributed to the development of rhyming patterns.

This article provides a quantitative corpus-based investigation of the Russian verb rhyme and how their use in the Russian poetic tradition from the beginning of the 19th century to the 1960s has changed. Section 2 introduces research methods and the corpus-based tool that our research group has designed specifically for the statistical analysis and computational modelling of poetic features. Section 3 contains a micro-diachronic analysis of verb rhymes supported by quantitative data. In Sect. 4, we summarise the identified historical trends and define acceptable and clearly avoided verb forms.

2 A quantitative corpus-based approach to Russian rhyme research: materials and methods

Russian versology has relied heavily on quantitative data for making predictions and generalizations about meter, rhyme, and other formal and linguistic features of poetic language (see Gasparov 2005; Taranovskij 2010; Jakobson 1973; Jarxo 2006, to name only a few; see also overviews in Semenov 2009; Kizhner et al. 2018). Slavic corpus linguistics (Kopotev,

Lyashevskaya, and Mustajoki 2018; Divjak, Sharoff, and Erjavec 2017) and formal methods in poetry in general (Scherr, Bailey, and Kazartsev 2011) add a new dimension to quantitative studies of poetry, providing more in depth annotations of linguistic and literary layers, and allowing larger data to be processed. At the moment, the Russian Poetry Corpus, which is an integral part of the Russian National Corpus, is the largest open-access online-collection of Russian poetic texts written between the 18th and the 20th-century. As Grišina et al. (2009, pp. 71–113) note, the Russian Poetry Corpus significantly reduces the amount of manual work involved in extracting basic quantitative data relevant for the versology studies, see more about it in (Orexov 2015).

Meanwhile, modern computational literary studies require a digital tool to maximize corpus output and provide pre-processed frequencies of lexemes, syntactic units, and other grammatical features. For this purpose, our interdisciplinary research group from the Higher School of Economics (Moscow) designed a new digital tool and an open-access web application for data summarising, filtering, and pattern structuring. We compiled the materials from the Russian Poetry Corpus and assembled the data with more accurate, in depth, and elaborated metatextual, versological and linguistic annotations, which have been validated by experts in modern literary studies and NLP (Natural language processing).

Every line contains tags indicating meter and foot, a number of words and syllables, ictuses and information on rhyming segments. We also marked up whether the sentence boundaries are placed at the beginning or at the end of the string.

Rhyme annotation, which is of particular interest in this article, comprises rhymed chains, an order of rhyming lines, the number of rhyming elements in a chain (usually two, less often three, four, or longer chain rhymes, e.g. monorim). Since the Russian National Corpus does not provide information on sublexical rhyming units, we retrieved this data automatically, based on information about elements of rhyming units and rhyming schemas.² We also marked up word-long rhymes and larger rhyming units (supralexical units), stress and ictus patterns, patterns of vowels and syllable structure, parts of speech and other grammatical features of rhyming elements.

Finally, elaborated annotation comprises char-grams and combinations of letters, their classification, combinations of vowels and consonants, as well as stress patterns. This information is determined using orthography and graphic elements.

The database includes poetic texts dating from 1800 to the present. The further analysis comprises about 80,000 poetic texts of both short and long genres, which together add up to more than 2 million verses and about 10 million words. Poetic metatexts, such as headings, dates, epigraphs, prosaic comments made by the author, and editors' notes belong to a separate part of the database and are subject to a separate investigation, see for example Kuzmenko and Orekhov (2016).

Figures 1a and 1b demonstrate how the size of the corpus varies if we group the texts by decades (Fig. 1a) or by authors (Fig. 1b). A user can gather information about the proportions of subcorpora and normalize the raw frequencies with regard to the size of the specified subcorpus. A user can also determine the period of poetic tradition to be included in their research. For example, the size of the subcorpus comprising the period after 1980 is rather small and cannot provide reliable statistical data in many cases. Furthermore, for the contributors and maintainers of the corpus, these graphs would also suggest which parts of the corpus need balancing and the addition of new data.

