A Corpus-Based Study of Adjectives that Describe People

Anna Shevlyakova [0000-0002-2659-1887], Valery Solovyev [0000-0003-4692-2564] and Vladimir Bochkarev [0000-0001-8792-1491]

Kazan Federal University, Kremlyovskaya str. 18, Kazan, Russia
anna_ling@mail.ru

Abstract. The computer methods were used in this article to study attributive constructions, namely, Russian adjectives that combine with the words *chelovek, muzhchina, zhenshchina*. The research was based on the Google Books text corpus. Besides, the Open Corpora morphological dictionary was used to automatically determine parts of speech of the words that contain in the obtained word combinations. Approximately 17 thousand adjectives were identified and classified into five groups. The center and periphery of the colour palette of the studied colour adjectives were identified. The analysis of the adjectives describing appearance and physical parameters of a person revealed the most significant body parts. The obtained data also allowed making a frequency portrait of a typical man and woman of the 19th and 20th centuries. It was also revealed how age is represented by adjectives in the Russian texts. Analysis of the opinion adjectives showed the most significant traits of character attributed to the words under study. Adjectives that describe physical condition, emotional state and social status demonstrated the most typical ones for Russian written culture. All the adjectives were analysed in terms of frequency. It was shown that most often the adjectives combine with the generic word *person*. As for the nouns *man* and *woman*, there are adjectives that combine more often with *woman* and never with *man* and vice versa. The diachronic analysis of the adjectives showed how their use in combination with the nouns under study can change with time and how it can respond to significant historical events.

Keywords: adjectives, attributive constructions, Google Books Ngram.

Introduction

Computational methods have been widely used in linguistics for the last several decades. This happens due to different factors. One of them is big data. Creation of extra-large text corpora such as Google Books or COHA allows one to use computational methods for linguistic studies. In this paper, these methods are used to study word combinability.

The problem of word combinability has been studied in linguistics for a long time. There are different views on the nature of combinability and the degree of its idiomaticity [1,2,3]. However, word combinability is semantically determined [4] and reflects how people perceive the surrounding reality. By studying lexeme combinability, one can reconstruct individual fragments of the linguistic worldview.

The issue of lexeme combinability has been investigated in various scientific papers. For example, attributive constructions are studied in [4,5]. Studies of colour terms have been also a productive field of research in terms of combinability [6,7].

This article studies attributive constructions, namely, Russian adjectives that describe the words *chelovek*, *muzhchina*, *zhenshchina* (person, man, woman). It was decided to study these constructions for the following reasons. Firstly, the fact of the anthropocentricity of the linguistic worldview has been known for a long time. However, it seems interesting to understand how a person perceives himself in the framework of this worldview and how this perception can change over time. Secondly, the word *chelovek* is a generic category and the words *muzhchina* and *zhenshchina* are basic categories. The basic level of categorization is the most significant one for the language since it is most fully reflected in the linguistic worldview [8]. Therefore, specific categories (e.g., names of professions and occupations, names of kinship, etc.) were not considered in the article.

Two Russian text corpora served as a study material. The first one is Google Books [9], the famous electronic library which contains millions of diachronically marked texts. The Russian sub-corpus of Google Books Ngram is based on the texts of 591 thousand books and includes more than 67 billion of words. This corpus is the largest one among other diachronic Russian text corpora. It contains data for the period 1607-2009. However, most of the texts were written within the last two centuries [9]. Google Books was used to identify attributive constructions and study the dynamics of their use. The second one is an electronic morphological dictionary Open Corpora [10]. Currently, this dictionary includes 391,268 lemmas containing 5,128,422 inflectional forms. Thus, it is one of the largest dictionaries of the Russian language. The second one is an electronic morphological dictionary Open Corpora [10]. Open Corpora was applied to automatically determine parts of speech of the words contained in these constructions. The LC system was used in the article to perform transliteration.

1 Method

The GBN corpus (the version of 2012) was used to find 2-grams of the *Adj+Noun* type. The nouns that combined with the adjectives were presented by forms of the words *person*, *man* and *woman* (there are 36 of such word forms, according to Open Corpora, including the suppletive form *liudi*). Since the POS mark-up of the GBN corpus contains a lot of errors [11], the Open Corpora morphological dictionary was used to verify parts of speech (adjectives). This also allowed us to exclude word forms with the ORC errors which often occur in the GBN corpus. Thus, 24504 2-grams were selected. These 2-grams contain 17183 adjectives of 3613 different lemmas. Lemmatization was performed based on the Open Corpora dictionary.

