

The Prosody of a Poet's Prose: Comparative Analysis of the Rhythmic Structure of A. Pushkin Prose

Ekaterina Nakonechnaya
National Research University
"Higher School of Economics"
Saint Petersburg, Russia
nackonechnaya.97@mail.ru

Abstract—In Russian metrics and prosody, the rhythm of prose is considered as a neutral language background. However, several works demonstrate that poet's prose is somehow different from the "pure" prose. The paper talks about the rhythm of prose written by the poet and the reflection of the poetic rhythm in the prosaic text. We use the probability model built on the base of the rhythm of prose and compare the model with the rhythm of prose.

I. INTRODUCTION

One of the types of the language representation in Natural Language Processing is a statistical one. In metrics and prosody, it is customary to use a probabilistic model (or Language Model, LM). This model was developed by A. Kolmogorov and based on the hypothesis of the independency of the rhythmic words [1]. This hypothesis proposes an occasional distribution of rhythmic words in a prosaic text. In other words, a probability of an occurrence of a certain rhythmic word in a certain place of a text does not depend on which rhythmic words surround it [2]. LM is a theoretical model of an iamb and it shows what an iamb would be if a poet writes a poetry like a prose writer does but trying to keep a meter. After LM was developed, its legitimation was proved. The researches of A. Prokhorov allowed to conclude that the model does work since the LM of "The Queen of Spades" was similar to the rhythm of occasional iambs [3]. An occasional iamb is a fragment of prose in which each second syllable is a strong one (or it has a stress) and the amount of all syllables in the prose line is 8 or 9. Each pair of syllables forms an ictus. This fragment could be a verse line in a poetry and is considered as a microsituation of verse in the prose. However, the results of the works of M. Krasnoperova [4] and E. Kazartsev [5] showed a remarkable difference between a poet's prose and a prosaic one. In [5], the rhythm of "Tales of Belkin" was explored. The rhythm of occasional iambs of this prosaic text differs from the LM. This fact leads to that the distribution of rhythmic words in "Tales of Belkin" is not occasional. In this paper, we continue a series of experiments in order to realize if there is a result of a poetic experience in a poet's prose. In other words, if a poetic experience can transform somehow a rhythmic structure of a prose.

II. LANGUAGE MODEL AND RHYTHM OF PROSE

A. Calculation of Language Model

The difference between a poet's prose and a prosaic one is, in particularly, the difference between the probability model

and the rhythm of occasional iambs. We can find remarkable deviations between the LM and occasional iambs of the poet's prose. The extensive studies of M. Krasnoperova demonstrate that the difference between the poet's prose and the prosaic one consists in a distribution of occasional iambs. In the "pure" prose, there is a coincidence in the indexes of the LM and occasional iambs, in the poet's prose, there is a deviation between them.

The probability model developed by A. Kolmogorov got a new comprehension in the theory of the reconstructive simulation of the verse rhythm of M. Krasnoperova. In this theory, the model becomes a cognitive idea [4], [6]. Relying on the description of the model given by M. Krasnoperova, we can consider the LM in the next way.

In metrics and prosody, there is a term 'rhythmic word'. A rhythmic word is a complex of syllables united by a common stress [7]. For example, a combination of two grammatical words 'at school' represents one rhythmic word with the stress on 'school'. A rhythmic word is characterized by that how many syllables there are in a word and what syllable a stress is on. For instance, a trisyllabic word with a stress on the third syllable is 'on the table', a disyllabic word with a stress on the second syllable is 'a phone' and so on.

The first step of the LM formation is a calculation of rhythmic words. The second step is a calculation of their frequency. Frequency is a quantity but not a probability. However, we consider the frequency of rhythmic words as their probability because the greater the sample of words is, the more their frequency tends to the probability.