²At the moment, we have processed 60% of the verses, taking only data into account in which the rhyming schema does not change throughout the whole text.

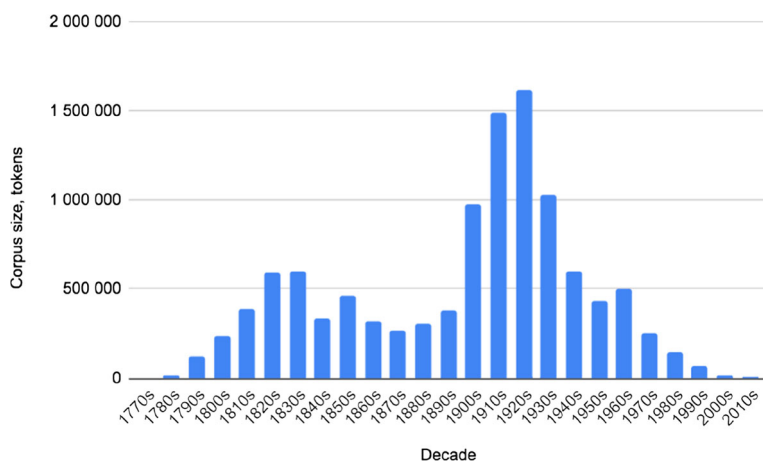


Fig. 1a Corpus size by decade

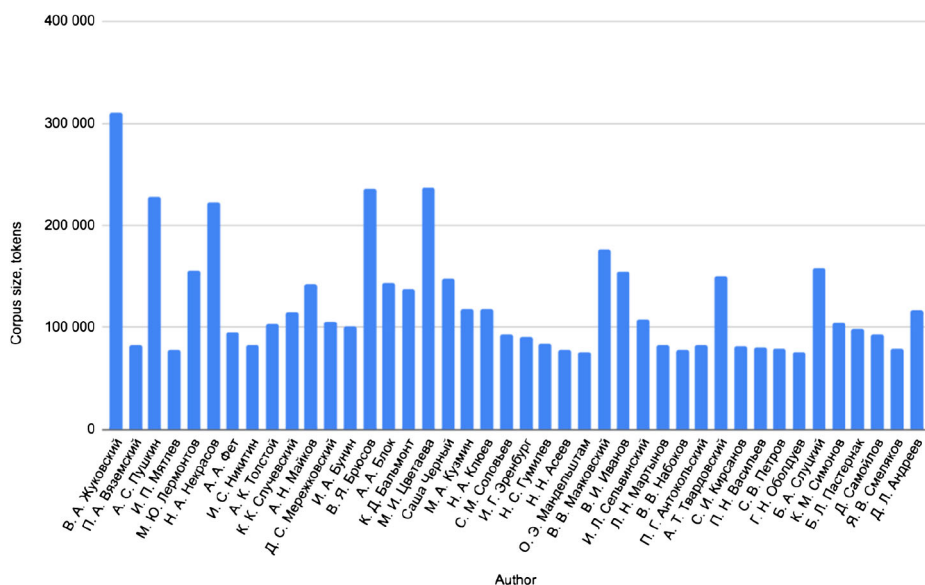


Fig. 1b Varying corpus sizes by selected authors

3 A quantitative diachronic analysis of verb rhymes

By using quantitative corpus data, we can study whether the authors have followed the above-mentioned restrictions on grammatical verb rhymes and if the trend has changed over time. We examine two alternative hypotheses:

- (i) Authors seek to avoid the verb-to-verb rhymes at the beginning of the 1800s, but this is less strictly observed during the later period;
- (ii) There are periods in which the authors follow the recommendations, but they alternate with periods in which these recommendations are less strictly observed.

