

Boris Yarkho's works on literary theory*

Mikhail Gasparov**

Boris Isaakovich Yarkho (1889–1942)¹ was a major and unique figure in literary scholarship during the 1920s and 1930s. However, his name is usually recalled less frequently than the names of many of his contemporaries. The circumstances of Yarkho's life and work prevented his most important research from being published. Therefore the general panorama of his completed work and his projects for further research escaped most of his contemporaries. And it is the width of his general intentions to perfect scientific methods in studying literature that has rendered the most services to literary scholarship.

Yarkho received his education at Moscow University and later studied in Heidelberg and Berlin. From 1915 to 1921 he taught at Moscow University, first as Privatdozent and later as a professor. Literary theory did not immediately become the main object of his scholarly activity. At first he specialized in folklore, Germanic philology, and the history of medieval literature (in this area he was a recognized authority, his articles were published in foreign journals, and the Medieval Academy of America elected him as a member). His first major work was dedicated to folklore: it was *The Tale of Sigurd and its Reflection in the Russian Epic* (published in *Russkij filologičeskij vestnik*, 1913–1916); his second long work was on skaldic poetry (*Mansongr*, published in *Sborniki Moskovskogo Merkurija*, 1917). He chose to write his doctoral dissertation on Hrotsvitha [of Gandersheim], a German poetess of the tenth century who composed dramas in Latin. The title of Yarkho's dissertation was *The Rhymed Prose of Hrotsvitha's Dramas*; he worked on it for over ten years; the dissertation was completed in two variants, Russian and German, but remained unpublished. In studying Hrotsvitha's use of poetic rhythm, which complexly oscillates between verse and prose, Yarkho had to work out detailed

* This paper was first published in Russian (Gasparov 1969; a revised version: Gasparov 1997). Translation by Michael Lavery and Marina Tarlinskaja.

** Mikhail Leonovich Gasparov (1935–2005) was a preeminent scholar of Russian and European verse. Among his works translated into English are Gasparov 1972; 1980; 1987; 1996. A special issue of the *Slavic and East European Journal* (2008, vol. 52, no. 2) was published as a tribute to Gasparov.

¹ Russian: Борис Исаакович Ярхо. Other transliterations: Iarkho, Jarcho, Jarxo, etc. — *Eds.*

methods of statistical approach to the phonetic layer of literary speech with its immensely varied phonic features. Here Yarkho worked out the foundations of what he called the “methodology of an exact study of literature”. The application of statistical methods in the study of versification, as we know, was traditional already in nineteenth-century classical philology, familiar to Yarkho. The novelty of Yarkho’s approach lay in the fact that he was probably the first to transpose the use of these methods onto other areas of literary studies.

Yarkho wrote his main series of foundational theoretical works during the 1920s and 1930s. From 1922 to 1930 he worked in GAKhN [Gosudarstvennaja akademija khudozhestvennykh nauk: the State Academy of Artistic Sciences], where he headed the department of theoretical poetics and the committee for translation of literature. Here he managed to organize a small scholarly group (Mikhail P. Shtokmar, Leonid I. Timofeev, Igor’ K. Romanovich and others) whose work from then on he always remembered with deep satisfaction. GAKhN was something like a club for the Moscow humanitarian intelligentsia; its members were forced to make a living on the side. Yarkho taught languages and stylistics in the educational institutions of that period, whose names were constantly changing. He worked for BSE [Bol’shaja Sovetskaja Entsiklopedija: the Great Soviet Encyclopaedia] and for three years served as an economist in the VSNKh [Vysshij sovet narodnogo khozjajstva: the Supreme Soviet of the People’s Economy]; he translated a great deal (*La Chanson de Roland*, the *Völsunga saga*, Goethe’s *Reineke Fuchs*, the plays of Molière and Schiller; his translation of *El Cantar de mio Cid* was published posthumously with annoying changes and without an introductory article; his translated anthologies of Carolingian poets and of “visions” of the sixth to twelfth centuries, and his translation of the *Cent nouvelles nouvelles* to this day remain unpublished).² During this busy period he also wrote a series of articles on the rhythm of Russian rhymed prose of the seventeenth and eighteenth centuries (partially published) and the unpublished works “The distribution of speech in five-act tragedies”, “The comedies and tragedies of Corneille”, and “The correlation of forms in the Russian chastushka”. He also wrote a small study on the history of the adaptations of *La Chanson de Roland* (“Young Roland”, Leningrad, 1926) that adjoins the works listed above. His work on these themes allowed him to systematize and generalize, first, the aesthetic prerequisites for the formal study of literature, and second, the specific devices for the exact methods of

² Yarkho’s anthology of Carolingian poets was eventually published in 2010. His preface to his anthology of medieval “visions” was published in 1989. — *Eds.*

such research. He achieved the former in the articles “The elementary foundations of formal analysis” (Yarkho 1927) and “The limits of the scientific study of literature” (Yarkho 1925–27); these articles deserve much greater renown than they currently enjoy. The latter was achieved in the book *The Methodology of an Exact Study of Literature (Introduction. Analysis. Synchronic synthesis. Diachronic synthesis)*, a large study (over 400 pages), summing up all that the author had accomplished and planned during the past 20 years of work. GAKhN was scandalously abolished in 1930; Yarkho was arrested in 1935 (in the “case of the Great German-Russian Dictionary” that entrapped a large group of the Moscow intelligentsia); he wrote his book in exile in Omsk and continued to work on additions to it until his death. The book remained unpublished; only short excerpts appeared in the periodicals *Sign Systems Studies* IV (Tartu, 1969)³ and *Context-1983* (Moscow, 1984), and it has not in any way become obsolete in the 60-plus years since it was written.⁴ The unfinished large studies of the poetics of the *Slovo o polku Igoreve (The Lay of Igor’s Campaign)* and contemporary Western European epics were to serve as a practical demonstration of the developed methodology. Yarkho worked on this study in Kursk, where he was a professor from 1940 to 1941, and in evacuation in Sarapul, where he passed away from tuberculosis on May 3, 1942.

