The shortest species: how the length of Russian poetry changed (1750–1921)

Artjom Shelya, Oleg Sobchuk*

Abstract: The paper studies long-term changes in the length of Russian poetry (1750–1921) to reveal the relation of poem length (counted in lines) to a poetic form and its evolution. The research has shown a dramatic decrease in the mean and median poetry lengths during the 19th century. This decrease was followed by the decline in length diversity, which resulted in short poems (8–20 lines) overpopulating the literature during the age of Modernism. We argue that this transformation towards the short form could be understood in the framework of cultural evolution: Russian poetry struggled to keep its literary niche, while being continuously under the pressure of successful large narratives of the 19th century. Therefore, it was forced to develop complexity while being highly constrained formally (accentual-syllabic verse and rhyme maintained for a long time) by the shrunk length of a lyrical poem.

Keywords: Russian poetry; length; 19th century; 20th century; cultural evolution; complexity; system of poetic genres

Introduction

The length of poetry is not a widely discussed topic in literary theory – neither formalist nor structuralist. However, changes in poem lengths are frequently addressed in the context of genre's system disintegration, which was happening in Russian poetry during the first half of the 19th century (see on that: Ginzburg 1974, Tynyanov 1977): since Russian formalists this disintegration has been seen as the decline of monumental lyrical genres (ode), followed by the dominance of elegy in the 1810s, and its decomposition towards genre-unspecific lyrical form. The genre system and its normative examples started to lose their influence on the subsequent generations of Romantic poems.

There are reasons to think that during these processes Russian poetry was becoming shorter: obviously, the ode was a large genre that vanished while

* Authors’ addresses: Artjom Shelya, Department of Slavic Studies, University of Tartu, Lossi 3, 51003, Tartu, Estonia. E-mail: artjoms@ut.ee; Oleg Sobchuk, Institute for Cultural Research and Arts, University of Tartu, Ülikooli 16, 51003, Tartu, Estonia. E-mail: sobchuk@ut.ee.

doi: doi.org/10.12697/smp.2017.4.1.03
literature drifted away from the court, but multiple evidence also suggests that the long elegy of the 1810s shrunk rapidly during the 1820s (Martynenko 2015: 25–27); Fyodor Tyutchev broke down the philosophical ode and invented an extra-short form of a “lyrical fragment” (Leibov 2000: 44–58; Rydel 1995; Tynyanov 1977b); Lermontov – in the situation of genre ambiguity – pushed narrative forms towards “lyrical novellas” (Eikhenbaum 1924: 115–128); ballads were frequently squeezed into short poems and this resulted in the poetry of cues: indirect leftovers of the stories in a lyrical poem (explicit already in Lermontov, but especially important for Fet).

If this change towards a shorter form can be traced, then why, when gaining the relative freedom from the genre system, does poetry need to be shortened? From the evolutionary perspective, change in size cannot happen independently from change in the whole organism: “For every type of animal there is a most convenient size, and a large change in size inevitably carries with it a change of form”, states the famous essay of J.B.S. Haldane (Haldane 1927: 18)\(^1\). Similarly, Tynyanov in 1923 wrote on the differences between large and small poetic forms: “The calculations for large form are not the same as for small form, every detail, every stylistic device has a different function, depending on the size of construction, has a different power, carries a different burden” (Tynyanov 1977).

Despite this manifest importance of the size of a poetic construction, the length itself was rarely studied with quantitative evidence, and it was never done on a large historical scale, using the whole population of poetic forms (or at least a large sample) instead of observing few canonical authors. In this paper, we try to trace the change in length of Russian poetry during almost the entire history of the modern Russian literature, to solve the question of the “convenient size” for poetic “species”, and to explain the possible reasons and pressures that were driving this change within the framework of cultural evolution theory.