In both cases, we need to identify in which period(s) the rule is violated most and under which conditions. In order to put this analysis in a broader perspective, we retrieved data on the use of verb rhymes, both in the pairs ‘verb form – verb form’ (V-V, cf. (2) from V. I. Ivanov’s poem *Ozim*’ written in 1904), and ‘verb form – non-verb form’ (V-non-V, cf. (3) from F. I. Tjutčev’s poem «*Ėti bednye selen’ja ...*» written 1855) (in any order):

- (2) Nezadrožavšij trepet *lovit*
 Mež kosnyx glyb, — tak Rus’ moja
 Nemotnoj smerti *prekoslovit*
 Gluxim začat’em bytija ...
 ‘Detects the trembling not yet begun
 Among the stasis of the clods of earth,—so my Russia
 Defies mute death
 Through the imperceptible inception of life ...’ (transl. Wachtel 2004, p. 137)
- (3) Udručennyj nošej krestnoj,
 Vsju tebjja, zemlja *rodnaja*,
 V rabskom vide Car’ *nebesnyj*
 Isxodil, *blagoslovljaja*.
 ‘Weighed down by the burden of the cross,
 In the guise of a slave, the heavenly king
 Walked through all of you, my native land,
 Giving blessing.’ (transl. Wachtel 2004, p. 135)

Rhymes consisting of more than one word in any rhyming unit (e.g. verbs followed by particles in pairs like *dotjanu li – Kaligule* ‘if (I) reach – Caligula’) are excluded; this also excludes pairs with the subjunctive forms with the particle *by*, *b* ‘would’. In order to simplify the calculations, rhyming chains with more than two rhyming lines were decomposed into simple pairs in which the following lines always rhyme with the first line of the chain. The data is limited by the time of creation from 1801 to 1960, the timeline is divided into steps of 20 year intervals. Texts with a creation timespan beyond the step boundaries, have also been excluded from examination. Since not all rhyming pairs have been annotated in the current database (see fn. 2), the normalized frequencies are calculated³ taking into account the size of subcorpora that only include texts annotated with regard to the rhyming pairs. Lastly, we set a threshold of three or more occurrences of a particular rhyming pair in the corpus, and two or more authors using the same pair, to exclude the author’s individual choice (cf. the rhyme *obmanut – ustanut* ‘deceived – get tired’ used only by Brjusov) and possible errors of automatic rhyme identification. The resulting dataset consists of 39,319 rhyming pairs, of which 9,172 are of the V-V type and 30,147 are of the V-non-V type.

3.1 General distribution of verb rhymes across time

Figure 2 shows a relatively complicated structure of verb rhyme distribution over the period 1801–1960, with peaks at 1841–1860, 1881–1920, and 1941–1960. Should we interpret this as a piece of evidence against hypothesis (i)? Not necessarily, there is a general decrease in the normalized frequencies of verb-to-verb rhymes. Three periods can be distinguished:

³In ipm (items per million words). Note that yet another natural way to perform normalization for the poetry data is to weight the occurrences of units per the number of verses (lines) rather than words or tokens. However, the rhymed chains can include three, four, even more than 10 rhyming lines, so weighting this data method per line would be difficult.

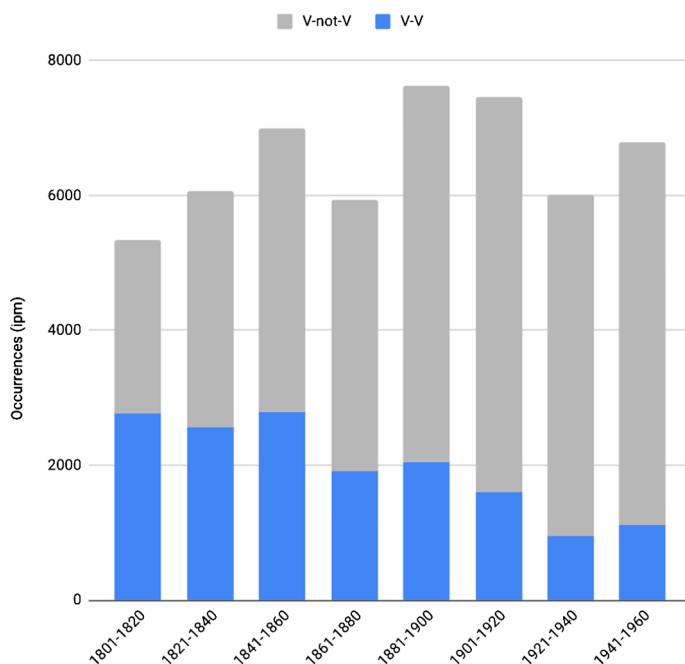


Fig. 2 Distribution of verb rhymes between 1801 and 1960: all verb forms, verbs rhyming with verbs (V-V) and verbs rhyming with other parts of speech (V-non-V)

