Quantitative Approaches to Versification: Conference report (June 24–26, 2019, Prague, Czech Republic)

Igor Pilshchikov, Vera Polilova

The Quantitative Approaches to Versification conference met on June 24–26, 2019 in Prague. The conference, which continued the tradition of the regular Slavic Verse (Slavjanskij stikh) and Comparative Slavic Metrics (Słowiańska metryka porównawcza) conferences dating back to the 1960s, was dedicated to Květa Sgallová in recognition of her 90th birthday. Sgallová, one of the most prominent Czech experts in versification, honored the conference with her presence. The conference was jointly organized by the Institute of Czech Literature (Czech Academy of Sciences) and the Institute of Russian Language (Russian Academy of Sciences) and gathered a group of international scholars who share an interest in the formal quantitative analysis of all aspects of versification. The Prague meeting featured more than twenty presentations, all in English. The program is available at http://versologie.cz/conference2019/.

The main topics of the conference included a broad range of issues unified by an umbrella concept of quantitative approaches to the study of verse: new approaches to comparative metrics, descriptions of ancient, medieval and modern European syllabic, accentual and accentual-syllabic meters, an analysis of texts with an unknown metrical status, an automated processing of versified texts, the use of corpora-based methods in verse studies, a statistical analysis of verse rhythm and specific rhythmic varieties of meters, as well as analyses of various measurable aspects of poetic texts, such as semantic patterns and lexical diversity, or the relationship between syntax and rhythm in verse. Not only written, but also spoken and sung poetry was analyzed. According to the organizers, the conference comprised “various methodological perspectives ranging from simple descriptive statistics to advanced machine learning methods (such as support vector machines, random forests or neural networks) as well as material covering a large span of time and languages: from very ancient versifications (Sumerian, Akkadian, Hittite; Ancient Greek), through medieval...
(Old English, Old Icelandic, Old Saxon) and Renaissance verse to modern experiments (free verse, concrete poetry); from English and Russian through Spanish and German to Portuguese and Catalan” (QAV 2019: 275). What all the presented papers share in common is that “versification is being studied in the context of other linguistic phenomena that may affect or determine it” (Ibid.). Our experience of the venue fully confirms this claim.

Conference organizers rejected the concept of keynote speakers; the panels were thematically structured but were all plenary. This helped the participants to combine the scope and representativeness of an international conference with the vivid nature of a workshop, where an affinity group of scholars meets to debate and exchange opinions. By the beginning of the conference a book of extensive synopses – in fact, articles – had been published in print and online (see QAV 2019), so that the audience had a chance to prepare themselves for almost every talk in advance.¹ This design proved to be very fruitful and the conference discussions were no less inspiring than the presented papers.

The conference was inaugurated by the director of the Institute of Czech Literature Pavel Janáček, who in his opening speech commemorated the work of Květa Sгалlová and her collaborator Miroslav Červenka, as well as their long-time participation in Comparative Slavic Metrics international projects.

The opening panel featured only one paper. Salvador Ros and Javier de la Rosa’s “Using deep learning paradigm to improve syllabic versification: A first approach” was presented by Javier de la Rosa. The presentation was part of the POSTDATA (Poetry Standardization and Linked Open Data) project funded by the European Research Council in Horizon 2020 Framework Programme for Research and Innovation. The coauthors are representatives of LiNHD (the Spanish National Distance Education University’s research center on digital humanities), which has long been working in the field of automatic text processing. They described the prospects of using neural networks for verse analysis, particularly for automatic scansion, and proposed using an inference-based approach instead of a rule-based approach that defines syllable boundaries, language stressing, metrical stressing and rhythmic patterns. The presenter listed the currently existing automatic rule-based scansion systems for Spanish and English created between 1996 and 2018 (such as Scandroid, Zeuscansion, Poesy, AnalysePoems, etc.), as well as recently created systems using neural scansion (IXA for English and Spanish, KAIST for English). He

¹ A few papers were not included in the book; a few others were included but were not delivered. The proceedings have been included in the Web of Science Core Collection. A copy of the book on academia.edu has been viewed more than 1,000 times.
Igor Pilshchikov, Vera Polilova

examined the principles of BiLSTM (Bidirectional long short-term memory) used in neural scansion and the four prosodic phenomena that affect the number of syllables in Spanish verse: synaloepha, hiatus, syneresis and dieresis. In their future work, the coauthors plan to use Words and Sentences Embeddings based on the language model BERT (Bidirectional Encoder Representations from Transformers by Google). The advantages of this model are many; it is multilingual, pre-trained and fine-tuning. The coauthors believe their approach is paving the way for an unsupervised multi-language scansion system.

The next panel was devoted to a structural and statistical analysis of Germanic and Slavic verse. It opened with Evgeny Kazartsev’s paper, titled “Probability and cognitive models of verse meter”, in which he reported on the establishment of iambic versification in early modern European poetry with a focus on iambic tetrameter. The study explores the development of iambic verse in Northern Europe and the Baltic area. The inclusion of Russia in this cultural zone after the Northern War and the reforms of Peter the Great predetermined, according to Kazartsev, the fate of Russian and East Slavic poetry. His comparative research is devoted to studying the mechanisms of intercultural and interlingual communication involved in the formation of syllabotonicism in different languages. Kazartsev studied the rhythmical structure of iambs against the background of probability language models of meter, which are constructed on the basis of rhythmical vocabularies of fiction prose. He distinguishes between two types of models: the Language Models of Independence (which are based on the presumption of the independence of rhythmical words in a metrical line) and the Language Models of Dependence. Kazartsev’s study revealed that models of dependence are better suited to describing the early stages in the development of metrical verse. Despite the similarity of certain tendencies in the use of rhythmic structures, he discovered significant differences in the rhythm of early Dutch, German and Russian iambs. The presenter’s hypothesis was that this difference is due to the fact that “German and Russian iambs arose under identical conditions not as a result of natural evolution, as was the case with Dutch verse, but due to a substantial and quite abrupt reform” (QAV 2019: 111).