A rhythmic word is an element of a general population of all rhythmic words in a text. We consider a line as an elementary event. A space of these elementary events is a separate test:

$$\{R', R_1, R_2 \dots R_m\} \quad (1)$$

Where R_m is a line of rhythmic words which is similar to a verse line, R' is a line which is not similar to a verse. In our case, a separate test is a line. If rhythmic words (e.i. an elementary event) occur in such sequence that stresses in them interchange in a certain way, then we consider this line as a verse line peculiar to a certain verse size. We take into consideration only iambic tetrameter lines.

The probability of the occurrence of a line is equal to the probability of the occurrence of rhythmic words:

A sequence of tests presents a space of elementary events:

$$\{\dots R^1_{i1} \dots R^2_{i2} \dots R^3_{i3} \dots R^s_{is}\} \quad (3)$$

between R^1_{i1} and R^2_{i2} (and so on) there are missing R and any number or R^i ; $i_k = 1, 2, 3 \dots 1$. The probability of each of these events is equal to:

$$P(R^t)^t P(R_{i1}) P(R_{i2}) \quad (4)$$

where t is a number of R^i . This probabilistic model (or the language model – LM) is determined by the one which corresponds to the prosaic text. So, the probability of the line of a certain size (for example, iambic tetrameter, or trochaic trimeter) is:

$$P(R) = \sum_{i=1}^l P(Ri) \quad (5)$$

where l is a number of variants of a size. Rhythmic words have a different number of syllables. A boundary between rhythmic words in a line may be in other places. The next examples are the different variants of the Russian iambic tetrameter: 1) *с|и|ж|у| на| сту|ле| у| ок|на|*, 2) *а| кто| при|шел| ко| мне| до|мой|?* So, here we talk about *the variants of one size*.

The probability that a line pertains to a certain variant is:

$$P(Ri/R) = P(Ri) / \sum P(Rk) \quad (6)$$

In our research, we work with such elements of the rhythm of prose as occasional iambs. As we have said already, an occasional iamb is a fragment of prose in which each second syllable is a strong one (or it has a stress) and the amount of all syllables in the prose line is 8 or 9. Each pair of syllable forms an ictus. This fragment could be a verse line in a poetry and is considered as a microsituation of verse in the prose. For example, in the next sentence of “The Queen of Spades” – “*Только Лизавета Ивановна успела снять капот и шляпу, как графиня уже послала за нею и велела опять подавать карету*” – we find the next occasional iamb: *успела снять капот и шляпу*. In this fragment each second syllable is strong, the whole number of syllables are 9, the fragment has syntactic and semantic wholeness. There are 8 configurations of an iambic tetrameter:

1. U-’ U-’ U-’ U-’ *она дала ему три карты*
2. U- U-’ U-’ U-’ *но сохраняла все привычки*
3. U-’ U- U-’ U-’ *и снова занялась работой*
4. U-’ U-’ U- U-’ *в её передней и девичей*
5. U- U- U-’ U-’ *не удовлетворить его* (the sample of N. Zhurbina)
6. U- U-’ U- U-’ *и приготовился метать*
7. U-’ U- U- U-’ *прищурилась и усмехнулась*

8. U- U- U- U-’ *и не из-за неурожая* (the sample of N. Zhurbina)

$$P(R_i) = p_{i1} p_{i2} \dots p_{im} \quad (2)$$

All the examples are from “The Queen of Spades”.

The probability of an iambic configuration is equal to:

$$P(Fj) = \sum_{i=1}^m P\left(\frac{R1}{R}\right) \quad (7)$$

We consider LM as a set of probabilities of iambic configurations and compare them with a relative frequency of occasional iambs [2].

B. Stress Profile

One more method to compare LM with occasional iamb is to calculate stress profiles based on the data of LM and occasional iambs. Stress profile is a frequency of stresses on each ictus in a verse size. In our case, the stress profile shows how many stresses has each ictus of the occasional iambs.

