Verse and Prose: Linguistics and Statistics

Tatyana Skulacheva
Institute of Russian Language of Russian Academy of Sciences
skulacheva@yandex.ru

Alexander Kostyuk
Philological Faculty of M.L. Lomonosov Moscow State University
kostyuk.ae@gmail.com

Abstract

Stable (statistically proven) linguistic differences between verse and prose were studied in a number of European Languages (Russian, French, Spanish, English). These differences regularly occur in verse of different meters, different systems of versification, different periods and literary trends.

1. Syntax. Parataxis is increased in verse, while hypotaxis is decreased as compared to prose, as show our calculations for 30,000 lines of Russian, French, and Spanish poetry and prose of the 17th–20th centuries.

2. Intonation. Analysis of intonation in the Praat program and the statistical analysis of the obtained data shows that Chervenka’s suggestion that verse has intonation of its own 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.

3. Other differences: there are number of other differences, connected with semantics, information processing, mistakes detection, etc. Most informative words in verse are scattered along the line, while in prose they occur most often at the end of the syntagm under syntagmatic stress. Participants of our experiment easily detected mistakes in prose, while in verse mistakes almost always remained undetected.

4. Recent data obtained by neurophysiologists at Carnegie Mellon University in the USA as well as British data from Liverpool and a number of other British universities enable us to suppose, that all these peculiarities observed in verse may be linguistic mechanisms of activating imaginative thinking at the expense of logical thinking.
1 Introduction

Determining the most stable linguistic features of verse as opposed to prose is one of the important tasks of modern verse study. We are primarily interested in those features of verse which consistently occur in verse of different languages, periods, literary trends, and presumably are required by verse form as such. As we know, predictable meter, rhyme, stanzas, alliteration, syntactic parallelism, enjambement and a big number of other verse elements can be totally absent from a verse text without it ceasing to be verse. The only parameter which cannot disappear from the text without the text loosing its verse character is its division into lines. Still, we do not so far possess enough knowledge about indispensable characteristics of a verse line as the basic unit of a verse text. We will describe some linguistic peculiarities of verse closely connected with its division into lines and suggest a hypothesis of how linguistic peculiarities of verse form a system aimed at a peculiar type of impact on the reader’s consciousness and brain.

2 Linguistic differences between verse and prose and their impact on the reader’s consciousness

If we look at the set of linguistic peculiarities of verse which most consistently occur in verse of different languages, periods and literary styles we will be surprised by the fact that the majority of them seem to be aimed not at a better understanding of a text, but, quite the opposite, at making the clear, precise, straightforward understanding of a text difficult, and, in some cases, even impossible. What might be the reason of writing a text and at the same time preventing the reader from a quick and effective understanding of it? Let us describe some of the linguistic peculiarities of verse structure and give possible explanations of their functions.

2.1 Syntax

We have analyzed about 30,000 lines of Russian verse of the 17th–20th centuries and French verse of the 18th–20th centuries and compared it with prose. We have analyzed parataxis and hypotaxis between clauses in the verse and prose by the same author for 14 Russian authors of the 17th–20th centuries and of 8 French authors of the 18th–20th centuries (Skulacheva–Buyakova 2010a; Skulacheva 2014). In verse we analyzed parataxis and hypotaxis between clauses in a position between lines, because our data show (Skulacheva–Buyakova 2010b), that though the parataxis grows in quantity in all positions within a verse line, the main place where the growth is most intensive is the position between lines. This probably happens because it is this position where parataxis is most functional in verse structure: it helps to present verse lines as equal or compatible in semantic weight and prominence—a feature mentioned by Gasparov as one of the main peculiarities of verse division into lines, while in prose the most important is the hierarchy of prominence, the demonstration of what is more important and what is less important in a text.
Data given in TAB. 1 and TAB. 2 show that there is very stable growth of parataxis in verse in comparison to prose within the texts of the same author. In fact there are so far no exceptions found to the fact that each author differentiates his verse from his prose by the important (normally a few dozen percent) growth of parataxis in verse.