The basic postulates of Yarkho’s methodological system laid out in these works can be shortly summarized in the following theses.

Literature is an independent phenomenon of objective reality and should be studied as an independent object of research rather than as a reflection or expression of any non-literary phenomena (social relations, psychic complexes, etc.). As is well known, trees, for example, may be studied as a reflection or expression of something extrinsic, for instance, of Christian virtues, and all of medieval botany was based on this presumption, but it did not move science one step forward.

As a phenomenon of objective reality, literature should be studied by scientific methods, such as are applied to objective reality – the methods of exact sciences. Of course, the level of precision intrinsic to mathematical sciences is not attainable in the study of literature, but some level of precision, achievable, for instance, in the natural sciences is totally accessible to literary research. The exact sciences know two main methods of research – observation and experimentation. Clearly the use of experiments in literary studies is difficult

³ This excerpt was translated into English (Yarkho 1977 [1936]). — *Eds.*

⁴ The book was finally published in 2006. All references to the *Methodology* in this article have been transposed to this edition and formatted according to this journal’s specifications. — *Eds.*

(though not impossible). Therefore observation remains the main method of literary research. Its success depends on the solution of two questions: "what to observe" and "how to observe".

The natural answer to the first question is as follows: we observe what is specific to a given object. The specifics of literature are its aesthetic effectiveness, that is, its power to invoke the feeling in a person (for lack of a more precise definition) that forces him to approach the object from the point of view "is it beautiful or not beautiful?" The aggregate of aesthetically effective elements of a literary work is called its literary form. The aesthetic effectiveness of each element of a literary work arises from its uncommonness, either quantitative or structural. The sound "p" in language is common and therefore aesthetically neutral, while the accumulation of this sound in Aleksandr Pushkin's line "*Pora, pero pokoja prosit*" is quantitatively uncommon, and therefore aesthetically potent. It enters a literary scholar's field of conscience as a phenomenon known as "alliteration". Stressed and unstressed syllables in language are common, but their regular alternation in the same line is structurally unusual; it is recorded in the consciousness of a literary scholar as a phenomenon known as "iambic tetrameter". Uncommon sound forms (that is, those that attract our hearing), similar to the examples above, comprise the field of *phonics*; uncommon linguistic forms (that is, those that affect our thinking) belong to the field of *stylistics*; uncommon images, motifs and plots (that is, forms that affect our imagination) are called the field of *iconics* or *poetics per se*. These are the three main fields of literary research; to these three Yarkho adds a fourth – *composition*, the study of their interaction. Each of these areas, independently or together, can be studied either synchronically or diachronically.

The question of "*how to observe*" is answered thus: start with an immediate impression, check it against an objective account of all features capable of producing this impression, and express the results of the observation in the form of quantitative indices. Only in this way can the results be considered trustworthy. Prior to Yarkho, literary scholars had limited themselves to the first of these three stages and therefore could not get themselves out of the abyss of subjectivism. Relying on the immediate impression they operated with such concepts as a "flowery style", "colourful", "lively", etc. However, by not worrying about an objective account of these features, they had been unable to fill these concepts with a universally recognized meaning: what seemed flowery to one person didn't seem flowery to another, etc. However, if we agree beforehand, say, to give the name "flowery" to a style saturated with morphological and syntactic figures, and call "colourful" a style saturated with sensuously characterized images, etc., then all these disagreements will end:

it will be sufficient to count the ratio of figures in this or that text to be able to say which of them is more flowery and which is less flowery, and to what degree. Thus, an expression of quality via quantity is the primary goal of any literary study if it wants to be a science and not a game of subjective tastes.

To reiterate: we are not talking about driving intuition out of scholarly cognition, but only about driving it out of scientific exposition. In the very first pages of his monograph, Yarkho writes:

Science comes from a need for knowledge, and its (basic and primary) goal is the satisfaction of this need. [...] This aforementioned need is as inherent in man as the need to reproduce: if a man leaves this need unsatisfied, he will not physically perish but sometimes will suffer extremely intensely. [...] An intellectual man is not a person who knows a great deal, but rather merely one who possesses an above-average thirst for knowledge. [...] The need for knowledge, however, is only the grandmother of science. Its mother is the need for communication of knowledge. If curiosity is the primary biological characteristic of the human individual, then “thirst for communication” is the secondary quality of man, already as a *zoon politikon*: in other words, the social aspect comes into force. And science, the daughter of this need, is first and foremost a *social act*. Indeed, there is no scientific (in contrast to unscientific) knowledge *per se*: when we discover the most reliable scientific principles, our intuition, fantasy and emotional tone play an enormous role alongside our intellectual processes. Science is a *rationalized statement of what has been cognized*, a logically formulated description of that part of the world which we have managed to understand; this means that science is a particular form of communication (exposition) rather than of cognition. The task of this exposition is to find a general, *objective language*, because logic is considered the most homogenous psychic function in people. Alongside with logic, sensations also have a high level of homogeneity: most people see red as red, wormwood tastes bitter to most people. Therefore *sensory demonstration* and *logical proof* are the basis of objective language. [...] In my view, science is a system of proof, a rationalized language, and I dreamed of converting my science, the science of literature, into precisely such a system. (Yarkho 2006: 19–20, 21)

Thus, statistics for Yarkho were by no means a goal in themselves, but only an instrument to discipline thinking, which the science of literature sorely needs.

I ought to say I found especially annoying that we, scholars of literature, practice this *saloperie* in treating the conclusions made from sometimes *extremely carefully collected material*. I have just taken a closer look out of the corner

of my eye at meteorology. What a contrast! What precision in processing the most approximate, arbitrary and superficial observations! What panache in the diverse methods of recording and accounting data! [...] If we exercised one tenth of this diligence and spent one hundredth of these means, our science would be able to compete with any other (Yarkho 2006: 30). — Placing quantitative data and microanalysis at the base of my research, I only suggest that we do for the science of literature what Lavoisier did for chemistry a century and a half ago, and I have no doubt that the results will be imminent (Yarkho 2006: 7).