1801–1860—higher, 1861–1920—middle, and 1921–1960—lower proportions of the V-V rhyming pairs. At the same time, the number of the V-non-V pairs increases consistently, and in this case, three different stages appear to exist: 1801–1820—lower, 1821–1880—middle, and 1881–1960—higher proportions of V-non-V rhyming pairs.

In this instance, the authors active between 1841 and 1860 are rather conservative in terms of their use of the V-V rhymes compared to the previous periods. This trend is in line with an increased interest in the use of the folk and vernacular motifs during this period, for which the simpler the rhyme, the more stylised it was. The V-V rhyme is most actively used by I. S. Nikitin (150 pairs, 5,009 ipm⁴), I. P. Mjatlev (22 pairs, 4,360 ipm), A. A. Fet (116 pairs, 3,451 ipm), L. A. Mej (47 pairs, 3,123 ipm), N. F. Ščerbina (72 pairs, 2,873 ipm), A. A. Grigor'ev (32 pairs, 2,611 ipm). Besides that, authors in the 1841–1860 subset followed the trend of the previous time periods: creating rhymes by combining, more actively, verbs with non-verb elements. In addition to those mentioned, the most noticeable authors, in this respect, include A. N. Apuxtin (97 pairs, 6,469 ipm), N. A. Nekrasov (98 pairs, 4,052 ipm), A. N. Pleščeev (87 pairs, 5,896 ipm), and V. G. Benediktov (56 pairs, 4794 ipm).

At the beginning of the Soviet era, in 1921–1940, there is a sharp decrease in the use of V-V pairs. Still, we can identify authors that use them comparatively frequently: V. V. Nabokov (48 pairs, 1,818 ipm), B. J. Poplavskij (48 pairs, 1,631 ipm), and A. T. Tvardovskij (34 pairs, 1,401 ipm). We are also able to examine whether this pattern correlates with the active use of the V-non-V combinations within this period and for the same authors (Nabokov and Tvardovskij are among the top-5 in this respect).

⁴The absolute frequencies are weighted by the size of the corpus of a given author in a given time period.

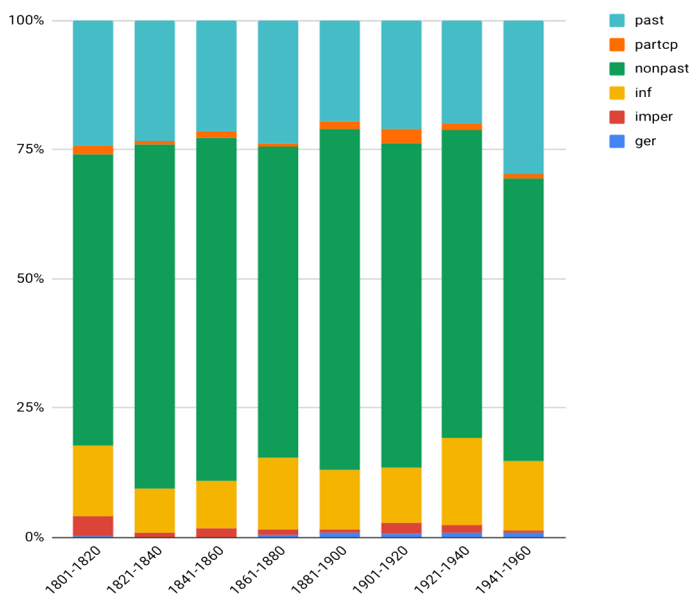


Fig. 3 Distribution of V-V rhymes between 1801 and 1960 by grammatical forms: ger—gerundive, imper—imperative, inf—infinitive, nonpast—non-past finite, partcp—participle, past—past finite

The Russian verb paradigm contains a wide diversity of verb forms with different rhyming potential and aesthetic value. Figure 3 illustrates the contribution of various grammatical forms to the V-V distributional pattern presented in Fig. 2.