<table>
<thead>
<tr>
<th></th>
<th>% of all ties between clauses</th>
<th>VERSE</th>
<th>PROSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Folklore of the 17th century</td>
<td></td>
<td>92.2</td>
<td>68.4</td>
</tr>
<tr>
<td>Simeon Polocky</td>
<td></td>
<td>62.3</td>
<td>46.0</td>
</tr>
<tr>
<td>Lomonosov</td>
<td></td>
<td>54.7</td>
<td>22.8</td>
</tr>
<tr>
<td>Zhukovsky</td>
<td></td>
<td>81.7</td>
<td>62.0</td>
</tr>
<tr>
<td>Pushkin</td>
<td></td>
<td>74.1</td>
<td>36.3</td>
</tr>
<tr>
<td>Lermontov</td>
<td></td>
<td>77.5</td>
<td>57.0</td>
</tr>
<tr>
<td>Tyutchev</td>
<td></td>
<td>78.6</td>
<td>28.6</td>
</tr>
<tr>
<td>Fet</td>
<td></td>
<td>79.8</td>
<td>30.6</td>
</tr>
<tr>
<td>Kuzmin</td>
<td></td>
<td>62.5</td>
<td>40.5</td>
</tr>
<tr>
<td>Blok</td>
<td></td>
<td>70.9</td>
<td>42.8</td>
</tr>
<tr>
<td>Khlebnikov</td>
<td></td>
<td>76.4</td>
<td>46.7</td>
</tr>
<tr>
<td>Akhmatova</td>
<td></td>
<td>77.6</td>
<td>36.4</td>
</tr>
<tr>
<td>Tvardovsky</td>
<td></td>
<td>65.1</td>
<td>32.6</td>
</tr>
<tr>
<td>Brodsky</td>
<td></td>
<td>51.5</td>
<td>29.0</td>
</tr>
</tbody>
</table>

TAB. 1: Percentage of parataxis in verse and prose of Russian authors (17th–20th centuries)

<table>
<thead>
<tr>
<th></th>
<th>% of all ties between clauses</th>
<th>VERSE</th>
<th>PROSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voltaire</td>
<td></td>
<td>71.7</td>
<td>38.7</td>
</tr>
<tr>
<td>Hugo</td>
<td></td>
<td>73.7</td>
<td>26.3</td>
</tr>
<tr>
<td>G.de Nerval</td>
<td></td>
<td>73.0</td>
<td>32.5</td>
</tr>
<tr>
<td>Baudelaire</td>
<td></td>
<td>57.8</td>
<td>34.5</td>
</tr>
<tr>
<td>Mallarmé</td>
<td></td>
<td>39.2</td>
<td>28.6</td>
</tr>
<tr>
<td>Verlaine</td>
<td></td>
<td>57.0</td>
<td>27.0</td>
</tr>
<tr>
<td>Valéry</td>
<td></td>
<td>54.5</td>
<td>26.8</td>
</tr>
<tr>
<td>Eluard</td>
<td></td>
<td>76.5</td>
<td>34.0</td>
</tr>
</tbody>
</table>

TAB. 2: Percentage of parataxis between clauses in verse and prose of French authors (18th–20th centuries)

We also have data with similar regularity on Spanish verse, obtained by A.V. Kruglova (Kruglova–Smirnova–Skulacheva 2017), which show that parataxis between lines is 79%, and hypotaxis is 21% in the verse of Francisco de Quevedo (1580–1645). There are also data by A. Ilyushina, V. Bayburtsyan, S. Zarudneva, presented at the Gasparov conference of 2019 in Moscow, which show the same regularity for English verse of the 18th–19th centuries and Spanish verse of the 17th century as compared to prose. Why does verse so consistently need parataxis between lines? If a person is most interested in presenting the logic of his speech to the listener he or she will obviously use hypotaxis to show what is more and what is less important and to reveal the logical connection between clauses in his or her speech. Instead, verse uses parataxis, especially in positions corresponding with line borders:
His hall scarce echoes with his wonted name,
His portrait darkens in its fading frame,
Another chief consoled his destined bride,
The young forgot him, and the old had died...