---

\(^1\) On the topic of size in literature (character networks in the novel and title length) see also Moretti 2013: 179–210; 231–237.
Data

For the study we used data of the poetical subcorpus of the Russian National Corpus\(^2\), which is the largest well-structured and marked up digital collection of Russian poetry of the 18–20th centuries\(^3\). We had access to a database dump of the poetical subcorpus (made in 2012): a collection of separate .xml documents with texts and metadata.\(^4\) Files were parsed to extract different data, which was combined into one large table: it contains information on \(\approx 43,000\) poems.\(^5\) The length in lines was counted for each poem. We chose to observe the number of lines instead of the number of words because lines provide more information about the form and genre: for example, a poem of 12 lines occupies the space of three quatrains, while that of 4 lines suggests a small genre of epigram, epitaph, or inscription.

Not everything in the corpus could be controlled. The main problem is that there is no way to separate texts based on their editorial status (finished / unfinished, published / left in manuscripts, etc.), and thus there is no control on the cultural impact of specific texts (the corpus includes many Alexander Pushkin's unfinished drafts, for example). Another notable source of noise is dates of composition: some of the texts are undated and represented in the corpus within the time interval in which they could have been written – often this period extends to 20–30 years or even more. To decrease the possible bias in chronology, we decided to exclude from the observation those poems that have this interval larger than 10 years. For the rest of uncertain dates the upper year mark was taken as the approximate time of composition.

For the most part of the study, we will be using texts written between 1750 and 1921. Before this period, the corpus data is simply too scarce, as the system of modern Russian poetry was just emerging; after that, the data is unreliable, since Russian poetic tradition clearly splits to three separate branches – Soviet, underground, and emigrée – and there is no way in the corpus to distinguish

\(^2\) http://ruscorpora.ru/search-poetic.html.

\(^3\) This corpus is not a whole population of the Russian poetry, but a large sample that, despite following certain rules of representativeness, is skewed towards the canonical texts and the texts available in the 20th century editions. Some principles of design of corpus contents are described in Korchagin 2015.

\(^4\) We are sincerely grateful to Boris Orekhov and Kirill Korchagin for the access to the database dump. Without it this study couldn't be done.

\(^5\) The processed data and the source code for plots used in this paper are freely accessible: https://github.com/perechen/russian.poetry.length.
between them automatically. Besides that, collection of the 20th century poetry may be not very reliable due to authorship restrictions and the lack of some key authors\(^6\).

### Results

In Figures 1.1 and 1.2, the mean and median poetry lengths were counted by both years and decades. There is a clearly observable decrease in length – it is more evident in Figure 1.3, where extreme values of the 18th century were excluded from the observation. However, there is also a decline in lengths diversity: mean values are rapidly closing the gap with the median.

![Figure 1.1. Mean and median values of length in lines counted for each year (1750–1921, sample size = 34,000)](image-url)

\(^6\) This problem is certainly important for our sample made from Russian National Corpora in 2012, since then the 20th century in poetic corpus became much more representative.
Figure 1.2. Mean and median values of length in lines counted for each decade. Smoothing lines added (loess)

Figure 1.3. Closer look on the 19th century (1800–1921, sample size = 31,000)
The median represents centrality, it could be seen as a dot that separates the higher and the lower halves of values, so this measurement is resistant to outliers. Because short poetry is dominant (95% of poems in our sample are shorter than 100 lines), median values remain relatively low, compared to the means that are sensitive to big numbers of large forms. However, this difference between the mean and the median was fading during the second half of the 19th century, and suggests that short forms were overpopulating Russian poetry, and that large forms had less and less impact on the means.

This difference could be also observed in particular authors: while Pushkin’s gap between the mean and the median is 29 lines (with values 43.4 and 14 respectively), poets of a predominantly short form have mean even lower than median: Afanasii Fet (16.99 and 18), Alexander Blok (17.5 and 18), Anna Akhmatova (11.56 and 18). This trend reaches its maximum in the Modernism – when reinvention and outburst of lyrical poetry happened. It is hard to address the latest situation with our sample as seen on Figure 2, where the timeline was extended to 1980s. Results remain uncertain; should we attribute the increase in median and mean length in the 1920–1930s to the Soviet period, when poetry was trying to bring the large form back – new epic and heroic narratives? Or is this just a flaw in the sample, which is small and could be seen as not representative (dispersion of values is particularly high in the second half of the 20th century, where data is very scarce)?