Interestingly, the infinitive forms are underused in 1821–1860 and 1901–1920, i.e. in the periods before there was a sharp drop in use of V-V rhymes⁵ (1861–1880 and 1921–1940). When the drop happens, it is followed by a decrease in the ratio of the non-past indicative forms. The contribution of the past indicative forms is the largest in the period from 1941–1960, associated with the war and post-war poetic narrative.

To sum up, the micro-chronological analysis identifies the stable historical trend that the pure verb-verb rhymes have gradually been replaced with V-non-V rhymes over time.

3.2 Lexical pairings

In the remaining part of the section we examine word lists and focus more on the V-non-V rhymes. The corpus' data analysis tool provides the user with an option to retrieve the frequency word lists for the category in question, in our case, for rhyming counterparts of the particular grammatical forms. The list of rhymes with an infinitive includes the pairs *opjat' – spat'* 'again – to sleep' (36 occurrences), *opjat' – ponjat'* 'again – to comprehend' (25 occurrences). The following rhymes top the list for non-past forms: *poju – svoju* '(I) sing – oneself' (35 occurrences) and *pojut – prijut* '(they) sing – a refuge' (34 occurrences). Among rhymes with the gerund the most frequent forms are *ljubja – tebja* 'loving – you' (267 occurrences) and *ljubja – sebja* 'loving – myself' (117 occurrences). Among rhymes with the

⁵Note that whereas A. Kantemir called the feminine infinitive rhymes like *spati* 'to sleep' as 'vile', by the 19th century the *-ti(s')* forms were replaced by forms with the endings *-t'(sja)*, *-č'(sja)*, consistent with the masculine rhyme.

Table 1 Top frequent rhyming verb pairs: V-V (right) and V-non-V (left)

F	V1	V2	F	V	non-V
87	<i>dýšit</i> 'breathe'	<i>slýšit</i> 'hear'	271	<i>ljubjá</i> 'love'	<i>tebjá</i> 'you'
66	<i>búdet</i> 'be'	<i>zabúdet</i> 'forget'	187	<i>dýšá</i> 'breathe'	<i>dušá</i> 'soul'
62	<i>govorít</i> 'say'	<i>gorít</i> 'burn'	169	<i>pomóč'</i> 'help'	<i>noč'</i> 'night'
60	<i>byt'</i> 'be'	<i>ljubit'</i> 'love'	147	<i>idítí</i> 'go'	<i>putí</i> 'way, path'
56	<i>móžet</i> 'can'	<i>trevóžít</i> 'be disturbing'	137	<i>najítí</i> 'find'	<i>putí</i> 'way, path'
49	<i>žit'</i> 'live'	<i>ljubít</i> 'love'	124	<i>spešá</i> 'be in a hurry'	<i>dušá</i> 'soul'
49	<i>bléščet</i> 'blister'	<i>trepéščet</i> 'flutter'	117	<i>ljubjá</i> 'love'	<i>sebjá</i> 'oneself'
47	<i>byt'</i> 'be'	<i>zabýť</i> 'forget'	110	<i>zvenjá</i> 'ring'	<i>menjá</i> 'me'
46	<i>byl</i> 'be'	<i>ljubil'</i> 'love'	104	<i>est'</i> 'be, eat'	<i>čest'</i> 'honour'
44	<i>byt'</i> 'be'	<i>žit'</i> 'live'	90	<i>zoví</i> 'call'	<i>ljubví</i> 'love'
43	<i>dýšit</i> 'breathe'	<i>kolýšet</i> 'flutter'	85	<i>otdoxnút'</i> 'relax'	<i>put'</i> 'way, path'
39	<i>lovljú</i> 'catch'	<i>ljubljú</i> 'love'	83	<i>šutjá</i> 'joke'	<i>ditjá</i> 'child'
30	<i>búdu</i> 'be'	<i>zabúdu</i> 'forget'	82	<i>mog</i> 'can'	<i>bog</i> 'God'
30	<i>byl</i> 'be'	<i>zabýl</i> 'forget'	78	<i>moglí</i> 'can'	<i>zemlí</i> 'earth, land, soil'