(G.G. Byron “Lara”)

Thus, on the syntactic level verse works against an easy and quick understanding of the logical hierarchy of the text by increasing parataxis at the expense of hypotaxis between lines, while prose facilitates understanding of its logic by intensive use of hypotaxis.

2.2 Intonation

Intonation serves the same purpose of hampering quick and easy understanding of a text. Regular educated readers as well as some authors read verse with very even intonation, without bringing out important words, without focusing on something, without theme-rheme differentiation. As has been shown by T.E. Yanko, such intonation is typical of verse as well as of prayer in different languages (Russian, Italian, Arabic). FIG. 1 gives the pitch movement in the verse “Sretenye” read by I. Brodsky himself as shown by T.E. Yanko (2010, 2015). Each even segment on the picture corresponds to a verse line.

As we see in FIG. 1, intonation of the verse line is very even and monotonous, without emphasis on important words.

It has been shown by our experimental phonetic analysis that diapason in the oral presentation of verse is twice as narrow as that of read-aloud prose. When one tries to make the logic of oral speech most explicit, he or she varies pitch considerably to bring out the most important words and to generally bring out the logical hierarchy of information in an oral text. Verse remains monotonous with as little emphasis as possible. There is also another type of reading verse, typical of poets themselves rather than of regular readers. Some poets tend to emphasize every word in verse. This seemingly opposite way of reading actually serves the same purpose as monotony and narrow diapason: one can't easily understand what is important and what is not. One cannot tell more important words from less important ones when every word is emphasized. So the very existence of these two types of reading suggests that the main aim of both types of reading is to neutralize an intonation based on a logical hierarchy of words in a text.

In FIG. 2 we suggest the invariant of verse intonation based on experimental data which we have obtained.

It is characterized by the technical rise at the beginning (as we technically cannot immediately reach a pitch high enough for intonation to go down from), then a smooth and even lowering within a rather narrow diapason from the beginning to the end, and no distinct drop at the end (Kostyuk 2017).

We also tried to prove experimentally that intonation specific for verse does exist and is recognized by listeners. We recorded and delexicalized 4 verse texts and 1 prose text
FIG. 1: “Sretenye” read by Brodsky, pitch (Yanko 2015)

FIG. 2: Generalized tone frame pattern of prosodic structure of a verse line with reduced values of coefficient of the tone frame pattern at the beginning of a line ($k_1$) and at its end ($k_2$)

(dactyl by A. Blok, dolnik by I. Brodsky, verse libre by M. Kuzmin, English iambus by J. Tolkien). By delexicalization we mean the processing of sound signal so that only pitch is left but particular words become unrecognizable. Rhythm also was not easily recognizable or did not matter because free verse by M. Kuzmin with no predictable rhythm was recognized even better than the dactyl by A. Blok. The delexicalized text (where only pitch was left) was played to 23 participants of the experiment. The group of participants consisted of University students of ages 20–22. The duration of segments played to participants of the experiment was about 30 seconds. The graph below shows the percent of those who judged correctly (green) and incorrectly (violet) whether a particular text was verse or prose.

As we see in FIG. 3, verse and prose is quite easily recognized by intonation (pitch movement) only, which proves that specific verse intonation, the existence of which was once predicted by M. Chervenka, does actually exist. The only text that presented more problems for recognition was I. Brodsky’s “Niotkuda s l’ubovju”, a very complicated text, for which 30 seconds did not seem to be enough for recognition. Another experiment of a similar kind has lead to really unexpected consequences. We increased the number of texts given to participants of the experiment up to 23, the majority of which were verse. Quite unexpectedly, listening to a bigger portion of delexicalized verse texts caused complaints by informants of headaches, depressed moods,
and feeling sick, so that between one third to one fourth of participants refused to finish the experiment. This unexpected consequence shows that verse intonation given separately from other linguistic levels has a pretty strong influence not just on our consciousness but on some physiological parameters which we hope may be tested in a biochemical experiment in the near future.