The quantification of aesthetic impressions constitutes the pathos of all theoretical literary work by Yarkho. Here is a small but striking example. Among the scholar's papers (Russian State Archive of Literature and Art [RGALI], fond 2186, op. 1, ed. khr. 83) a page is preserved with the title "The programme for the study of syntactic boundaries between verse lines". Yarkho describes in passing the history of this research programme in his *Methodology* (Yarkho 2006: 34). He happened to debate with "a Leningrad professor"⁵ about "normal diction": which pause is stronger in Lermontov's verse line, "*Beleet – parus odinokij*" ["Gleams – a sail lonely"] or "*Beleet parus – odinokij*" [Gleams a sail – lonely]? Yarkho maintained that the first break was stronger, while his opponent felt that it was the second. Such debates are usually resolved without arguments. Yarkho began to argue. The first pause in the line in question separates the subject and the predicate, the second pause – a modified noun and its modifier. It is obvious that the stronger of the two syntactic pauses will be the one that most often coincides with the strongest rhythmic pause of the verse, the final pause between the lines. Yarkho compiled a list of all types of syntactical pauses possible at the end of a verse line; there turned out to be five such types (ten with subtypes included): 1) between sentences, 2) between a subject and a predicate, 3) between a verb and its object or adverbial modifier, 4) between a noun and an attribute or appositive, 5) other instances – forms of address, auxiliary particles. And then he calculated how often each of these five pauses coincided with the end of a verse line. He used as material the first hundred lines from five texts: A. K. Tolstoy's "The dragon", Pushkin's "The little house in Kolomna", Vladimir Benediktov's translation of [Schiller's] "The gods of Greece" (all in iambic pentameter), Pushkin's "Ruslan and Liudmila" and "The Bronze Horseman" (both in iambic tetrameter). The results are displayed in the table below (all numbers in percentages):

⁵ Lev V. Shcherba, a prominent linguist who worked in Leningrad. — *Eds.*

Table 1

Type	The dragon	The little house in Kolomna	The gods of Greece	Ruslan and Liudmila	The Bronze Horseman
1	52	59	67	60	39
2	25	14	18	22	17
3	18	25	13	15	36
4	2	1	2	2	8
5	3	1	—	—	—

Thus we have the means to answer the original question: the type “*Vot uzh na more beleet / Parus odinokij*” [predicate phrase / subject phrase] is encountered on average six times more often than the type “*Vot uzhe beleet parus / Odinokij v sineve...*” [predicate phrase, subject / subject’s modifier]. In proving his point Yarkho obtained for the first time numeric data to help address such an important problem as enjambment (cf. the data from “The Bronze Horseman” and other texts). These numbers required further precision and examination, and the page records a plan of “excisions” for the subsequent work: how the distribution of the types of end pauses changes according to the author, metre, stanza type, genre... Thus a fortuitous debate about declamation grew into a conceived plan of how to investigate a large and important subject: rhythm and syntax. In 1929 Yarkho planned to carry out such a study with the help of his group at GAKhN, but in 1930 GAKhN was dissolved and the project did not materialize.

Yarkho planned his research in such a way as to gradually test his method in all fields of literary studies.

In the field of *phonics* this allowed him to study in depth a number of complex forms intermediary between verse and prose, both in Russian and foreign material. It wouldn’t be necessary to focus on the details of these works in our current overview; firstly, the majority of them have been published (“The rhymed prose of the so-called *Novel in Verse*” and “Pushkin’s free sound forms” in the collection *Ars poetica* 2, 1928; “The mystery play of the ten virgins” in the collection *Pamiati P. N. Sakulina*, Moscow, 1931; “The rhymed prose of Russian intermezzos and interludes” in the collection *Teorija stikha*, Leningrad, 1968), and secondly, because in the field of versification a successful application of statistic methods (one of the pioneers of which was Yarkho, who started his work independently not only of Boris Tomashevsky, but also of Andrei Belyi) had long ago ceased to be a novelty.

In the field of *stylistics*: Yarkho’s most effective illustration of how to apply the new method to stylistics can be found in the introductory article to his

translation of the poem *La Chanson de Roland* where he described his analysis of the poem's stylistic devices (Moscow, 1934). The common opinion had been that the style of *La Chanson de Roland* was impoverished in verbal embellishments, that it is dry, simple, and terse. A calculation showed that both points of view were wrong; the number of verbal ornaments, that is, stylistic figures, comprises about 20% of the number of lines, which is no insignificant amount (in *El Cantar de mio Cid* – only about 11%); among these figures the majority (60–65%) is pleonastic; thus, the style of *La Chanson* is not dry but, on the contrary, intentionally verbose. The traditional view of the scantiness and dryness goes back to the judgments of great scholars, whose “feel for language” was undoubtedly no less acute than Yarkho's. This example demonstrates that a victory in the controversy was not won by the superiority of the scholar's talent, but rather by the superiority of his method – the statistical over the intuitive.

In the field of *iconics* on its lowest levels – *image, motif, plot* – Yarkho worked relatively little. One can probably mention the expressive comparison of military motifs in *La Chanson de Roland* and *El Cantar de mio Cid*: although both poems are nearly identical in length, in *Roland* we find 76 battle scenes (of which 46 are single combat), while in *Cid* – 25 battle scenes (of which only 5 are single combat): the plot dynamics of *Cid* is achieved not by the battle scenes, but by changes of location (in *Roland* the place of action changes only 18 times, while in *Cid* – a countless number of times). One can also note the intriguing analysis of sensuous colouring of images: the statistics of colors in five medieval epics. In *Cid* only three colours are used, and colour epithets comprise only 0.024% of the text (calculated from the number of syllables); *Beowulf*, correspondingly, has seven colours and 0.053% of the text; *Slovo o polku Igoreve* – nine colours and 0.433%; *Roland* – ten colours and 0.217%; the *Nibelungenlied* – eleven colours and 0.065%. The growth in the richness of the palette from the semi-barbarian *Beowulf* to the courtly *Nibelungenlied* could have perhaps been predicted a priori (although *Cid*'s placement here too is unexpected). However, even the most refined intuition could hardly have been able to predict that the author of *Slovo o polku Igoreve* used his nine colours six times more generously than the author of the *Nibelungenlied*'s eleven colours.