C. The Law of Regressive Accent Dissimilation

In 1953, Kirill Taranovsky explored the law of regressive accent dissimilation. According to this law, weak and strong syllables in a verse interchange through one from the last strong position to the first one. The difference between the frequencies of stresses on strong positions gradually decrease: it means that there is a dissimilation of stresses. The process of regressive accent dissimilation started in the first half of the 18th century. The law of regressive accent dissimilation is reflected in the poems of the 20th years of the 19th century [8]. This law describes changes in Russian iambic tetrameter which occurred after the 20th years of the 19th centuries. Thus, if the stress profile of occasional iambs of the work is alternative, we can claim that the law of regressive accent dissimilation affects on the rhythm.

III. OBJECT OF RESEARCH AND METHODS

Our research is based on the analysis of the rhythmic structure of five texts – “The Queen of Spades”, “Dubrovsky”, “The Captain’s Daughter”, “Egyptian Nights”, and “Kirdzhali”. Our hypothesis is that the poet’s prose would be created by the influence of the poet’s writing experience.

There are two methods how occasional iambs can be obtained from the text. The first method is when syntagmacy of a phrase is not taken into consideration, but what is important is a number of syllables among which each second one is strong. As well as crossed line segments are considered. Syntagmacy means syntactic and semantic wholeness of a phrase. The second method is that we do take into account syntagmacy of fragments: both syntactic and semantic wholeness of a phrase are considered. Crossed line segments are not under consideration.

The first method was followed by Alexander V. Prokhorov. Prokhorov analyzed the rhythm of “The Queen of Spades”. He took these fragments in which there was not necessarily syntagmacy and the fragments could cross each other. The language model and the rhythm of occasional iambic tetrameters were similar that is why it was concluded that “The Queen of Spades” is an example of a “pure” prose where there

is no dependency between rhythmic words and they are distributed occasionally. Then Prokhorov's results were not verified.

The first research based on the second method was done by Vladislav E. Kholoshevnikov [9]. Based on the amount of rhythmic words he built the language model of prose written by both poets and prosaic writers. The results demonstrated that the language models of poet's prose significantly differ from the models of "pure" prose.

His researches were continued by Evgeny Kazartsev and Marina Krasnoperova which follow the same method of obtaining occasional iambs [10]. In their works, they take into consideration the syntagmacy of occasional iambs. Kazartsev calculated the language model and the rhythm of occasional iambs of "Tales of Belkin". The results are significant and show that the distribution of rhythmic words in "Tales of Belkin" are not occasional and there is some dependency between them because the indexes of the language model and the rhythm of occasional iambic tetrameter are different. Moreover, the stress profile of "Tales of Belkin" is similar to the alternative one which is character to the latest poetic works of A. Pushkin. So, it was suggested that the poet's prose is influenced by the verse.

We follow the second method of getting occasional iambs. In order to be sure that a fragment is got accurately we can put it in a context of verse. This research is a continuation of the previous one [11] in which we analyzed only "Dubrovsky", "Captain's Daughter" and "The Queen of Spades".

The calculations of the models and the rhythm of occasional iambs were made with the help of the new program created by V. Vashchenkov. The program was developed under the guidance of professor E. Kazartsev. The development is supported by RGNF, project No. 15-04-00331.

IV. RESULTS

1) "Dubrovsky"

The research of A. Prokhorov got support to the hypothesis of the independence of rhythmic words in "The Queen of Spades". We got similar results on the material of "Dubrovsky". Follow our calculations, the rhythm of occasional iambs is similar to the data of the language model (LM was calculated by K. Taranovsky), even though there is a slight difference between the indexes of the third and sixth iambic configurations.