These measurements, however, give only a rough general picture, because everything is counted together within the corpus: large drama in verse and Pushkin’s drafts. Another possible problem that could bias the results is that large forms in corpus sometimes can be hard to detect because of its structure, where long texts are frequently broken into smaller parts according to their internal organization by “sections / chapters / parts”. For example, Mikhail Kuzmin’s poem The Trout Breaking through the Ice (Форель разбивает лед) is represented in the corpus not as one poem of 524 lines, but as 15 smaller poems-parts (in this case, the split is a good decision – and it will be discussed later).

We also expected that the drop in length would accompany the disintegration of the genre system theorized by formalists – but there is no visible change in the length during the first half of the 19th century. A probable answer is: comparing to the second half of the 19th century, the first half provided a vast amount of space for large poetic forms: drama, romantic narrative poems, even a “novel in verse”. Short forms existed in the diverse population, and that diversity had disappeared, thus Figures 1 and 2 depict this most visible change, but they say little about short lyrical poetry. Obviously, short forms have always existed, but there are reasons to expect that their “short” length
wasn’t constant, and in fact became even shorter. To trace this change and switch to the scale of “small” poems, some control over the data should be acquired.

Poetical corpus is tagged for genres and we can remove large poems (поэма), dramas and verse novels from our observation. Also, one case requires a special treatment. According to the markup used by the creators of the Russian poetic corpus, a “cycle” (цикл) means: any poem that was considered a part of a larger text unity, or this larger unity itself. Since there is no way to automatically distinguish between these two meanings, we decided to exclude all the texts marked as “cycle” from our corpus.

That will exclude many texts, not only large ones, but also small poems, however, this operation can remove possible bias that emerges by treating actual large forms as a number of small texts. Finally, all texts larger than 100 lines and smaller than 5 (4 lines stand for traditionally small genres like epigram and include unfinished pieces and drafts) were excluded to access the lowest part of the size scale, where the unidentified “lyrical poem” is arguably located.

---

7 Corpus genre tags sometimes only reflect genre attribution in the title of the poem (so “song” tag is added for poems, which have the word “song” in their title). Given the extreme uncertainty in the discussion on genre problems, this decision allows to have only basic and rough control of different poetic forms in the corpus.
The shortest species: how the length of Russian poetry changed (1750–1921)

Figure 3. Mean and median for shorter poems (lines < 100, > 4; 1750–1921, sample size = 32,000). In this cleanest sample we calculated Pearson's correlation for variables in time, which is highly significant ($r = -0.80$ (mean), $r = -0.72$ (median); $p < 0.001$)

In Figure 3 two significant decreases in poetry length during the 19th century become clearly visible. The first happens in the 1820s, right where we expected the decomposition of genre system and the beginning of a new lyrical form; poetry gets compressed even more in the 1880s, losing about one quatrain in size. Roughly, the average short poem became two times shorter during this 150 year period of modern Russian literature.

Here we were tracing an “average” length poem, but what does this average consist of? To trace this population of short forms, a distribution of poems of small sizes (number of lines below 50, 90% of corpus) was made, as shown on Figure 4.
Figure 4. Quantities of poems of small size (lines < 50; 1750–1921, sample size = 30,000)

Note how all the most common lengths are the multiples of four (8, 12, 16…), which points to the high sustainability of a quatrain’s syntactic and compositional scheme in Russian poetry. This is the outcome of rhyme preservation: rhymes are inducing the underlying 4-line structure (especially ABAB, ABBA patterns) to verse. However, there is an important exception: the number of poems of 14 lines is unexpectedly high. This has a simple explanation: a sonnet.