past indicative forms the most frequent forms are *byl* – *ljubil* 'was – loved' (45 occurrences) and *ljubil* – *sil* 'loved' – 'force' (38 occurrences). Among rhymes with the imperative the most frequent forms are *prosti* – *puti* 'forgive – paths' (75 occurrences) and *živi* – *ljubvi* 'live – love' (61 occurrences). The most frequent rhymes with a participle include *davno* – *dano* 'long ago – given' (35 occurrences) and *dano* – *odno* 'given – the one' (30 occurrences). Moreover, only the participle *dano* 'given' is attested in combination with these types of rhymes.

It can be seen that verbs which are most frequent in such word lists are those with a 'poetic' meaning related to topics like life and love (cf. *ponjat* 'understand', *pet* 'sing', *ljubit* 'love', *prostit* 'forgive', *žit* 'live'). Besides that, the verb *byt'* 'be' is frequently used, both as a common verb and as an auxiliary. The list of the most frequently used V-V pairs includes a lot of similar verbs (see Table 1). Verbs such as *byt'* 'to be', *ljubit'* 'to love', *pet'* 'to sing' are the most frequent verbs in both types of rhymes: V-V and V-non-V.

Tables 2 and 3 illustrate the distribution in pairs in which one element is represented by a particular grammatical form of the verb and another by a particular part of speech. All in all, verbs rhyme most often with nouns (23,594 rhymes (60%), of which 3,624 are unique), followed by verbs (9,172 rhymes (23%); 1,639 unique), adverbials (2,442 rhymes (6%); 370 unique), and nominal pronouns (1,704 rhymes (4%); 164 unique). The remaining part-of-speech combinations make up 6% of the data. Thus, the verb-to-verb rhymes are the second most used and productive forms.

Both Tables 2 and 3 show the same clear trend: the most common morphological rhyming pattern in Russian poetry is made up of a combination of a verb form and a noun. As for verb-verb rhymes, there was a clear constraint on verb pairings, including the combinations of two past forms or two infinitives. Meanwhile, rhyming of non-past verb forms with other verb forms is as frequent as the heteromorphological rhyming using verbs with substantives.

Infinitives often make up a rhyme with nouns (3,629 occurrences; 465 unique pairs), verbs (1,163 occurrences; 208 unique pairs), and adverbs (347 occurrences; 52 unique pairs). The non-past forms demonstrate a similar distribution of pairs: the most common combinations are those with nouns (6,818 occurrences, of which 1,100 are unique), verbs (6,186 occurrences; 1,054 unique), adverbs (1,003 occurrences; 135 unique), and adjectival pronouns

Table 2 Rhyming pairs: verb forms and part of speech matches (number of occurrences)^a

Form	A	ADV*	APRO	CONJ	INTJ	NUM	PART	PR	S	SPRO	V
ger	297	105	108			4		6	2357	1108	51
imper	64	214	63			25	6		2729	54	101
inf	4	347	1		1	40	1		3629	2	1163
nonpast	224	1003	723	19	23	20	76		6818	158	6186
partcp	194	154	55	7			12	2	1043	273	119
past	364	619	48	10		8			7018	109	1552
Total	1147	2442	998	36	24	97	95	8	23594	1704	9172

^aHere, and below, the part of speech tags from the Russian National Corpus were used: A—adjective, ADV—adverb, APRO—adjectival pronoun, CONJ—conjunctive, INTJ—interjection, NUM—numeral, PART—particle, PR—preposition, S—noun, SPRO—substantive pronoun, V—verb. The category of ADV* includes adverbs, predicatives, and parentheticals (cf. Ljaševskaja and Šarov 2009).