In the field of *iconics* on its highest levels, however, Yarkho worked a great deal, and he achieved great success in the difficult task of formalizing the main *emotional and ideological concepts* embedded in a literary work. He accomplished this feat by strictly isolating the object of his research. The emotional concepts of a work (tragic elements, optimism, etc.) derive only from the statements contained in the text, moreover, from the author's speech, and by no means from the researcher's own interpretations. The “rules” for

identifying the ideological concept of a literary work formulated by Yarkho in his *Methodology* deserve to be cited in full:

It is important for me to show that it would be possible to set the work of finding the dominating idea or emotion of a literary work in a way different from how it is being done now. Usually, in the best case, the “researcher” takes some statement from his object at random, subjects it to an arbitrary exegesis, adjusts it to several important passages from the story, and then admires how it “all” (!), in a coherent whole, is “subservient to one idea”. If such a hasty intuitivist were to carry out an analysis that he despises so much (people simply do not like to work), then he would see what this “all” comes down to. But this, as mentioned, is still the best case: the “scholar” often himself thinks up an idea and attaches it to the a literary work, and then already without any hesitation declares it dominant, finding the metre and meaning superbly in agreement with it (in what manner?) etc., etc.

I propose to bring this activity into a stricter framework *in order to explain some genuine, and not fictitious, relations*.

α. Prerequisites:

1. An idea is not necessarily a fundamental (dominant) feature in the general structure of a complex.
2. An idea may be considered a “concept” if it surpasses in verbal quantity all other ideas of the same complex in the image material it encompasses.
3. An idea is dominant if it encompasses over 50% of all verbal material (expressed in numbers of any kind of volume denominator).⁶
4. A concept may be composite, ideological-emotional, i.e. one and the same material can express both an idea and an emotion. [...]
5. An idea should be articulated *expressis verbis* in the text of a work: only then can one say that the idea is present in the text. To derive it through arbitrary exegesis is a fruitless endeavour.

β. Method:

1. Proceed to a statistical weighting of a conception only after a sufficiently detailed analysis of the poetics of a given work (iconics, motifs, plot).

⁶ A volume denominator is any unit of measuring the text volume: line, word, syllable, etc. (Yarkho 2006: 118–119). — *Eds.*

2. One must evaluate not just a single idea, but all ideas clearly formulated in a given text. If an idea is clearly episodic, i.e. connected with an insignificant volume of verbal material (a proverb cited in passing, the opinion of a secondary character never later reaffirmed, and the like), then it may be disregarded.
3. One must take an idea only in the expressions in which it is given by the author. An idea must never in any case be paraphrased at random; but it may be complemented with a different variant contained in the same text. By no means should one ever transpose an idea from one work of the same author to another without mention.
4. Take an idea only in its context. If, for example, an idea is expressed by an antagonist and is constantly refuted by the course of action and the remarks of other characters, then obviously it is the negation of this idea and not the idea itself that is dominant.
5. In order to weight an idea, it is necessary to count words and indivisible phrases relating to it [...].
6. Motifs (i.e. [verbs and] verbal expressions) that are not relevant to the plot but support an idea are included in the general set of topics as nouns and adjectives; a verb with its object or adverbial modifier is counted as one unit. For example, units supporting the idea "wealth is the highest good" could be: "...he amassed wealth..."; "...managed to become wealthy"; "...sold profitably".
7. Motifs relevant to the plot are counted separately. It may turn out that the images mostly relate to one concept (to a panegyric of military valor, for example), while the course of action relates to another (for example, to the affirmation of the power of fate); in this case, the synthesis will contain two dominants: one for ideology, the other for the subject-matter.
8. The number of the units obtained must be weighted with the help of some volume denominator. (Yarkho 2006: 122–124)

Yarkho's analysis of the ideological conception of *La Chanson de Roland* may serve as an illustrative example of such an approach. At the time of his scholarly activity there were two theories about the *Chanson*: the old theory maintained that the *Chanson* was created in a military environment, while a newer theory affirmed that it was composed in a clerical environment. Yarkho adhered to the old theory but was the first to affirm it with statistical data, comparing the *Chanson* with a later reworking of the story composed by a clerical author (Priest Konrad's *Rolandslied*). It turned out that Christian motifs (i.e. all lines from which it was clear that the poem's author was a Christian) in the

Chanson comprise about 10% of the text, while nearly twice this proportion of such verse lines (about 20%) occur in the clerical version. Thus, the ideology of *Roland* is mainly secular and knightly. Moreover, with the help of statistics it is possible to specify both the nuances and the concentration of this ideology. If we single out only three of its elements – “courage”, “military honor” and “patriotism”, then the vocabulary related to these elements in *Roland* composes 0.29%, 0.24% and 0.12% of the text (by the number of syllables); in *Cid* – 0.14%, 0.07% and 0%; in *Beowulf* – 0.48%, 0.48% and 0.08%; in *Slovo o polku Igoreve* – 0.82%, 0.82% and 0.40%. From this it is clear that, first of all, in concentration of military ideology *Roland* surpasses *Cid* but is far behind *Slovo o polku Igoreve*, and, secondly, in patriotic coloration all three western epics cannot even come close to *Slovo o polku Igoreve*, “This is the mathematical expression of ideological specificity”, Yarkho concludes (RGALI, fond 2186, op. 1, ed. khr. 49, fol. 10).