In their “Verse and prose: Linguistics and statistics”, Tatyana Skulacheva and Alexander Kostyuk discussed the most stable (i.e. statistically proven) linguistic differences between verse and prose found in texts written in several European languages. These differences regularly occur in verse of different meters, different systems of versification, different periods and literary styles. What, in particular, is different, according to the presenters? First, syntax. Parataxis predominates in verse, whereas hypotaxis is less frequent in poetry than in prose, as the calculations for Russian, French and Spanish poetry and
prose from the 17th to the 20th century demonstrate. Then, intonation. An analysis of intonation in the Praat program and a statistical analysis of the obtained data shows that Miroslav Červenka’s idea that verse has a specific intonation (“intonation of enumeration”) is correct. Perceptive analysis (listening by 23 informants) has also shown that people easily recognize verse intonation in contrast with prosaic, even in a delexicalized text. There are many other differences relating to semantics, information processing, mistake detection, etc. In particular, participants in the experiments easily detected mistakes in prose, while mistakes in verse almost always remained undetected (Kimmelman 2012). The recent data obtained by neurophysiologists in the U.S. (the Carnegie Mellon group) and the U.K. (the group based at Liverpool and other British universities) enabled the presenters to suppose that the peculiarities observed in verse may be linguistic mechanisms of activating imaginative thinking at the expense of logical thinking.

Igor Pilshchikov’s “Rhythmically ambiguous words or rhythmically ambiguous lines? In search of new approaches to an analysis of the rhythmical varieties of syllabic-accentual meters” reexamined the problem of rhythmical ambiguity in Russian poetry. Analysis of the rhythm of Russian syllabic-accentual verse confronts the problem of how to accentuate words whose natural-language stress is weaker than that of fully-stressed words. Viktor Zhirmunsky called such words “ambiguous” and formulated a rule: they should be considered stressed in “strong” (ictic) positions and unstressed in “weak” (non-ictic) positions (Zhirmunsky 1966 [1925]: 93–113). Pilshchikov noticed that in many poems the fully-stressed lines that contain rhythmically ambiguous words are often isomorphic with the form that predominates in the poem. He argued that what is important here is the distribution of the ambiguous forms in the line, i.e. not rhythmically ambiguous words, but rhythmically ambiguous lines. The statistics of rhythmical forms may be complemented with an analysis of ambiguous forms, better reflecting what the Russian Formalists called the rhythmical impulse of the poem. A revised and extended version of this paper was published in the previous issue of Studia Metrica et Poetica (Pilshchikov 2019).

Petr Plecháč and David J. Birnbaum’s paper “Assessing the reliability of stress as a feature of authorship attribution in syllabic and accentual verse” was built on a recent study by Plecháč, Bobenhausen and Hammerich (2018), which shows that statistics relating to versification may be used for authorship attribution. One such statistic is the stress profile of a poem, i.e. a vector consisting of frequencies of stressed syllables at particular metrical positions. The initial hypothesis was that, because syllabic versification regulates the number of syllables in a line but not the distribution of stresses, it allows authors to
individualize their rhythmical style much more than accentual-syllabic versification, where the distribution of stresses is primarily determined by meter. For that reason, Plecháč and Birnbaum expected the stress profile to be a more reliable indicator of authorship in Spanish syllabic versification than in Czech or German accentual-syllabic versification. This hypothesis, however, was not supported by their analysis. This failure led them to another idea: measuring the tonic entropy of the data described above (a modification of the procedure of using the Kullback entropy as a measure of rhythmic diversity proposed in Dobritsyn 2016). They demonstrated that “Czech is in this respect closer to less regular Spanish than to more regular German”. As a consequence, versification-based attribution dealing with strongly regulated versification systems (such as German) “would require more features than languages with less regular versification systems” (QAV 2019: 209).

The late afternoon panel discussed the correspondences between metrical, prosodic and phonetic structures in verse.

In “Some characteristics of sound patterns in English verse”, Olga Barash and Georgy Vekshin discussed a certain kind of sound repetition in English poetry. Their task was to determine whether the concept of phonosyllabic patterns, consisting in repetition of groups of identical or similar consonant sounds within a potential syllable, is valid when applied to poems written in English. This method of analysis was previously tested on Russian poems using the Phonotext 2.1 software presented in Vekshin’s introduction to Vekshin and Barash’s joint talk.2 The material of the research comprises randomly chosen accentual-syllabic and accentual poems by English and American poets of different centuries and literary schools, from Thomas Wyatt to Dylan Thomas. Lord Byron’s poem “Sun of the sleepless! Melancholy star...” was discussed in detail. The methods used for measuring sound cohesion were based on a “syllabocentric” view of the sound composition of a poetic text. This approach proceeds from the fact that the elementary speech constituents are syllables or syllable-like entities, outside of which no sound and rhythmic associations of text segments are possible (Vekshin 2006, 2008). The sound chain of verse is granulated in accordance with the syllabic organization of speech and its prosody. The functional significance of individual granules is provided by analogies in the structure of the chain – by sound repetitions that, because they are embedded in the rhythmic and prosodic contours of words, phrases and text, complicate and redistribute its syntactic and morphological links.