Table 3 Rhyming pairs: verb forms and part of speech matches (number of unique pairs)

Form	A	ADV*	APRO	CONJ	INTJ	NUM	PART	PR	S	SPRO	V
ger	66	11	15			1		1	333	55	9
imper	15	35	15			3	2		361	11	20
inf	1	52	1			8			465		208
nonpast	47	135	83	3	5	3	12		1100	30	1054
partcp	47	25	7	1			2		210	55	28
past	66	112	8	3		1			1155	13	320
Total	242	370	129	7	5	16	16	1	3624	164	1639

(723 occurrences; 83 unique). The past tense forms most often rhyme with nouns (7,010 occurrences; 1,155 unique), verbs (1,152 occurrences; 320 unique), adverbials (619 occurrences; 112 unique). However, there is a larger proportion of past forms rhyming with adjectives (364 occurrences, 4% of all past forms). It is likely that the rhyming potential of adjectives is connected with the combination of the short *l*-adjectives and *l*-forms of the verb (52 of 66 unique forms: *bylo* – *milo* ‘was – sweet’, *byla* – *mila* ‘was – sweet’, *byla* – *svetla* ‘was – bright’, *zametil* – *svetel* ‘noticed – light’). The same applies to the combinations of the short adjectives and short participles. Besides that, the infinitives rhyme with numerals, mostly with *šest* ‘6’, *pjat* ‘5’, and *desjat* ‘10’, cf. *est*’ – *šest* ‘eat – six’ and *sest*’ – *šest* ‘sit down – six’, *spat*’ – *pjat* ‘sleep – five’, *povesit*’ – *desjat* ‘hang – 10’). The non-past forms also rhyme with conjunctions and particles.

The analysis revealed a relatively uniform hierarchy of the rhyming groups, with nouns, verbs, and adverbs being among the most frequent non-verb elements of pairs. In contrast, gerunds stand out among the rhyming verb forms since they frequently rhyme with nominal pronouns (*maja* – *moja* ‘concealing’ – my’, *visja* – *vsja* ‘hanging – all’) and adjectives, but rarely combine with verbs. In general, the observed distribution indicates a high activity of verbs in the rhyming zone and, accordingly, there are no particular limitations on verbs combining with other parts of speech. Obviously, some frequent and stable cliché, such as *idti* – *puti* ‘walking – paths’ (147 occurrences) and *zovi* – *ljubvi* ‘call – love’ (92 occurrences) may affect (both positively and negatively) the use of particular grammatical forms and lexemes.

4 Conclusions

The chronological corpus-based analysis has demonstrated that Russian verb rhymes fall into two patterns, and each pattern has developed independently over time. The general grammatical rhyme pattern is made up of a verb paired with a noun, and eventually agrees with Kantemir's ideal. However, another pattern involving verb-to-verb rhymes is still in common practice; poets have not completely rejected this pattern, which is significantly more frequent than pairing verbs with other parts of speech, nouns notwithstanding.

Within the group of verb-verb rhymes, the most frequent rhyming forms are indicative forms, past forms, and infinitive forms. The indicative forms rhyme with other verbs as often as with nouns, whereas past forms and infinitive pairings dominate in rhymes containing different parts of speech. Thus, the implied poetic restriction on the use of certain verb forms in Russian rhymes has not been supported by the corpus data. Although Kantemir, the author of famous guidelines to Russian versology, recommended not to use infinitives in a rhyme zone at all, the data shows that rhymes that include infinitives are the third most frequent, after the non-past and past forms.

The general historical trend shows that grammatical patterns of verb rhymes gradually changed over time towards an increase of rhymes containing different parts of speech. The fall in the frequency of use of verb rhymes, at the beginning of the 20th century, could be associated with language experiments made by avant-garde poets rather than the influence of certain 'rules' of poetry.