Sonnets, of course, became popular in the Romantic Age, were widespread in Modernism, but had existed long before Russian modernity started. It was probably the most successful and widespread form in the new European lyrical poetry; invented in the literature of the court (Wilkins 1915) a sonnet initiated a shift from singing lyrical poetry to reading it (Oppenheimer 1989: 1–29). By isolating canzone (or canso) stanza (Kleinhenz 1986) a sonnet acquired its short structure of 8+6 lines, and became a vital device for formally diverse expression of lyrical emotion in a constant, limited size. The size of the most widespread short forms of Russian poetry was approximating the length of a sonnet (12 and 16 lines) – and in fact never compressed more than this.

---

8 See the overview on early history of the sonnet in Kennedy 2011.
Figure 5 shows how these extra-short forms (8–20 lines, 60% of all poems in the observation) were overpopulating Russian poetry in time, taking more and more relative space. This suggests why, among other things, the Sicilian-born poetical form of the 13th century was successful, its luckily coined size was of high importance – probably convenient form for European lyric poetry that was rediscovered many times in different national literatures.9

![Figure 5](image_url)  
Figure 5. The weight of the most spread short poems (8–20 lines, poems of 14 lines excluded) in the sample size measured for each decade (1750–1921)

Cyclization of short forms

As Russian poetry was becoming shorter on average, and the genre system was losing its influence over it, specific forms developed to constitute larger unities of lyrical texts. The shortest species couldn't fit in the old genre structure that divided poetic space into clear sectors (odes were united with odes,  

---

9 Of course, poetry length should be investigated in different literary traditions to explore the hypothesis of “convenient” size for the European lyrics.
elegies with elegies, etc.). So instead of the unity of grouping by genre, these short forms found more ambiguous and flexible ways to be united in a book of poetry, its sections, or a thematic cycle (Leibov 2000: 44–46). This way of short poetic forms living in lyrical “colonies” within a book was called the “large form” by Tynyanov. This “large form” was widespread in Modernism, and this can be clearly seen on Figure 6, where the weight of poems with the tag “cycle” was counted – it gives only a rough estimate of this process, but nevertheless shows the increasing cyclization of lyrics, which was happening while Russian poetry experienced a continuous decline in length.

Figure 6. The weight of the poems with a “cycle” tag in the sample (1750–1921)

Modernism presented a wide diversity of “large forms”, with different strengths and structures of links between the parts: from Annensky’s “shamrocks” of three poems, to the lyrical “trilogy” of Blok. Andrey Belyj and Mikhail Kuzmin invented a specific type of larger poem: links between poems-parts are stronger than in a cycle but weaker than in a unified text (e. g. the Romantic poem). This type survived in the Soviet underground poetry in a genre that Elena

---

Schvarts called in her works a “small narrative poem” (маленькая поэма). But in all these cases the “large form” should remain a metaphor, as larger unities weren’t particularly stable, they were still made of small blocks of short poems that differed from each other significantly. These forms introduced to the literature what could be called the serialization of lyrical poems, and trained both readers and scholars at perceiving lyrical colonies as inseparable entities. That is why the corpus structure that splits Kuzmin’s *Trout* into parts is highlighting the way how lyrical poetry was assembled to form larger unities. The spread of lyrical “large form” in Modernism shows how short poetry expanded in literature and freely moved to the spaces that were previously occupied by large poetic constructions. In some sense, short poetry invented a way of being large by remaining small.\footnote{Another instance of lyrical expansion in the 20th century is the “large lyrical poem” (большое стихотворение) – an important poetic form for Joseph Brodsky, clearly opposed to the large narrative poems (песни) in some of his self-descriptions (Volkov 1998). While this form in Brodsky’s poetry was clearly dependent on English tradition of a long metaphysical poem and bears lesser resemblance to Russian neo-archaic large elegy (like *Autumn* of Baratynsky), the possible growth of the small lyrical form into the larger build in the 20th century should be investigated closely. If there is a trend (beyond Brodsky), it could mean forces that were at work during the 19th century disseminated at some point – and that the lyrical poem was expanding freely as role of a poetry in culture and its social impact declined.}