---

2 This introduction was initially announced as a separate paper under the title “Automatic detection of sound repetitions and quantitative measurement of sound cohesion in Russian verse: the Phonotext 2.1 software.”
and thereby ensure the multidimensionality of the poetic form. Another complicated issue discussed in the paper was the question of grapho-phonetic and graphic repetition (when letters rather than sounds are repeated).

Patrik Bye presented a paper on “English iambic meters and the tension asymmetry”. The corpus for the study consists of 125 isometric poems (47 in tetrameter, 73 in pentameter and 5 in hexameter) from the last 400 years. It has been revealed that, in comparison with iambic tetrameter, iambic pentameter allows greater metrical tension – specifically, a higher relative frequency of prominence mismatches. The presenter attributed this Tension Asymmetry to the uneven metrical structure of the pentameter and supported his hypothesis by the finding that iambic pentameter does not only show inversion more frequently, but tolerates inversion at lower levels of the prosodic hierarchy. The crucial role of even vs. uneven structures in explaining the Tension Asymmetry is also supported by the discovery that (balanced) hexameter is at least as stringent as tetrameter. Iambic pentameter is more tolerant of initial inversion at lower levels of the prosodic hierarchy (inside the phonological phrase and even the procliticized prosodic word), which tetrameter seems to disallow. The hypothesis that best explains the facts seems to be “the Nonlinear Metrical Hypothesis, according to which the difference between even and uneven metrical structure[s] plays a crucial role. The Tension Asymmetry is thus an interesting example of a statistical expression of an underlying structural asymmetry” (QAV 2019: 47). The presenter hypothesized that this fact explains the expressive capabilities of the meter, which led to the cultural dominance of iambic pentameter in English poetry.

The first panel of the second day opened the floor for a discussion of oral poetic traditions.

“Homeric formulas and meter” by Sophia Sklaviadis and James K. Tauber was a visualization of the formulae in the Homeric corpus, through which it is possible to explore stylistic features such as the correspondence between formulae boundaries and metrical boundaries. The coauthors used a Python script to count all repeated n-grams in the Homeric corpus, visualized n-grams at a high level across the entire corpus and constructed a reading environment in which repeated n-grams in the text are indicated by color. N-gram variants can be defined in terms of equivalence classes: n-gram equivalence class may be set by the user, based on tokens or on any function of surface forms. For example, n-grams can be visualized over orthographic normalization, lemmatization, part of speech and meter. The n-gram analysis is potentially useful for semi-automated treebank correction based on partial n-gram matches at
Alexander M. Petrov presented a paper on the “Metrical types of bylinas (Russian epic folk songs) in the collection of the Institute of Linguistics, Literature and History at the Karelian Research Center of the Russian Academy of Sciences”. He discussed the typology of epic verse exemplified by the material stored in the Audio Archive of the Karelian Research Center in Petrozavodsk. The problem of analyzing folklore verse is complicated by two factors: 1) until recently, audio recordings of folklore texts were not available to researchers and 2) metrical studies lacked a developed method for the analysis of folk verse. Petrov applied the method for analyzing folklore texts proposed by Mikhail Gasparov and developed by Tatyana Skulacheva (see Skulacheva 2012, 2014). The following five indicators were used: metrical structure, anacrusis, clausula, number of stresses and number of syllables. A total of 25 texts and fragments collected from the Russian traditional singer Fyodor Andreevich Konashkov (1860–1941) that present Russian epic stories (about Duke Stepanovich, Ilya of Murom, Dobrynja, Vol’ga and others) were analyzed. The result is in many aspects different from the established view of the bylina meter. The author acknowledges that the material (441 lines from 25 texts) is not sufficient for definitive conclusions. He intends to continue this research with the recordings of another 100 epic songs (texts or fragments) that are available but require transcriptions and markup.

Mari Sarv’s paper “‘The meter of Estonian folksong is always the same’: How to visualize regional variation of folk meter” was devoted to Estonian runosongs, a branch of common Finnic musical-poetic traditions usually described as Kalevala-metric (with some deviations). Kalevala-meter is a set of versification rules discovered and noted down by Finnish scholars on the basis of Karelian folksongs. These rules have been overtaken by Estonian research tradition without much critique. As a result, the significant variation of folksong meter has remained almost unnoticed in scholarly literature, in school-books and other writings on the Estonian runosongs. In some cases, the Estonian runosong-texts have even been “repaired” according to the Kalevala-model, and literary imitations of folksongs are often written in Kalevala-meter deviating from real Estonian folksongs. Variation is an essential feature of folklore, therefore it is natural that in the process of transmission songtexts are in constant renewal and adaptation. Sarv’s studies, based on an analysis of folksongs from 101 Estonian parishes, have shown that the meter in the songs noted down mainly between 1880–1920 has been in transitional stage from specific quantitative meter to syllabic-accentual one (see, in particular, Sarv 2019). In the case of Estonian folksongs, metrical variation reflects
linguistic variation and the prosodic structure of dialects. On the other hand, the constant re-creation inherent to folklore also manifests itself in meter, and metrical variation reflects the changes in the prosodic structure of language. The aim of the paper was to detect the “metrical dialects” of the Estonian runosong on the basis of several metrical features and compare different possibilities of detecting and visualizing metrical grouping, from bundling isolines on the maps to various calculated alternatives.