After all, the results only partially confirm hypothesis (i) realting to the rejection of verbs within rhymes: The authors avoid verb-to-verb rhymes more in the later periods than at the beginning of the 1800s, however, they experiment with V-non-V rhymes more actively in later periods of time. Hypothesis (ii) holds true with respect to verb rhymes in general: there are periods during which the use of the verb rhymes is more limited, which alternate with periods during which there was an increase in their use. The detailed corpus statistics allow the user to identify authors who use such rhymes more frequently compared to others, and to examine grammatical forms, parts of speech and words which are used more actively in the creation of verb rhymes.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Divjak, D. S., Sharoff, S., & Erjavec, T. (2017). Slavic corpus and computational linguistics. *Journal of Slavic Linguistics*, 25(2), 171–198. <https://doi.org/10.1353/jsl.2017.0008>.
- Gasparov, M. L. (2000). *Očerki istorii russkogo stixa*. Moskva.
- Gasparov, M. L. (2005). *Sovremennyy russkij stix: Metrika i ritmika*. Moskva.
- Gasparov, M. L., & Skulačeva, T. V. (2004). *Stat'i o lingvistike stixa*. Moskva.
- Grišina, E. A., Korčagin, K. M., Plungjan, V. A., & Sičinava, D. V. (2009). Poetičeskij korpus v ramkax Nacional'nogo korpusa russkogo jazyka: obščaja struktura i perspektivy ispol'zovanija. In V. A. Plungjan et al. (Eds.), *Nacional'nyj korpus russkogo jazyka: 2006–2008. Novye rezul'taty i perspektivy* (pp. 71–113). Sankt Peterburg.
- Jakobson, R. (1973). *Questions de poétique*. Paris.
- Jarxo, B. I. (2006). *Metodologija točnogo literaturovedenija. Izbrannye trudy po teorii literatury*. Moskva.
- Kizhner, I., Terras, M., Manovich, L., Orekhov, B., Bonch-Osmolovskaya, A., & Rumyantsev, M. (2018). The history and context of the digital humanities in Russia. In J. G. Palau & I. G. Russel (Eds.), *Digital humanities 2018. Puentes-Bridges. Book of abstracts* (pp. 416–419). Mexico City. Available online at <https://dh2018.adho.org/abstracts/> (30 June 2020).

- Kopotev, M., Lyashevskaya, O., & Mustajoki, A. (2018). *Quantitative approaches to the Russian Language*. London, New York.
- Kuzmenko, E., & Orekhov, B. (2016). Geography of Russian poetry: Countries and cities inside the poetic world. *Digital Humanities 2016: Conference Abstracts*. Jagiellonian University & Pedagogical University, Kraków (pp. 830–832). Kraków.
- Ljaševskaja, O., & Šarov, S. (2009). *Častotnyj slovar' sovremennogo russkogo jazyka (na materialax Nacional'nogo korpusa russkogo jazyka)*. Moskva.
- Orexov, B. V. (2015). Ešče raz ob issledovatel'skom potenciale poetičeskogo korpusa: metr, leksika, formula. *Trudy instituta russkogo jazyka imeni V. V. Vinogradova*, 6, 449–463.
- Samojlov, D. (2005). *Kniga o russkoj rifme*. Moskva.
- Scherr, B. P., Bailey, J., & Kazartsev, E. V. (Eds.) (2011). *Formal methods in poetics. A collection of scholarly works dedicated to the memory of Professor M. A. Krasnoperova*. Lüdenscheid.
- Semenov, V. B. (2009). Statističeskie metody v russkom stixovedenii XX veka: Andrej Belyj i A. N. Komogorov. *Vestnik Moskovskogo universiteta. Serija 9: Filologija*, No. 6, pp. 36–44.
- Taranovskij, K. (2010). *Russkie dvuxsložnye razmery. Stat'i o stixe*. Moskva.
- Tomaševskij, B. V. (1959). *Stix i jazyk*. Moskva, Leningrad.
- Wachtel, M. (2004). *The Cambridge introduction to Russian poetry*. Cambridge.
- Žirmunskij, V. (1923). *Rifma, ee istorija i teorija*. Petrograd.