Thus, by applying exact methods of research, we have won the right to speak of “an emotion that permeates the entire poem”, of “an idea that directs all the characters’ actions” and other such high words that I did not at all plan to eliminate from the study of literature but into which I merely wanted to insert a concrete meaning. Besides, such words are often nothing but hot air, and if the reader becomes accustomed to demanding numerical evidence, then our desire to throw around pompous phrases will gradually pass away. (Yarkho 2006: 48)⁷

Everything discussed above concerned separate fields within literary studies. The question of the *interaction* between these fields was, understandably, much harder to subject to exact analysis. Nevertheless, the first steps were taken in this direction as well. In his article “The correlation of forms in the Russian *chastushka*” (printed in German in 1935) Yarkho studied the style of the *chastushka*⁸, to be more precise, the stylistic device of repetition in the *chastushki*

⁷ At the beginning of the war in 1941 a wave of conferences on patriotic themes took place in Moscow scholarly institutions. Yarkho volunteered to give a talk on the patriotism of *Slovo o polku Igoreve*, presented the results mentioned here and said that he acknowledges his mistakes: until then he had counted the percentage of patriotic vocabulary by number of words; now he counts by number of syllables instead and sees that the indicator of patriotism of the *Slovo* is even higher. This presentation caused a storm of indignation: the motherland was dear to the scholars in attendance, but their methodological innocence was even dearer. (From the recollections of Boris V. Gornung and Mikhail P. Shtokmar).

⁸ A *chastushka* is a humorous folk song, a trochaic tetrameter quatrain usually rhymed abcb or abab that is sung at village and inner city youth gatherings. — *Trans.*

in connection with their phonic and thematic sides. It turned out that the rhyming lines of the chastushka were less frequently combined with device of repetition than non-rhyming lines, and that chastushki with socio-political themes were poorer in repetitions than those with a love theme: this is the law of compensation; phonic and thematic wealth are, as it were, reimbursed with stylistic poverty. In his article on the Serbian *tužbalica*⁹ of long lines (printed in Serbian in the journal *Slavia*, 1924), Yarkho wondered if the repetition in these songs generates alliteration, or whether alliteration generates repetition. A statistical count showed that repetition without alliteration does not exist, while alliteration without repetition exists in 70% of cases, so therefore alliteration must be considered primary. (However, this is not a causative relationship, as Yarkho pedantically insists, because, for example, in Old Germanic verse the same kind of alliterative phonics did not generate the style of repetition, but rather the style of synonyms and metaphors.)

If one does not go into the difficult question of the interaction between different elements of form but only speaks of their *coexistence*, then huge possibilities open up before the statistical method. After all, the concepts of literary genre and literary trend are nothing other than a complex of varied but coexisting features.

The poetics of *genre* got a detailed study in Yarkho's work "The comedies and tragedies of Corneille" (Yarkho 2006: 403–549). It is generally believed that comedy differs from tragedy in its greater liveliness, greater saturation of action, more ordinary feelings and thoughts of its characters, etc. "Greater" – of course, but how much greater? In order to answer this question, one must fill each of these concepts with concrete content. "Liveliness" – this is, obviously, the frequency of exchanged utterances (cues); it follows that the indicator of liveliness will be the ratio of the number of utterances to the number of verse lines in a play. Indeed, in Corneille's tragedies this ratio equals 0.150, and in his comedies – 0.276. This should be distinguished from the level of coherence – the degree of necessity with which a remark causes the following one. The indicator of phonetic coherence will be ratio of utterances breaking in the middle of a verse line or couplet to the total number of utterances (in tragedies – 67%, in comedies – 77%). Similarly, an indicator of stylistic coherence could become the ratio of the number of utterances connected by a question and answer, anaphora, antitheses and the like to the general number of utterances. Saturation of action is the percent of personages that act in their own

⁹ A *tužbalica* is an improvised folk dirge, usually in octosyllabics, that is sung as part of ritual mourning practices in South Slavic cultures. — *Trans.*

interests (in tragedies – 37%, in comedies – 54%: heroic characters in tragedies usually entrust their actions to others). Saturation of action is also the number of “moments of action” (a concept introduced by Yarkho) per character (in tragedies – 2.6, in comedies – 4.2); it is the number of physical actions (killings, embraces, etc.) indicated in dialogue or stage directions per play (in comedies this indicator is 1.9 times higher). Saturation of action is also the ratio of the amount of pre-plot exposition to the general length of the play (in comedies this indicator is 2.7 times higher: the intrigue-motivating point [*zavjazka*] occurs sooner). It is, in addition, the ratio of the number of actions on stage to the number of actions that occur behind-the-scenes (in comedies, which do not make use of messengers’ speeches, this indicator is 16 times higher!). The difference in the characteristics of feelings is clear, firstly, from the fact that the number of characters who act out of fraud and deception is 49% in comedies and 12% in tragedies, the number of characters acting out of jealousy is 25% in comedies and 10% in tragedies, those acting out of bravery – 13% in comedies and 32% in tragedies, out of patriotism – 0% and 15% respectively; and, secondly, from the fact that the vocabulary of love, joy, and cunning and the vocabulary of fear, woe, valour, and hatred (about 160 words, about 4000 cases of usage) appear in tragedies at a ratio of 31:69 and in the comedies at a ratio of 64:36 (the indicator of transgression in the former is 15% and in the latter – 39%, i.e. the comic character more often acquires a tragic colouring than the other way around). And, finally, the difference in the kinds of thinking in tragedies and comedies follows, for instance, from the themes of utterances pronounced by the characters in a play: such themes as state and society, blood lineage, woe and fear, and references to time comprise 61% of utterances in tragedies, but are totally absent from comedies; such themes as love and marriage, religion, crime, valor and fate are present in both tragedies and comedies, but they comprise 39% of all utterances in the former and 78% in the latter; finally, such themes as joy and happiness, deception and lies, wealth and poverty, literature and art are present only in comedies and comprise 22% of all utterances.

At first glance these data may appear trivial: even without any calculations it is well known that tragedy is characterized by loftiness, while comedy – by liveliness, etc. But if we recall just how many dramatic works in world literature are located in between typical tragedies and typical comedies, and how important it is to establish their gravitation to a certain genre, then it will become clear how necessary it is here to have firm “reference points” – the characteristics of genres in their purest form. To be sure, the statistical data mentioned above characterize the genre system of only one epoch: classicism. But nothing prevents us from deriving such characteristics for other epochs,

and this will already be a transition towards the next problem – the poetics of literary trends.