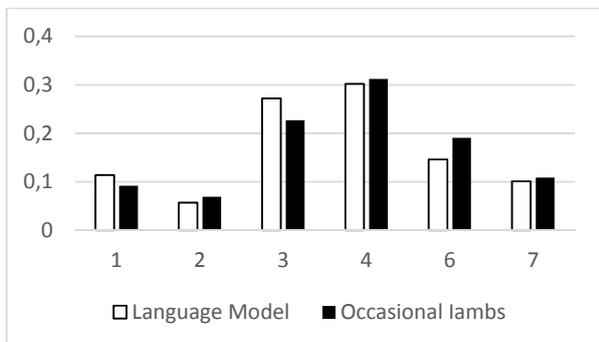


Fig. 1. The language model and the rhythm of occasional iambs of "Dubrovsky"

As well as the stress profile of the occasional iambs resembles to the one of LM. The frequency of stresses on the first ictus in occasional iambs is a bit lower than in LM. The frequency of stresses on the second ictus in occasional iambs is a bit of higher than in LM.

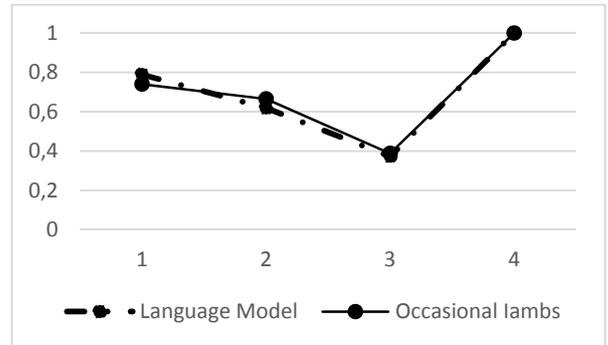


Fig. 2. The stress profiles of the language model compared with the rhythm of occasional iambs of "Dubrovsky"

The data get support to the hypothesis of the occasional distribution of rhythmic words because the proximity of the LM and the rhythm of occasional iambs is observed. As the author of the text is, first of all, a poet, we can suggest that he very controlled himself when he wrote the novel and did not admit an influence of the rhythm of verse to occur in the text.

2) "The Queen of Spades"

Since the Prokhorov's work was published, the data of the LM and occasional iambs of "The Queen of Spades" were not verified. Thus, we decided to obtain occasional iambs using the alternative method and compare them with the LM.

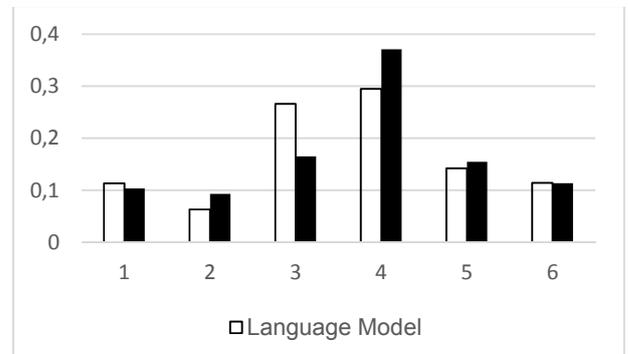


Fig. 3. The language model and the rhythm of occasional iambs of "The Queen of Spades"

The LM was calculated by A. Prokhorov. From the graph, the LM and the rhythm of occasional iambs are not similar. There is a difference in the index of the third, fourth and sixth configurations of occasional iambic tetrameter.

For the third iambic configuration, it is characterized that the second ictus has a gap on the strong position. The more amount of the third iambic configuration in the model, the more probably the model will be similar to the Russian iamb of the beginning of 19 century and the less probably it will be alternative one. In our model, this index is lower than in Prokhorov's model. It leads that in our sample there are more occasional iambs which have stress on the second strong position. This difference is observed on the graphics. As for the fourth and sixth iambic configurations, the predominance of

these forms in the rhythmic structure leads the stress profile to be alternative.

The stress profile of the LM was calculated by K. Taranovsky. This tends to be alternative. The fourth ictus always has a stress. The third one has the least number of stress. The frequency of the stress on the second ictus is much higher than on the previous one. On the first ictus there is a decrease of a stress frequency again. This tendency shows us the alternation of the distribution of stresses on each ictus (or the regressive accent dissimilation).