After the coffee break the participants reunited to discuss less studied metrical forms, from the world’s earliest poetic traditions to the avant-garde experiments. What these papers had in common was an uncertainty about the prosodic basis of these poetic texts, or their rhythmical specificity, or – in case of the first paper devoted to what possibly was the world’s earliest form of verse – even doubts as to their poetic status.

The paper by Rim Nurullin, Nadezhda Roudik, Maria Molina, Andrei Sideltsev and Tatyana Skulacheva “The most ancient verse in the world (Sumerian, Akkadian, Hittite): quantitative analysis” was presented by Tatyana Skulacheva. The paper discussed the methods suitable for analyzing the most ancient verse in the world, which is not strictly organized. The discussion was exemplified by a quantitative analysis of verse texts in Sumerian, Akkadian and Hittite languages. Though there are many hypotheses as to the system of versification for proto-Indo-European or even all-world verse at its initial stage of development, the most ancient examples of verse have not yet been thoroughly studied in terms of their versification. That is why Sumerian, Akkadian and Hittite texts (some of which were written from the 23rd to 13th centuries BC, before the much better known Ancient Greek) were examined using the following indicators: number of syllables, number of stresses, number and distribution of long vowels and distribution of heavy syllables within a line. Comparing quantitative data for the three languages enabled the coauthors to suggest that “both syllabic (Sumerian) and accentual (Akkadian, Hittite) systems” coexisted and that the choice of a particular system of versification was dependent on the peculiarities of the given language (QAV 2019: 173).

Andrew Cooper opened “A quantitative analysis of the Old English verse line” with the statement that scholars, beginning from Eduard Sievers, have produced many analyses of Old English verse, but the only point of agreement among them is that each line has four stressed syllables marked by alliteration. At the same time these lines show great variation in the quantitatively measurable indicators of metrical structure. Cooper presented “a metrical model of the Old English verse line using entirely quantitative criteria and without the need for a typology of acceptable forms” (QAV 2019: 54). To formulate it, he created a large corpus of Old English verse (14 texts for a total of ca. 7000 lines)
and annotated it for stress, syllable quantity and alliteration patterns. Cooper demonstrated that line length, foot size and metrical prominence are functions of a verse structure based around a quantitative line of 8–16 moras (as in Golston and Riad 2003), with a prototypical line of 12 moras. Metrical prominence is determined by foot length, which is distributed around 10 syllables (from 8 to 14) and 12 vocalic moras (from 8 to 16). To place the Old English data in the context of its two nearest traditions, two small balanced and representative samples of Old Saxon verse and Old Icelandic fornyrðislag verse forms. It was revealed that, despite superficial similarities, the presented model is applicable only to Old English. These findings allow all Old English verse lines to be described with a single metrical pattern, which is “based on a preference for verse feet to be congruent with prosodic words, and for prosodic words to be measured in length by the combined syllable weight of metrical feet” (QAV 2019: 72).

Mirella De Sisto in her “Micro- and macro-variation in verse: A typology of Romance Renaissance meter” presented the study of the main Romance Renaissance forms, traditionally divided into three groups: 1) the purely syllabic poetry of the French tradition; 2) the meter of Italian and Spanish hendecasyllable, tending towards iambic rhythm, and 3) the meter of Catalan and Portuguese decasyllable, tending towards syllabicity. The paper proposed a quantitative approach to verify or disprove this grouping and to answer the ongoing question about the presence of the iambic tendency in Renaissance meter (Gasparov 1981 [1978]; 1987; Piera 1980/1981; Nespor, Vogel 1986; Akimova, Bolotov 2009). De Sisto considered 20 samples from Italian, French, Spanish, Catalan and Portuguese, as well as less investigated varieties: Occitan, Neapolitan, Sicilian and Venetian (130 lines for each author, ca. two authors per language). The aim of the study was to investigate the extent to which the “tendency towards iambic rhythm” deviates from the iambic scheme, the microvariation within the groups, and the difference between traditions with a pre-Renaissance decasyllabic meter (Catalan and Portuguese) and those that directly adopted the new form. It was shown that, in terms of the two poles (“pure syllabicity” and “pure iambicity”) the two Portuguese samples are far apart: the pre-Renaissance poet Diniz is the furthest from iamb, whereas Camões is close to the samples with an iambic tendency, next to the Venetian and Neapolitan/Spanish sub-branches. French (Peletier and Du Bellay) is a sub-branch of the syllabic group, closely related to the Catalan (March and Torroella) and Occitan (Ventadorn and Guacelm) sub-branches. On the “iambic” pole we find two Neapolitan poets (Velardinello and Cortese) and the Spanish poet Garcilaso, whose poetry was deeply influenced by his trip to Naples and subsequent stay there. At the same time all forms under consideration are not constrained on the foot-level but should be considered
colon-based: strictness increases on the right edge of the two cola, in the middle and at the end of the line. “The comparison with the two Germanic samples shows how distant the Romance traditions are from an iambic meter and highlights the Germanic versus Romance distinction in terms of poetic instantiations” (QAV 2019: 90).