The poetics of *literary trends* is dealt with in Yarkho's work "The distribution of speech in five-act tragedies" (Yarkho 2006: 550–607). This is a reconnaissance effort along a narrow front: out of the thirty features by which tragedies and comedies were compared, only four have been here selected: *A* – the percentage of scenes with 1, 2, 3... speakers (monologues, dialogues, dialogues...); *B* – the uniformity in the distribution of this percentage in the plays of each period (sigma); *C* – the average number of scenes in a play (the coefficient of mobility of action); *D* – the number of characters in a play. The material was limited to four periods: 17th century classicism (24 tragedies of the two Corneilles, Racine, Quinault and others), 18th century classicism (80 tragedies by Voltaire, Crébillon, Sumarokov, Lessing, Alfieri and others), early romanticism (15 tragedies by Kleist, Werner, Schiller), late romanticism (22 tragedies of Grillparzer, Kerner, Hugo, Vigny and others); 45 more tragedies from Euripides to Byron and Hebbel were added as comparative material. The following data were obtained (see Table 2; percentages for the most commonly found type of scene – dialogues – were chosen for the feature *A*):

Table 2 (Yarkho 2006: 589)

Feature	17 th c. classicism	18 th c. classicism	Early romanticism	Late romanticism
<i>A</i>	73.7	55.2	33.8	43.0
<i>B</i>	±0.61	±0.88	±1.95	±1.33
<i>C</i>	29.5	29.6	73.9	55.1
<i>D</i>	8.3	7.5	24.6	16.7

Table 3 (Yarkho 2006: 599)

Feature	17 th c. classicism	18 th c. classicism	Early romanticism	Late romanticism
<i>A</i>	0	46.3	100	76.9
<i>B</i>	0	20.1	100	53.7
<i>C</i>	0	0.2	100	57.6
<i>D</i>	4.6	0	100	54.1
<i>Mean indicator</i>	1.1	16.6	100	60.6

Before us is the law of wave-like evolution: the maximal constraint of the classicists is replaced by the maximal freedom of the early romantics, while the late romantics begin a return to strictness. If before the start of the curve

of the consolidated (mean) indicator (the concept of “consolidated indicator” is dealt with below) we add an indicator of a preceding period, 18.2 (early Corneille), and after its end an indicator of the following period, 93.0 (neo-romanticism of Hebbel, Grabbe, Immermann), then the wave-like character of the curve becomes even clearer. (“Know the changing tides that rule the lives of men” – Yarkho used this line by Archilochus as the epigraph to his work.)¹⁰ The observation of the wave in the change of literary epochs (Gothic – Renaissance – baroque – classicism – romanticism – realism) is far from new, but Yarkho was the first to propose a numerical expression (even of only one parameter) of this wave. Statistics allow us to make even more subtle observations. Thus the average distribution of scenes with 1–2–3–4–5 speakers in 17th century classicism is 10–70–16–3–1; for Corneille in particular: 8–67–20–4–1, and for Racine: 13–73–11–2–1. In other words, both classicists in equal measure deviate from the average amounts but in opposite ways: Racine prefers monologues and dialogues, Corneille – trialogues and tetralogues. The similarities between the writers are clearer to us than the differences between them, but for their contemporaries, who readily juxtaposed and contrasted Corneille with Racine, the differences were more obvious than their similarities.

The problem of the change of literary trends naturally leads to the general problem of *literary dynamics*. Here, too, Yarkho proposes the path of quantification. How do we express in numbers the overall pace of changes from epoch to epoch, from classicism to romanticism? In order to answer this it is necessary to derive consolidated (mean) indicators for all features under analysis, regardless of their heterogeneity (as it is impossible to prove, for example, that the percentage of dialogues is a feature more noticeable or, alternatively, less noticeable than the number of scenes). This can be accomplished in the following way (Table 3): the minimal indicator of each feature is designated by 0, the maximal indicator is designated by 100, the intermediate indicators are designated by percentages within this range, and the arithmetic mean is derived from the percentages of indicators of each feature inside a given period. The index reveals the place of this period on the evolutionary path. And after that it is already easy to define the historical place of any given writer within his epoch. If, say, a certain author at the end of the 18th century gives the consolidated indicator 70, then it is obvious that he is closer to the early romantics than the late classicists, and more precisely, from the distance between their indicators (16.6 and 100) he has covered about 64%. One may call this amount

¹⁰ Yarkho inaccurately cites V. Veresaev’s Russian translation of Archilochus. A recent authoritative English translation of Archilochus’ fragment 128 yields: “Know what sort of pattern governs mankind” (Archilochus et al. 1999: 167). — *Trans.*

the coefficient of progressiveness of a given writer – that is, the indicator of his gravitation to newer forms rather than those that are becoming obsolete. It will be even easier to quantify an enrichment or impoverishment of the metric repertoire of a certain period (for example, from Pushkin to Lermontov), the repertoire of stylistic figures, a set of images and motifs used (for example, from early to late adaptations of the plot of *La Chanson de Roland*), etc. In the process of identifying the sources of a literary work, one should first isolate the percentage of metrical, stylistic and iconic forms that have a correspondence in each proposed literary source; only afterwards one can attribute the residual elements to non-literary sources originated in everyday life. It should be axiomatic for the scholar of literature that literary sources are primary and non-literary sources are secondary. To adhere to the opposite order (as Yarkho explains in his favourite biological comparison) would be the same as to affirm that man comes from milk, porridge, bread and beef: of course, without all of this man cannot live, just as literature cannot live (without declining into epigonism) without an inflow of material from everyday life, yet a human being arises all the same not from food, but from human reproductive cells. (We will remember, by the way, that in his main research on literary genesis in “Young Roland” Yarkho proves, contrary to tradition, that the tale has a non-literary, rather than literary, origin.)