Cultural evolution

Quantitative observations on the length of Russian poetry reveal a long-term trend of change towards shorter forms. Features and scope of this trend could be further clarified by running measurements on different samples and acquiring better control of the data, but the pattern is clearly present. Similar long-term trends were also found in different cultural domains: various examples provide observation on technological complexity and diversity (increase in the number of technology parts, lines of codes, etc.; see: Kelly 2010: 278–292); average shot length decreased during entire film history (Salt 2009: 366–368), while the visual activity and darkness of the screen increased (Cutting et al. 2011); title length in an English novel decreased as the literary market was growing and new genres emerging (Moretti 2009); Anglophone literatures experienced continuous decline in the use of lexicon of positive emotions (Acerbi, Morin 2016).
These observed cultural trends could be understood within the theory of cultural evolution, a study field that aims to explore and explain cultural change with basic Darwinian principles and provide coherent empirical framework for the humanities (see: Richerson, Boyd 2005; Mesoudi 2011; Mesoudi 2016 for recent review of the field). Mechanisms that drive cultural change, transmission and diffusion (see Morin 2016), success and failure of particular traditions and cultural forms are the domains studied by cultural evolution. While the field itself remains rooted in anthropology, social sciences, and technology studies, several attempts were made to include literature in the discussion on the evolutionary process of culture (Morin, Acerbi 2016). One of the key innovations of cultural evolution theory is the introduction of population thinking to the domain of cultural forms: these are the scope of population, where variations occur, and the selection mechanisms (or drift) drive the successive change. The expansion of various forms of “distant reading”, which were recently amplified by the accessibility of large digital text collections and computational approaches to textual analysis, was actually accompanied by the analogues of evolutionary thinking in literary theory12: with studies of the formal variety and branching in literature (Moretti 2000), and accessing the literary population (e. g. going beyond canon to inspect literary diachrony (Algee-Hewitt et al. 2016, Jockers 2013, Underwood 2016)). When the change in length of Russian poetry is considered a population change, several ways of explaining it emerge.

The shortest species

Russian poetry during the 19th century was obtaining its “convenient form”; this form appeared to be small and its size resembled that of an old European invention in the field of lyrical poetry – a sonnet. Partly, the success of these short forms is related to the changes in literary environment and the fall of large poetic constructions. Russian poetry in its modern version appeared in the 18th century, when European literary space provided lots of open opportunities for poetic language: from drama to philosophical treatises and guides on garden design. These niches, however, collapsed rapidly during the course of the 19th century. Highly successful realistic conventions put a constant

---

12 An early and influential attempt to provide literary studies with quantitative evidence and empirical background made by Boris Yarkho was also rooted in naturalistic and evolutionary explanations (Yarkho 2006).
pressure on poetry. Appearance of a novel with its complex narration made large poetic narratives (constrained on various levels) irrelevant; the same happened in the realistic drama, and only a small niche for simplified poetic language was left in theatre – it was within opera where poetry was largely a subordinate of voice.

However, lyrical poetry appeared to be sustainable; shrunken to its shortest form it could withstand a novel and expand dramatically in Modernism. As a literary form, it has provided something that other literary forms couldn't conquer or replace. There could be different explanations for what it was, but the inherent features of a small-sized lyrical space suggest a few important ones. First, short form could be argued to be a good attention controller – it demands very little effort from a reader in terms of time (to leave other “costs” of poetic language out of discussion). Roughly speaking, a short lyrical poem of an advanced poetic tradition could be seen as a short and sophisticated rhymed advert, radio jingle, children rhyme or limerick, which are easily moving in and out of our attention window.