What is problematic here is, to our opinion, De Sisto's definition of the "iambic template": “The 01 sequences of every sample were processed in order to calculate the percentage of deviation from a perfect iambic pattern for each position” (QAV 2019: 82). This definition presupposes that, in accentual-syllabic verse, stresses cannot be skipped, and the template for iambic pentameter is 0101010101(0). But they can. Therefore, 01001010101(0) or 010100010101(0) (with skipped stresses) are as perfectly iambic as 010101010101(0). Furthermore, stresses can be not only skipped but also added. However, the possibility of skipping and adding stresses depends on the word-boundaries: a phonetic word-unit in a syllabic-accentual line can either skip a metrical stress or add an extrametrical stress, but cannot do both simultaneously. One version of this rule, as applied to Russian binary meters, is widely known from Roman Jakobson’s “Linguistics and Poetics” (Jakobson 1960: 361; see Pilschikov 2019: 62–64 for the discussion and references). Simplified, clarified and restricted to iamb, it can be reformulated thus: “iamb is a verse line, in which the stress of non-monosyllabic words may fall only on even syllables” (Lotman 1975: 302; cf. Shapir 2019 [2005]: 123). It follows from this rule that, in iamb, stressed monosyllabic words can occupy odd syllables without violating iambicity. Therefore, 1001010011(0) can be iambic or not depending on the word-boundaries: 110010100111(0) is iambic, whereas 1010101010111(0) is not. Thus, in Aleksandr Pushkin’s Mozart and Salieri (famously written in iambic pentameter) Mozart says:

\[
\begin{align*}
Něť – ták; bezdělitsu. Namědni nôč’ju & 1|1|0100|010|10 \\
Bessónnitsa mojá menjá tomíla, & 0100|01|01|01 \\
I v gólovu prishlí mne dvé, trí mýsli. & 0100|010|1|1|10 
\end{align*}
\]

None of the five syllables marked in bold violates the iambic template. The same rules work in German iamb. Moreover, both Russian and German meters are examples of strict iamb. Other traditions, such as English or Czech, allow for even more divergences from the “01” template.³

³ See, in particular, Červenka 1973, Tarlinskaja 1987 and Lotman 2019, among many other valuable studies.
Adiel Mittmann, Paulo Henrique Pergher and Alckmar Luiz dos Santos’s paper “What rhythmic signature says about poetic corpora” was presented by Adiel Mittmann, who analyzed various sets of verses in Portuguese and Spanish in order to establish the relative frequency distribution of rhythmic patterns, which the coauthors term the rhythmic signature, and to describe how such signatures relate to each other. The researchers built a corpus with more than 250,000 verses: first, they compared the poems written in Portuguese; then, they examined several books written by the same poet; finally, they compared Portuguese and Spanish poets. Two methods were used: the Gini coefficient and agglomerative hierarchical clustering. The results show that many authors have a distinguishable, sometimes unique rhythmic signature. At the same time a poet can deliberately mimic the rhythmic style of another poet. The third experiment showed that verses in Portuguese and Spanish are rhythmically similar and that differences in style can mostly be ascribed to the authors, not to the languages. The distribution of stressed syllables within a Portuguese and Spanish verse line is fairly varied, which allows poets to use the rhythmic patterns in specific ways. Longer verses tend to impose constraints as to which syllables are allowed to carry stress, but there is still room for variation. For instance, 10 syllable lines can belong to one of 22 standard patterns (or rhythmical forms, as the scholars of Russian verse habitually called them in the discussion).

Anne-Sophie Bories presented an analysis of Raymond Queneau’s free and strict verse using the Queneau database as a research tool. An extensive database was developed (ca. 8,000 lines) specifically to analyze Queneau’s versification, and in particular the relationship between the form and the meaning of his meters. An important component of Queneau’s poetics is his blending of trivial language and everyday grammar with rare and pedantic vocabularies, of humor and silliness with grave themes, and of scientific subjects with intimate psychoanalytical autobiography. This eclecticism is mirrored in his versification choices, with nearly half of his poems in strict verse, nearly half in free verse, and a fair amount of his lines being more difficult to classify. Queneau uses light meters for grave subjects, epic alexandrines for the glorification of plastic, and expected rhymes to create puns. This mismatching of subject and tone is typical of Queneau’s writing practices. The presenter emphasized some of the challenges of gathering good quality data, along with the utility of an economical approach, the necessity of quantitative methods to get a good sense of versification choices from a stylistic point of view, and the relevance of a dual method combining the quantitative, computational approach with a more traditional close reading of texts in order to achieve a truly hermeneutical understanding of the texts and their poetics. She argued that the quantitative
method is both acutely necessary to the study of versification and insufficient on its own, as versification analysis should be used to serve interpretation.