In mentioning the “coefficient of progressiveness”, we have touched upon the last problem of the traditional study of literature – the problem of *evaluation*. For Yarkho this problem in general does not enter into the field of scientific cognition and stays within the realm of criticism. If [the critic] Vissarion Belinsky considered Pushkin a good writer and [another critic], Dmitry Pisarev, considered him a bad writer, then these judgments help us very much in understanding Belinsky and Pisarev and very little in understanding Pushkin. One may also classify trees as beautiful and non-beautiful, but how does this help botanists? However, half-joking and half-serious, he allowed that someday it might be possible to quantify even the concept of talent. The concept of talent can be broken down into the concept of the richness of the forms used (metrical, stylistic, iconic) and that of their originality. The former is already accessible to statistics now, while the second will be quantitatively expressed only after we have compiled for each epoch a frequency catalog of forms and will be able to evaluate the originality of each form (the rhyme “liubov’ – krov” [love-blood], the metaphor of “love as flame”, the idea that “love is stronger than death” and the like) as a quantity inverse to its frequency, and there you have it.

Yarkho’s goal in the study of literature was not evaluation, not an answer to the question “is it good or bad?” but rather an answer to the question “what

and how?” i.e., an exhausting summary of statistically derived information about the repertoire of poetic forms of all times and peoples. He planned works of the widest scale – for example, a series of reference books on the metrics, stylistics and iconics of Russian poetry (a metrical reference book on Pushkin was published, a reference book on Lermontov was prepared for print) as well as a history of the filiation and migration of love topics in European poetry from the early Latin medieval period to the high Renaissance. Against the objection that one man alone could not handle all these themes, he answered that the time of researchers working alone in all sciences had already passed, and that for a small research group such a theme would be a task for a year or two, and no more. But the conditions of scientific work in the 1930s turned out such that these plans remained unrealized.

Yarkho named his method “formal”. But unlike the OPOYAZ Formalists, he considered this method not a revolution in science, but on the contrary, a direct development of positivist methodology of the 19th century. He polemically opposed contemporary methodologies to his own methodology, the goal of which was the search for objective truth rather than replacing it with subjective conviction or up-to-the-minute utility. Accordingly, it became necessary to wage this polemic on two fronts: first, against the “organic poetics” of intuitivism (defended in GAKhN, for example, by Gustav G. Shpet and Mikhail P. Stoliarov) and, secondly, against the sociological literary studies of the 1920s and 30s (Petr Kogan, Vladimir Friche and others). Let us bring to mind a curious characteristic of the “organic poetics” that has not lost its relevance to this day:

It has often occurred that I hear from the mouths of people who firmly stand on the point of understanding the study of literature as “blabology” accessible to everyone, that analysis is unlawful, that it’s impossible to split the “living organism” of a poetic work into parts, to break up the whole which is an indivisible product of inspiration etc. etc. They propose to “proceed from the whole”, to define immediately (intuitively, that is, simply speaking, without any expenditure of effort or time) the “essence”, “core”, “dominant”... [...] But every time such a deductivist gives his definition of “essence”, he only snatches out one feature out of many and arbitrarily considers it predominant, for he has absolutely no means of proving the dominant character of this feature: not having analyzed the complex, he does not know its other features and he cannot judge to what degree they are more important or immaterial than the characteristic he has ripped out. In addition to this, in abstracting one feature, he himself is making an analysis, but he does this incompletely and poorly. (Yarkho 2006: 66–67)

Yarkho reproached sociological poetics for its constant striving to directly explain the better known by means of the less known – explaining literature by means of social consciousness. The path of research should be the opposite: from a particular literary work through a particular writer, through the actual cultural environment of people who knew each other (troubadours, skalds, the court of Louis XIV, Pushkin's St. Petersburg – something that Yarkho calls “literary ecology”) to the broader and less defined cultural environment, and while going through this process one should not forget that the cultural environment and class environment are not identical by far. Of course, a writer writes with a reader in mind, but their relation is not a “social demand” [or “social order”], but a “social market” [or “social sale”]; the reading public is not a client, directly dictating to the master all characteristics of a requested product, but rather a free consumer, who more often than not is attracted to the shop window exactly by the unexpectedness and unusualness of the goods. Literary works are indeed sometimes made to order, but this is usually “newspaper-template babble”; to “order” Pushkin's *Eugene Onegin*, the client should himself have the talent of Pushkin.¹¹

The positivist prerequisites of Yarkho's methodology fully explain his weak sides, which can be seen clearly enough from the aforementioned: his descriptiveness, mechanicism, and biologism. Yarkho sees in a literary work first and foremost a sum of atomistic devices, an aggregate of elements of form independent from one another. He readily accepts the concept of structure (in his manuscripts we find even a curious approximation to the concept of a generative model), but the concept of an organism is closer to him, and the argumentation with biological analogies is his favourite device. He does not reject the problem of the dynamic connections between features, but in practice he usually replaces it with accounts of statistic proportion of features (he says, justifying his approach, that in order to speak of a functional relationship, one must clearly separate cause and consequence, *prius* and *secundum*, and this is possible only through a diachronic study and almost never through synchronic analysis). He draws the boundaries between literary categories mechanically: if a certain feature in a text produces over 50% of rhythmic repetition, then it is poetry, if it is less than 50%, then it is prose (though it is obvious that in certain poetic cultures 10% of rhythmical repetitions would

¹¹ Yarkho argues against the concept of *sotsial'nyj zakaz* [‘social demand/order’], which was central to the ideology of the Left Front of Art (LEF). A leading LEF and OPOYAZ critic, Osip Brik, put it thus: “[...] a great poet does not reveal himself, but only fulfills a social demand”. “Had Pushkin not existed *Eugene Onegin* would all the same have been written” (Brik 1977 [1923]: 90–91). — *Eds.*

be sufficient, and in others even 90% would not be enough). And when Yuri Tynianov asserts that the specific metrical position of a word gives it a new semantic content, Yarkho skeptically requests a formulation: what content, exactly?

However, all of this does not detract from the positive sides of Yarkho's methodology: its demand for complete exactness and immutable proof for each assertion. Yarkho understood well that without exhaustive analysis any synthesis risks being arbitrary, and structural synthesis demands it in the same way as any other. He developed his programme of statistical discipline in order to avoid arbitrariness. All of his grandiose undertakings in descriptive literary studies were not ends in themselves, but rather only preparation of material for future generalizations. Several of his generalizations were correctly predicted from the first stages of his work: the law of compensation, the law of wave-like changes; most of the others had to be put aside for future researchers.