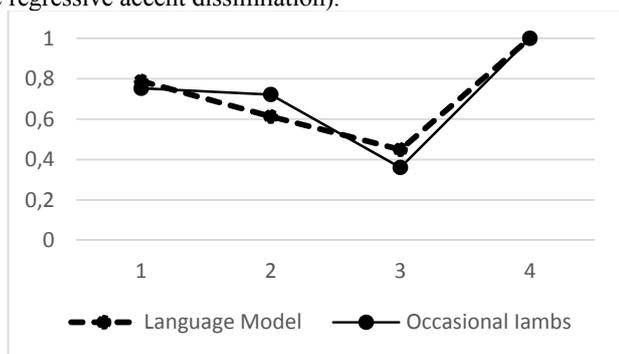


Fig. 4. The stress profiles of the language model compared with the rhythm of occasional iambs of "The Queen of Spades"

The results are similar to the one obtained by E. Kazartsev based on the material of "Tales of Belkin". They get support to the idea that the poet's prose is under the influence of the poet's writing experience because the stress profile of occasional iambs resembles the one of the latest poetic works of Pushkin.

3) "The Captain's Daughter", "Kirdzhali", "Egyptian Nights"

For "The Captain's Daughter", we calculated the language model and the rhythm of occasional iambs. For "Kirdzhali" and "Egyptian Nights", we calculated only the rhythm of occasional iambs. The LM of "The Captain's Daughter" is very similar in the indexes to the LM of "Dubrovsky" and "The Queen of Spades", therefore, we consider the LM of "The Captain's Daughter" as the common model for all three texts.

The same tendency as in the previous text is observed in the rhythm of occasional iambs of "The Captain's Daughter", "Kirdzhali", "Egyptian Nights". In the table below, there are indexes of the rhythm of occasional iambs of three latest works of A. Pushkin.

TABLE I. THE LANGUAGE MODEL AND OCCASIONAL IAMBS OF PUSHKIN'S PROSE

| Configu-ration | "The Captain's Daughter" (LM) | "The Captain's Daughter" | "Kirdzhali" | "Egyptian Nights" |
|----------------|-------------------------------|--------------------------|-------------|-------------------|
| 1 | 0,0893 | 0,1108 | 0,1200 | 0,1591 |
| 2 | 0,0469 | 0,0747 | 0,1600 | 0,0455 |
| 3 | 0,2774 | 0,1933 | 0,2000 | 0,1136 |
| 4 | 0,2980 | 0,3402 | 0,2800 | 0,2273 |
| 6 | 0,1670 | 0,1985 | 0,1200 | 0,3409 |
| 7 | 0,1193 | 0,0799 | 0,1200 | 0,1136 |

The indexes of the 2nd configuration increase in the rhythm of occasional iambs in comparison with the LM. The frequency of the first iambic configuration is higher in all three texts. The number of the third iambic configuration in occasional iambs is lower. This results in that the stress profiles of the texts will be similar to the one of "The Queen of Spades".

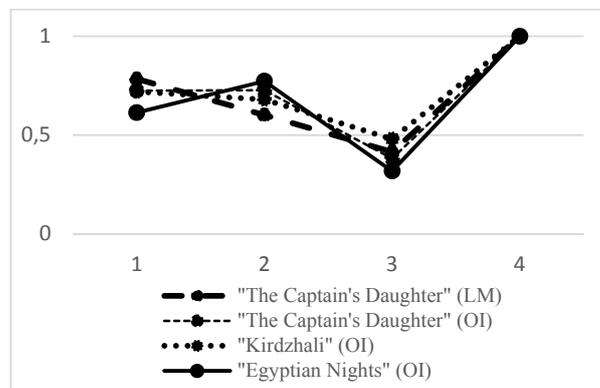


Fig. 5. The stress profiles of the language model compared with the rhythm of occasional iambs of "Captain's Daughter", "Kirdzhali", and "Dubrovsky"