2.3 Semantics and information significance hierarchy

There is also a hierarchy of the significance of words in a sentence (or, in our case, also in a verse line) because different words can be of more or less help for us when reconstructing the information of the sentence. If we need to meet somebody at the airport and we speak to that person on the phone in the noisy environment, it is of primary importance what words we hear and which we do not. For example, if out of the sentence “I come tomorrow at five” we hear “tomorrow at five” the major part of the information has reached us, if we hear “I” we got almost no information: it may mean “I have not got a ticket for tomorrow, I arrive in an hour”, or “I fell ill and do not come at all”. We asked our participants of the experiment to mark words in four-word sentences and four-word lines as “1” (“if you had a chance to hear only one word out of four, which will help you most to reconstruct the information delivered by the sentence), “2” (“if you had a chance to hear only two words out of four, what would be the second word which will help you most to reconstruct the information delivered by the sentence), same with “3” (“third word to hear”) and “4”. In prose participants of the experiment had very similar opinions on the hierarchy of the significance of words in a sentence apprehended in a noisy environment. Most often the most information-loaded words turned out to occur at the end of a sentence, where, as shown by phoneticians, phrasal stress most often appears. With verse, the situation was completely different. There were much more difference in opinion among the participants of the experiment regarding which words are more or less significant.
The most significant words turned out to be scattered randomly throughout a verse line, with even more significant words at the beginning of a line than at the end, and with almost unpredictable positions of important words in a verse line. Prose is aimed at quick understanding and we know where to look for the most informatively significant words, while in verse their position varies and becomes unpredictable. We once again have a mechanism which prevents easy and quick comprehension instead of facilitating it.

Yu. D. Apresyan once mentioned in an oral discussion that the main difference between verse and prose from the point of view of a semantician is that while in prose context normally helps to choose one meaning of a polysemantic word, in verse it often deliberately prevents the choice of one meaning. For example, a well-known poem by A. Blok “V goluboy dalekoy spalenke” (“Far away in a blue bedroom...”) a situation is described when a woman loved by the author is in bedroom of her child. The word describing the child’s condition is the Russian verb “opochil”, which in Russian may mean ‘asleep’ or ‘dead’. We questioned a number of professors and students of the philological faculties of Moscow State University and Russian State University for the Humanities and it turned out that opinions fall into almost equal groups: half imagined a peaceful and cozy room with a woman loved by the author and a sleeping child; the other half saw a tragedy in which a child has just died and the woman is grieving his death. This is not a common situation in prose, when the state of one of the main characters is directly described in the text and still almost half of the people sees one meaning (‘asleep’) and another half sees a completely different one (‘dead’), and both groups of people are fully satisfied with their understanding.

There are also cases in verse when an object is named in a few different ways which are normally mutually exclusive alternatives in everyday life. For example, there is a poem by Mikhail Kuzmin “Trzimenskie trostniki” (Trasimene reeds”), where the same object is called a lake, a sea and a river within a short text. What would be viewed as an obvious mistake in prose remains unmentioned by a reader in verse. Readers imagine some features common to all three objects,—for example, water glittering under the sun, a shore and reeds in the water near it. In verse, the illogicality remains undetected and a reader remains fully satisfied by the text.