The afternoon section of the second day was devoted to Russian verse. In his “Rhythmical structure of Russian iambic tetrameter and its evolution”, Sergei Liapin reexamined the evolution of the most popular, most frequently used and most studied Russian meter: iambic tetrameter. For many years, it has been considered that the rhythm of Russian iambic tetrameter is formed under the influence of two tendencies: 1) stabilization of the first ictus after at least one unstressed syllable and 2) regressive accentual dissimilation, that is, the alternation of strong and weak ictuses from the end of a line towards its beginning (Taranovsky 1971). The doubts to this theory expressed as early as 1973 by Miroslav Červenka were recently confirmed; a number of studies have discovered the reality of linguistic factors (syntax is a main one) that form the rhythmic structure of a verse text. Liapin gathered statistical data for longer Russian narrative poems from the middle and the second half of the 19th century, that have never been investigated before. These data seriously question existing views regarding the evolution of Russian iambic tetrameter. For example, the second ictus of the iambic tetrameter of Konstantin Sluchevsky – an important poet of the late 19th century studied by James Bailey (1975) – is even more intensively stressed than in the second half of 19th century in general. The highly stressed second ictus discovered by Taranovsky can be explained by the fact that the second half of the 19th century is poorly represented in Taranovsky’s data, and longer narratives are almost fully absent from the calculations. Meanwhile, Liapin’s data show that the rhythm of the narrative poems of this period is markedly different from that of lyrical poems. The reason is that, in non-stanzaic narrative verse, sentences are less uniform in length and often have a more complex structure than in stanzic lyrical verse, and very often the boundaries of a sentence do not coincide with the boundaries of a line. Also, the so-called third rhythmical form of iambic tetrameter (with stress omitted on the second ictus) appears much more frequently, which, of course, lowers the number of stresses on the second ictus. The presenter posited that there is a correlation between the increasing frequency of sentence borders within a line (often followed by enjambments) and the increasing frequency of the third form and, accordingly, the lowering number of stresses on the second foot. Liapin also argued that investigating stress profiles alone cannot provide sufficient information for describing the structure and evolution of verse rhythm. The analysis should be improved by the use of more advanced linguistic and statistical methods.

In her “Taktovik or mixed meter? Rhythmic features of Russian non-classical verse (1890–1920)” Vera Polilova presented preliminary results of the
study of the four-ictus taktovik (Tk4) of Russian Symbolists. Recent studies of texts by Aleksandr Blok and Konstantin Balmont have already demonstrated that the Tk4 of Russian Symbolists can be productively interpreted as a combination of two separate hemistichs. Regarding Balmont’s verse, it has also been established that his lines corresponding to Tk4 have a visibly stricter metric structure (Polilova 2017). This structure can be described as a combination of segments of classic binary and ternary Russian accentual-syllabic meters (i.e. lines composed of disyllabic or trisyllabic feet). The anacrusis and clausula of each segment may vary freely, while the hemistichs are separated by word boundaries and syntactic breaks that may be considered as a quasi-caesura. Based on this fact, Polilova compiled a corpus of four-ictus tonic texts dated between 1890 and 1920 and carried out an analysis of rhythmic structure, not only for each line as a whole but also separately for the left and right hemistichs. The resulting data enables a new vision of the genesis of Russian strict stress-meters and the question of dolnik/taktovik separation. The Symbolist corpus demonstrates the extreme paucity of dolnik-structured lines and hemistichs, as well as a significant number (no less than half) of hemistichs written in binary meters. These established facts lead to a reappraisal of Mikhail Gasparov’s theory that taktovik originated in dolnik (which, in its turn, allegedly descended from ternary meters and has a trisyllabic basis). The structure of Balmont’s and Blok’s Tk4 points to an entirely different genesis of the Russian taktovik. It is much more reminiscent of mixed or heteromorphic meter that allows for a free interchange of iambic, trochaic, dactylic, anapestic and amphibrachic lines.

Tatiana Yanko’s “Prosody of poetic reading in comparison with prosody of everyday speech (Joseph Brodsky)” was aimed at analyzing the prosody of so-called poetic reading. It was exemplified by Joseph Brodsky’s reading of his own verse and the writings of other authors. Poetic reading was analyzed in comparison with ordinary spoken speech, where the prosody designates a variety of meanings that form sentences as speech acts and texts as connected units of discourse. The paper was illustrated throughout with frequency tracings of sound recordings (Yanko analyzed the sound data with the Praat software). Brodsky consistently followed a rhythmical manner of reading, with all accents, pauses and length variations aimed at forming the rhythmic units of verse. He ignored all linguistic meanings – topics, foci, discourse continuity – that are designated by prosody. In this manner of recitation, the obligatory linguistic meanings remain unexpressed. Nevertheless, the prosodic

---

system used by Brodsky is connected not only with the rhythmic structure of verse but is also aimed at designating some specific, non-linguistic, variety of meanings. These meanings are not random – they are organized in a system. However, they do not belong to the Russian language but to a specific strategy of recitation. Brodsky developed a highly consistent and very simple prosodic organization for the spoken poetic text based on the following markers: a rise on the tonic syllable of each phonetic word in a poetic text; an increase in the medial frequency in each minimal poetic unit; the final fall in the frequency at the end of each minimal poetic unit; and an increase in the medial frequency from one minimal poetic unit to the next minimal poetic unit in advancing towards the end of the entire text. Brodsky’s manner of reciting his own poetry forms a kind of semiotic system, which treats the spoken poetic text as a structure. Discussants conjectured that this description can serve as an example of what Roman Jakobson called delivery pattern (Jakobson 1960: 366–367). Unlike the other three elements of Jakobson’s dichotomies (verse pattern, verse instance and delivery instance), delivery pattern is rarely examined.

The third and last day of the Prague conference was also devoted to Russian verse.