Second, Russian poetry did not give up on rhyme and for a long time remained accentual-syllabic, never being actually transformed fully into vers libre. Mikhail Gronas argued that the longevity of rhyme and verse structure in Russian poetry is related to the spread of specific practices of memorization in Russian poetic culture, including the conservative school education, the continuous pressure of censorship, the space of “unpublishable” literature, and the wide diffusion of oral traditions in prisons and later in Gulag (Gronas 2011: 71–96). Poetry has been pushed to remain memorizable, and short form provided a convenient package for mnemonic devices. If this hypothesis is correct13, then “memory bias” (how culture evolution theory would call it) was notably influencing the evolution of Russian poetry, similarly to how memory shaped folklore texts in oral cultures (Rubin 1997). If the spread of short forms is an outcome of this pressure towards memorability, then in traditions where poetry transitioned to predominantly free verse, and, accordingly rejected mnemonics the average poem length in this period of transition should have increased.

But neither size, nor memorability potential alone can explain the success of short lyrical form. Other poetic genres, from children rhyme to the popular song, are both short and mnemonical (probably even more aggressively) – and we know that these intuitively labeled “simple” forms do not make up the core

---

13 See the discussion on Gronas’s “mnemonic hypothesis” in Russian publication of the book chapter (Gronas 2012).
of Russian poetry. The important distinctive feature for the short lyric could be then located in its “complexity”.

Short lyrical poetry provided quick emotional output and was forced to master this emotional control in the very limited space, leaving aside other constrains of the poetic language such as meter structure and rhyme patterns. The poetic capability to carry complex messages in a highly constrained formal environment is often discussed as “informational paradox” (Abernathy 1961; Lotman 1976: 29–36). General information theory implies that increased number of constraints leads to a decrease in unpredictability and, thus, to a decrease in information capacity that message could carry – apparently, this is not what was happening in advanced poetic traditions. Dmitrii Manin’s experimental study, of the relation of constraints and unpredictability in Russian poetry found evidence for compensatory mechanisms for this information loss:

If the low metrical unpredictability of poetry indicates a narrowed space of possibilities for poetry compared to prose, then the higher non-metrical unpredictability means the compensatory expansion of this space in a different dimension, by relaxing the semantic and syntactic constraints on word combination. (Manin 2012: 293)

To rephrase – by maintaining many constraints on various formal levels, poetry could acquire extreme freedom in semantics while not becoming incomprehensible. If the assumed pressure for memorability is taken to account, then Russian poetry, especially in the 20th century, was shaped by the increased complexity within highly constrained mnemonic form – poetry struggled to be both sophisticated and “simple”14, e. g. to be widespread and to exploit a popular taste for conservative forms.

The decrease of average length of poetry could have also contributed to this “complexity under constrains” by keeping the available poetic sandbox very small. It seems relevant that important inventions in Russian lyrical semantics of the 20th century were made in extra-short forms (and while the average length became shorter than ever) – by Akhmatova and Mandelshtam (Levin et al.). Some of the consequences of poetry being short were noted by Tynyanov (on Tyutchev’s verse):

14 In a way in which mnemonic device could be “simple” by increasing redundancy of a message.
...words, that are being lost in the vast space of a poem, gain an extraordinary significance in the small space of a fragment [...] Ode became microscopic and have focused all the power on the small scale [...] Fragmentarity, the small form, which narrows the field of view, is empowering its stylistic features enormously. (Tyutchev 1977b: 46)

Arguably, this “focusing of power” happened not only in the lyrical fragment of Tyutchev. Any compressed lyrical form of a tradition in a certain time could have experienced this “empowering” that triggered semantical shifts and inventions, which lead poetical language to an increased complexity.
References


Martynenko, Antonina I. 2016. “Stikhotvorenija E. Baratynskogo” 1827 g.: istoria i predystoria sbornika. [manuscript of the thesis]. Higher School of Economics, Moscow.