V. Kimmelman, who was our student, has shown that while participants of the experiment easily detected mistakes in prose, same mistakes remained undetected in verse (Kimmelman 2012). One serious mistake (he is given three objects—two are enumerated) was deliberately introduced into a verse text and the same mistake was introduced into a prosaic retelling of the text. The prosaic retelling was made as close to a verse text as possible—the only changes were aimed at destroying meter and rhyme by changing word order and, when this didn't suffice, by using one synonym instead of another (compare Skulacheva 2004). Mistakes remain undetected in verse, which signals the suppression of critical thinking, which psychologists consider one of the features of the altered states of consciousness; but we think that it may be not just an altered state of consciousness, but also a normal feature of imaginative thinking. We use the term “imaginative thinking” here in an old-fashioned everyday meaning, for lack of the correct modern term (fully different from how it is customarily used in
psychology). Thus, deciding how many chairs will fit into a room is what we will consider logical thinking, in contrast with modern use of this term. Logical thinking for us is when we think of a New Year from the point of view that a year contains 365–366 days, 12 months and so on; we put objects into strict logical hierarchy, we know what is hyponym, what is hyperonym and what are the logical relations among objects. Critical thinking is an indispensable part of this type of work with information: if we say that there are 13 months in a year everybody would immediately detect it. On the contrary, we can think about the New Year in a quite different way: imagining tangerines, the smell of fur-trees and cinnamon, Christmas tree balls, etc. In this case, we will not be interested in whether tangerines are more/less important than cinnamon, or concentrated on the logical relations between objects, or care about any hierarchy between objects and we will not be concentrated on mistakes. We will deal with objects like with equally significant images, and our critical thinking would obviously not reach the level of activity that occurs when we think that a year contains 12 months. Thus we may suggest that critical thinking is more typical for logical thinking and is suppressed at the point when imaginative thinking starts to dominate.

3 Conclusion

Thus, we can conclude that the linguistic structure of verse at all linguistic levels seems to provide deliberate obstacles to the quick and easy understanding of a text, and that the result of this strange text organization is that the readers get into the physiological condition when they seem, among other things, to loose critical thinking and not mention even obvious mistakes and illogicalities.

Recent neurophysiological studies suggest some possible explanations to this. There are a few articles from two different groups, one is Carnegie Mellon in the US (Mason–Just 2007; Mitchell et al. 2008), another is a group from Liverpool and a number of other British Universities (Keidel et al. 2013; Thierry et al. 2008; see also Falikman 2017; Jacobs 2015). The Carnegie Mellon group studied the reaction of the brain to lexical ambiguities, and the British group to syntactic ones. Both show that when some difficulties occur when processing information, segments of the right semisphere become more strongly activated. This may mean that we are dealing here with the linguistic mechanisms of the activation of imaginative thinking.

Our preliminary studies may suggest that the same elements (or at least partially same, i.e. the specific monotonous intonation without bringing out important words, the growing number of paratactic constructions and the lowering number of hypothetic ones) are not confined to verse but may be characteristic of speech in a number of situations when logical thinking is suppressed by stress, the application of specific linguistic mechanism working against the easy application of logical thinking, the use of certain chemicals (alcohol or drugs) suppressing logical thinking as well as mistakes detection.

Yu. I. Aleksandrov and his coauthors (Aleksandrov et al. 2017) describe a set of conditions for which they use Freud’s term “regression” (meaning that in such conditions
people fall back to some previous state of their consciousness). These conditions include stress, borderline emotional conditions, learning, and alcoholic intoxication. Learning seemed to us at first a stranger on this list but listening to students learning to speak about something scientifically new to them we noticed even intonation, growth of parataxis, eyes fixed on the ceiling and some other elements. Also, as it has been mentioned to us by V. Nurkova, some similar phenomena may also be observed in situations dealing with autobiographic memory. We specially observed video-recordings of people being interviewed about events in their pasts and we noted that intonation became more even, parataxis grew, etc. All phenomena mentioned (verse, prayer, meditation, stress, borderline emotional conditions, speech in alcoholic or drug intoxication, possibly situations dealing with learning and autobiographic memory) are so far brought together only by our hypothesis and consistent longterm cooperation with physiologists of the brain and even biochemists may provide us with information which will enable us to see what is correct in this hypothesis. It is also obvious that the physiologists with whom we have started to work at present could formulate this in the much more precise terms of the natural sciences. Still, we thought it useful to mention this possible explanation of the difference between verse and prose and some similarities in the speech characteristics in this pretty broad set of phenomena, all of which are presumably characterized by the suppression of logical thinking, with imaginative (in the old, everyday sense of this word) thinking being left as the dominating mode of thinking.

Acknowledgment

This work is supported by the grant of Russian Foundation for Basic Research N 19-012-00534 under the supervision of T.V. Skulacheva.

References