Roman Leibov, Boris Orekhov and Artjoms Šela’s paper “Trochaic travels and iambic landscapes: using topic modelling to reveal semantic patterns within poetic meters” was presented by Artjoms Šela. The paper addressed an established theory in versification studies that is known as the “semantic halo of a meter”. Mainly popularized in the works of scholars of Russian poetry (Kiril Taranovsky and Mikhail Gasparov), in general form this theory states that the distribution of meanings across different metrical forms (and their variations) is non-random. For example, iambic trimeter over the course of its history would most likely retain certain semantic features, and the configuration of these features would never completely overlap with other metrical forms (Gasparov 1979; 1999: 88–119). The existence and accumulation of these metrical differences could, as it was suggested, form a “semantic valency” of a meter or its “expectations horizon” for a reader. Despite being well established, the theory of the “semantic halo” is one of the less rigorous work venues in quantitative versification studies. Tracing high-level semantic patterns across the whole tradition was a meticulous task which was hard to formalize. Scholars were able to explore more distinctive, less populated metrical forms (notably, Russian trochaic pentameter) but struggled to describe mechanisms

---

behind the appearance of the halo and the overall structure of relationships of meters in the semantic space. In their presentation the coauthors proposed an operationalization of the “semantic halo” using topic modelling on an 1800–1950s corpus of Russian poetry. They modelled each meter as aggregated topic probabilities of individual poems (composed in the corresponding meter) and then used an entropy metric to compare probability distributions. They demonstrated that, based on the vectors of semantic features, meters cluster together in a non-random fashion: this strongly suggest that the “semantic halo” theory holds true for a large-scale corpus when accounting for every metrical variation in it. Topic modelling could be used to answer the fundamental questions of the nature of the semantics/semiotics of meters and even to compare the “halo effects” across languages and national traditions.

**Olga Lyashevskaya, Ekaterina Vlasova and Kristina Litvintseva** discussed the “Lexical diversity and color hues in Russian poetry”. The paper is subtitled: “A corpus-based study of adjectives”. The coauthors described the distribution of color adjectives in Russian poetry of the Silver Age and defined individual preferences as to poetic tradition, syllable structure and metrical restrictions. Their research method combined a lexico-semantic approach, formal literary analysis and quantitative metrics obtained via the frequency database of the Russian Poetic Corpus (over 10 million words, including one million adjectives). The database allows users to create and compare subcorpora and build graphs of timeline distribution. The graphs demonstrate that the lexical diversity and relative frequencies of color adjectives start to grow rapidly in the 1890s, as the Modernists employed color adjectives to enlarge the poetic inventory. The adjectives referring to non-banal hues, including borrowed words such as *fioletovyj* ‘violet’ or *lazorevyj* ‘azur’, belong to the middle part of the ranked wordlist. A correspondence analysis of the data revealed individual color preferences and stylistic similarities among the most important poets of the Silver Age (for example, Anna Akhmatova and Aleksandr Blok are similar in respect to their use of the white hues). The distribution of the selected color hue adjectives across the metrical types highlights the strong association of multisyllabic adjectives with certain meters, although some words have a more complex distribution.

**Barry P. Scherr** presented a paper on “The Russian quintain”. This stanza occurs far less often than the quatrain (the predominant stanza in Russian verse). As a rule, quintains account for no more than several percent of a

---

Ksenia Tverianovich’s “Rhythm and syntax in Aleksandr Sumarokov’s odes” continued the topic of stanzaic rhythm, but from a different perspective. The paper was focused on the accentual rhythm and grammatical structures in the lines of iambic tetrameter in odic stanzas of Aleksandr Sumarokov, one of the most prolific and influential 18th-century Russian poets and verse theorists. The presented research was based on the concept of Russian iambic tetrameter as a stereotype, a cliché, or a formula. This concept was introduced in the 1920s by Osip Brik, who called this phenomenon “a rhythmical-syntactic figure” (Brik 1971 [1927]: 120) and later developed by Mikhail Gasparov in relation to individual lines of the meter (see, in particular, Gasparov 1984 [1982]). Tverianovich demonstrated that, in Sumarokov’s odes, not only do the individual lines of iambic tetrameter become formulaic, but so does the whole odic decima. Formulaic lines tend to occur in certain positions in stanzas. Moreover, they tend to fit the key images and concepts of Sumarokov’s solemn odes, such as the idea of absolutist state power and the greatness of the monarch.

Anastasia Kruglova, Olga Smirnova and Tatyana Skulacheva presented a paper titled “Syntax and pauses in a verse line: statistical analysis”. In regular human speech, pauses are of different lengths; this phenomenon is based on the strength of syntactic breaks between words. As it turns out, this peculiarity
of speech has special importance for the structure of a poetic line, a principle now firmly established in a number of European languages (Russian, Spanish, French, English). In their presentation, the coauthors concentrated on the regularities observed in Russian iambic tetrameter (exemplified by Aleksandr Pushkin’s *Evgeny Onegin*) and compared the data on the distribution of close and loose syntactic ties within the line and between lines, as well as the distribution of longer and shorter pauses in the reading of the same text. For acoustic analysis, the text of the second and the third chapters of *Evgeny Onegin* were read by a skilled public reader, a woman with higher non-linguistic education. The duration of the recording is 60 minutes. Pauses were measured using the Praat program. The coauthors’ analysis of this recording showed that close syntactic ties occur at the beginning (between words of the first and the second feet) and at the end of a line (between the words of the third and the fourth feet).