The methodological revolution in science has made available to scholars much more precise and dialectic methods than what Yarkho had at his disposal. But in order to apply these methods it is necessary to arrange the analyzed material accordingly. The study of literature had lagged behind other sciences: it was not adequately equipped with formalized material (the only exception was the humble field of versification studies). And without adequately formalized material the structural method in literary studies threatens to degenerate into jugglery with facts arbitrarily ripped out of context in the way of its predecessors. Modern structuralism is correct when it underscores that a device is not a fact by itself; the fact is the relationship between the fact and the background onto which it projects; that the absence of a device may be more telling than its presence.¹² But this means that for the ascertainment of a device we must know the background context just as well as the fact itself: in order to evaluate "minus devices", say, of Pushkin, it is necessary to have a comprehensive picture of the "plus devices" of the preceding epoch. We do not have such a picture yet, but it is indispensable: indices of the poetics of individual authors are just as irreplaceable for the study of literature as concordances and author's dictionaries are for the study of their language. Work in this direction will demand still more effort from researchers, and Yarkho's experience will often prove helpful in these undertakings.

1969

¹² Gasparov refers to Juri Lotman's concept of a "minus device". — *Eds.*

P.S. (1997). Two traits struck Yarkho's contemporaries: his phenomenal erudition and fantastic energy and ability to work in the most unsuitable conditions. When his brother Grigory Isaakovich Yarkho translated *Gargantua and Pantagruel* and reached an impasse in attempting to translate unclear passages and difficult realia, Yarkho – in exile, without books – sent him explanations, even with illustrations. At 25 years old, preparing to become a Privatdozent, he brought back from abroad a store of materials for 18 university lecture courses. This material was prepared so thoroughly that his work on *Mansongr* and “Young Roland” mentioned above each comprised only one part of only two of such courses. After his exile, reaching out to Narkompros [Narodnyj komisariat prosveshchenija, People's Commissariat for Education] with a request for work, he listed his specializations: medieval literature (Latin, French, Provençal, German, Anglo-Saxon and old Scandinavian); stylistics, metrics, poetics, Russian and Slavic folklore, Serbo-Croatian literature, the history and theory of drama; “in addition, I translate from approximately 20 (new and old) Slavic, Germanic and Romance languages”. (He still had to wait a year for employment.) He was not an absolutist in his scientific ideals: just as medieval science died out because it reduced everything in the world to good and evil, so is science of the modern era dying out because it is reducing everything to truth or falsehood, and these no man is capable of distinguishing either.

They will ask me why I constructed my *Methodology* on a principle that was fated to have no future [...]. I will answer thus: first, I made this for my own “ego” thoroughly imbued with a hypertrophied sense of truth and justice; secondly, I believe that our “sciences of truth” are destined to live; and if not, let my theory be the swan song of the old “Philalethist” study of literature. (Yarkho 2006: 27)

References

- Archilochus, Semonides, Hipponax 1999. *Greek Iambic Poetry: From the Seventh to the Fifth Centuries BC*. Edited and translated by Douglas E. Gerber. (Loeb Classical Library 259). Cambridge, Mass.: Harvard University Press.
- Brik, Osip 1977 [1923]. The so-called formal method. Translated by Ann Shukman. In: *Russian Poetics in Translation* 4, 90–91.
- Gasparov, Mikhail Leonovich 1969. Raboty B. I. Yarkho po teorii literatury. In: *Trudy po znakovym sistemam* 4 (Uchenye zapiski Tartuskogo gosudarstvennogo universiteta 236), 504–514.

- Gasparov, Mikhail Leonovich 1972. The metric repertoire of the Russian lyric in the eighteenth through twentieth centuries. In: *Soviet Studies in Literature* 8(4), 365–389.
- Gasparov, Mikhail Leonovich 1980. Quantitative methods in Russian metrics: achievements and prospects. Translated by Gerald Stanton Smith. In: *Russian Poetics in Translation* 7, 1–19.
- Gasparov, Mikhail Leonovich 1987. A probability model of verse (English, Latin, French, Italian, Spanish, Portuguese). Translated by Marina Tarlinskaja. In: *Style* 21(3), 322–358.
- Gasparov, Mikhail Leonovich 1996. *A History of European Versification*. Translated by Gerald Stanton Smith and Marina Tarlinskaja. Edited by Gerald Stanton Smith and Leofranc Holford-Strevens. Oxford: Clarendon Press.
- Gasparov, Mikhail Leonovich 1997. Raboty B. I. Yarkho po teorii literatury. In his *Izbrannye trudy*. T. 2: *O stikhakh*. Moskva: Jazyki russkoj kul'tury, 468–484.
- Yarkho, Boris Isaakovich 1925–27. Granitsy nauchnogo literaturovedenija. In: *Iskusstvo* 2 (1925), 45–60; 3(1) (1927), 16–38.
- Yarkho, Boris Isaakovich 1927. Prostejshie osnovanija formal'nogo analiza. In: Petrovskij, Mikhail Aleksandrovich (ed.), *Ars poetica I* (Trudy Gosudarstvennoj Akademii Khudozhestvennykh Nauk. Literaturnaja sektsija 1). Moskva: Izdanie GAKhN, 7–28.
- Yarkho, Boris Isaakovich 1977 [1936]. A Methodology for a Precise Science of Literature (Outline). Translated by L. M. O'Toole. In: *Russian Poetics in Translation* 4, 52–70.
- Yarkho, Boris Isaakovich 2006. *Metodologija tochnogo literaturovedenija: Izbrannye trudy po teorii literatury* (Philologica russica et speculativa 5). Edited, with notes (pp. 611–807), by Marina V. Akimova, Igor A. Pilshchikov and Maksim I. Shapir (general editor). Moskva: Jazyki slavjanskikh kul'tur.