From the graph, it is remarkable that the later the prose was written, the more alternative the stress profile is. The stress profile of "Egyptian Nights" is the most alternative one of all the profiles. This is very similar to the one of poetic works of the 20th years of the 19th century, particularly, the poems of Pushkin. These stress profiles resemble the ones of poetic works from the beginning to the 20th years of the 19th century. For the poetic works of the beginning of the century, the stress profile, like the one of "Kirdzhali" (1834), was characterized. Then the stress profile transformed to the one like "The Captain's Daughter" (1833-1836). In the 20th years, the stress profile of poetry was very alternative, like the one of "Egyptian Nights" (1835). Thus, we can claim that the law of regressive accent dissimilation is reflected in the rhythm of the later prose of the poet.

V. CONCLUSION

In the paper, we analyze the rhythmic structure of the poet's prose. In the researches of M. Kranoperova and E. Kazartsev, it was demonstrated that the poet's prose is influenced by the verse. We continue the experiments and obtain occasional iambs from the prose using the method of V. Kholshenikov. We consider occasional iambs as a microsituation of verse in prose. The results show that the rhythm of occasional iambs differ from the language model in all five texts. Not all differences are significant but they all tend to change in the way of verse. Therefore, we can suggest that the author somehow controlled himself when he wrote "Dubrovsky". He intentionally restrained himself in order not to write the lines of verse. However, in other works, A. Pushkin did not think of "strong" prosaic characteristics, that is why we observe an influence of poet's writing experience, or the influence of verse on the rhythm of prose.

This article is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE).

VI. REFERENCES

- [1] *Kolmogorov A.* Primer izucheniia metra i ego ritmicheskikh variantov // *Teoriia stikha*. 1968. P. 145-167.
- [2] *M. A. Krasnoperova.* *Osnoy sravnitel'nogo statisticheskogo analiza ritmiki prozy i stikha*. St. Petersburg, 2004.
- [3] *Prokhorov A.* About an occasional versification (to the question of theoretical and language models of verse) // *The Problems of Verse Theory*. 1984. P. 89-98.
- [4] *Krasnoperova M.* K voprosu o raspredelenii ritmicheskikh slov i sluchainykh iambov v khudozhestvennoi proze // *Russkii stikh: Metrika, ritmika, rifma, strofika. V chest' 60-letia M. L. Gasparova*, edited by D. P. Bak. 1996. P. 163-181.
- [5] *Kazartev E.* The Rhythmic Structure of The Tales of Belkin and the Peculiarities of a Poet's Prose // "A Convenient Territory". Russian Literature at the Edge of Modernity. Essays in Honor of Barry Scherr. Edit. J. Kopper & M. Watchel. 2015. P. 55-65.
- [6] *Krasnoperova M. A., Kazartsev E. V.* Reconstructive Simulation of Versification in Comparative Studies of Texts in Different Languages: Theoretical Aspects and Applications // *Frontiers in Comparative Prosody*. In memoriam: Mikhail Gasparov. Linguistic Insights. Studies in Language and Communication. Vol. 113, 2011. P. 97-120.
- [7] *E. V. Kazartsev,* *An Introduction to Comparative Metrics and Prosody*. St. Petersburg, 2015.
- [8] *K. F. Taranovskii.* *Russkie dvuslozhnye razmery. Stat'i o stikhe*. Moscow: Iazyki slavianskoi kul'tury, 2010.
- [9] *Kholshevnikov V.* Sluchainye chetyrekhstopnye iamby v russkoi proze // *Russkoe stikhoslozhenie: Traditsii i problem razvitiia*, edited by L. I. Timofeev. 1985. P. 134-143.
- [10] *Kazartsev E.* Comparative Study of Verse: Language Probability Models // *Style*. 2014b. Vol. 48. P. 119-139.
- [11] *Kazartsev E., Nakonechnaya E.,* The Prosody of a Poet's Prose and the Rhythm of Verse // *Frontiers in Comparative Metrics 3* in memoriam Jaak Polmae. 2017. P. 26