According to the presenters, closer ties and shorter pauses normally occur near the borders of a verse line (to form a contrast with the longest pauses and weakest syntactic ties between lines), while weak ties and long pauses within a line are concentrated in the middle of the line. This mirror-like dichotomy of close and loose connections between words at both the syntactic level and the corresponding phonetic one constantly occurs in verse and disappears in prose (various examples of accentual-syllabic meters, syllabic meters and free verse were examined). Loose syntactic ties show a mirror-like opposite distribution: they are not very numerous at the beginning, are extremely scarce near the end of a line, and reach a maximum in the middle. The same distribution has been observed in different periods of Russian (accentual-syllabic verse, dolnik, *vers libre*), English, and French verse (Skulacheva 1996).

The reason for such a distribution may be the necessity to keep words of a line together as one unit: it is much easier to tear away one word at the beginning or the end of a line than to break a line in the middle. Loose ties between lines support division into lines – the basic feature of a verse text – whereas a long pause within a line, too close to the interlinear pause would distract attention from the latter. The length of a pause depends on the type of syntactic tie. Long pauses at different positions within a line behave like the weak ties to which they correspond. Long pauses are not very numerous at the beginning, their maximum within a line is in the middle and they are very infrequent closer to the end, i.e. between words in the last two feet of a line. There is often no pause between words in the first and second feet of a fully stressed iambic tetrameter line, an absence of pause is most typical for the position between words in the last two feet of a line, and the absence of pause falls to its minimum in the middle of the line. Longer pauses are concentrated
in the middle of lines, where they can not be mistaken for the end of a line. The longest pauses appear between lines, thus supporting the division into lines. The position between lines in Russian accentual-syllabic verse is normally the position for testing the strength of a tie: weak ties reach their maximums there, and strong ties – their minimums. The longer the pause, the more often it occurs between lines.

The concluding paper of the conference, Svetlana Matyash’s “Rhythmical-syntactic formulas in the position of enjambment in Russian poetry”, was the only paper presented in Russian (and simultaneously translated into English by Tatyana Skulacheva). Matyash studied enjambments in Russian 19th- and 20th-century narrative poems written in iambic tetrameter. She investigated the same corpus that she previously employed to study enjambment in Russian poetry at large (Matyash 2017). The texts were divided into six groups (a total of 48 long poems containing 43,584 lines) according to authors and periods: 1) Zhukovsky's innovative translation of Byron's *The Prisoner of Chillon* (1822); 2) Pushkin's *Evgeny Onegin* and eleven narrative poems (1820–33); 3) nine narrative poems of Pushkin's epoch (1824–32); 4) fifteen narrative poems by Lermontov (1828–41); 5) nine poems from the mid-19th century (1845–58) and 6) two 20th-century poems: Blok's *Vozmezdie* (1911) and Tvardovsky’s *Za dal’ju – dal’* (1950–60). As contrastive material, non-four-foot-iambic texts, such as fables of the late 18th and early 19th centuries, Zhukovsky’s ballads from 1808–22 and Pushkin’s *Boris Godunov* and *Domik v Kolomne*, were used. Matyash adjusted Gasparov's concept of “rhythmical-syntactic formulas” (Gasparov 1984 [1982]) to a description of enjambments. Gasparov’s term refers to different lines by one or more authors where rhythmical structure, syntactic structure and one or more words fully coincide. To describe certain types of enjambment the presenter identified enjambments with identical final words in the line. The repertoire of formulas, their frequency and their structure were described against the background of the data on all enjambments of a particular period, author or poem. The frequency of enjambment formulas was calculated as a percentage of all enjambments and their structure was analyzed using the following features: 1) the type of enjambment (*contre-rejet*, *double-rejet*, but not *rejet* – this type was excluded from the study); 2) the type of the upper line clausula (masculine or feminine); 3) the type of the word-boundary after the formula in the subsequent line; 4) the number of words between the syntactically connected words of the upper and lower lines; 5) the type of interlinear syntactic ties and their frequency (by poem, author or period); 6) the rhyming/non-rhyming of the upper line (QAV 2019: 148). The analysis of syntactic ties was based on the hierarchy proposed by Gasparov and Skulacheva (1993). Matyash drew several conclusions about the functioning of
the enjambment formulas: 1) the repertoire of enjambment formulas is rather stable and has a tendency to broaden; 2) they occur as a result of borrowing, imitation and indicating adherence to a certain tradition; 3) their occurrence is often stimulated by the search for a rhyme, but the presence of these formulas in blank verse shows that the role of rhyme should not be overemphasized; 4) the number of formulas does not depend on the number of enjambments in a particular poem.

If the outcomes of Quantitative Approaches to Versification were to be summarized in one sentence, we could say that it was a memorable example of joint efforts of researchers from different countries who greatly contributed to the rapprochement of insulate scholarly traditions. The Prague conference demonstrated that verse theorists are willing and able to integrate into a diverse but vibrant and collaborative academic community.7

References


7 This paper is part of a research project based at Lomonosov Moscow State University and supported by Russian Science Foundation Grant № 17-18-01701. The authors are grateful to Robert Kolár, Petr Plecháč and Lydia Roberts for their help.


Quantitative Approaches to Versification: Conference report

https://doi.org/10.1515/jlt-2019-0004