After that, data on frequencies of the selected 2-grams were extracted from the GBN corpus and total frequencies for each lemma were calculated. To present the graphs of the time series of the frequencies, smoothing was used with a Kaiser window (with the parameter β equal to 3.5) with a half-width of 3 years. The selected adjectives were analysed and classified manually.

2 Results

The automatic analysis allowed us to identify approximately 17 thousand forms of adjectives. In other words, different case-forms of the nouns *chelovek, zhenshchina* and *muzhchina* are described by 3613 adjectives (without considering the case-forms of the adjectives). Major part of the relative adjectives was not considered in the article and qualitative adjectives and the rest smaller part of the relative adjectives were classified (according to their semantics) for a more detailed description. There are various approaches to classifications of adjectives [12,13,14,15,16]. Based on these classifications and the obtained data, the adjectives were divided into five groups:

1. Colour adjectives (*belyi, raznotsvetnyi*).
2. Adjectives describing appearance and stature of a person (*kosmatyi, vysokoroslyi*).
3. Adjectives describing age and "temporal relevance" of a person (*molodoi, sovremennyi*).
4. Opinion adjectives (including adjectives denoting traits of character and abilities: *vydaiushchiisia, milyi, luchshii*).
5. Adjectives describing emotional and physical state of a person, his social status (*ustalyi, peshii, odinokii, bezdetnyi*).

However, it should be noted that the proposed classification is not strict, since the same adjective can belong to different groups depending on the context. The main objective of this work is not to analyse semantics detailly but to reveal language patterns and trends.

2.1 Colour

The first group included adjectives that denote colour and combine with the nouns under study. Most of them have a figurative meaning (including idioms), and not just name a colour of a person. If talking about colours, any object of reality has its colour. However, people talk about the colour of a particular object if its colour can change and is significant for a person [4]. The amount of colour adjectives combined with the considered words are relatively small. The analysis revealed the following adjectives: *chernyi, belyi, seryi, ryzhii, krasnyi, tsvetnoi, zheltyi, ryzhevatyi, zelenyi, goluboi, korichnevyi, sinii, rozovyi, raznotsvetnyi, chernen'kii, zolotistyi, prozrachnyi, bagrovyi* (they are presented in the frequency descending order).

The basic colours are at the top of the colour frequency list. *Belyi* (white) and *chernyi* (black) are used in 68 and 56 thousand of word combinations, respectively. They dominate over the other colours in the stated combinations making the centre of the colour palette. Moreover, if a comparison is made by gender, it is seen that these adjectives are more often used with the word *zhenshchina*. The adjective *seryi* (grey) follows white and black in the frequency list. However, it combines with the words under study only in 6300 cases. The word combination *seryi chelovek* (grey person) in its generic sense is used more often than the other combinations. *Seryi* (grey) combines with the word *zhenshchina* only in 187 cases and the combination *seriy muzhcjina* (grey man) was not found in the corpus. The adjective *ryzhii* (ginger) traditionally used to describe people was found in 6000 cases.

The adjectives *krasnyi* (red) and *tsvetnoi* (coloured) are used in 2559 and 2446 cases, respectively. It should be noted that sometimes the word *krasnyi* means beautiful and there is a word combination *tsvetnaia zhenshchina* but there is no combination *tsvetnoi muzhchina* (coloured man) in the Russian corpus. *Zheltyi* (yellow) occurs in 1200 cases and does not combine with the word *muzhchina*. Starting from *zelenyi* (green), the use of colours in the studied word combinations significantly decreases. *Zelenyi* is used in 766 cases, mostly with the word *chelovek* and never with the word *muzhchina*. *Goluboi* ((light blue) (663 cases)), *korichnevyi* ((brown) (487 cases)) and *sinii* ((blue (409 cases)) can combine with a *person* but not with *man* or *woman*. The adjective *rozovyi* (pink) is used in 344 cases and can combine with all the words under study. *Raznotsvetnyi* ((multicoloured (328 cases)) and *chernen'kii* ((blackish (196 cases)) combine with *chelovek* (person) and *zhenshchina* but not with *muzhchina*. There is no word combination *nebelyi chelovek* (not white person) in the corpus.

Without going into details of semantics, it can be concluded that the core of the colour palette in this context is represented by the basic colours *chernyi* and *belyi* (black and white), closer to the center are *seryi* and *ryzhii* (grey and ginger). *Krasnyi*, *tsvetnoi* and *zheltyi* (red, color and yellow) are the periphery of the palette, and the remaining colours are the periphery of the periphery. Moreover, colour adjectives are more often used with the generic word *chelovek* than with the other words under study. The use of colour adjectives is influenced by the gender of the nouns they combine with. This is due to word semantics and perception of the world.

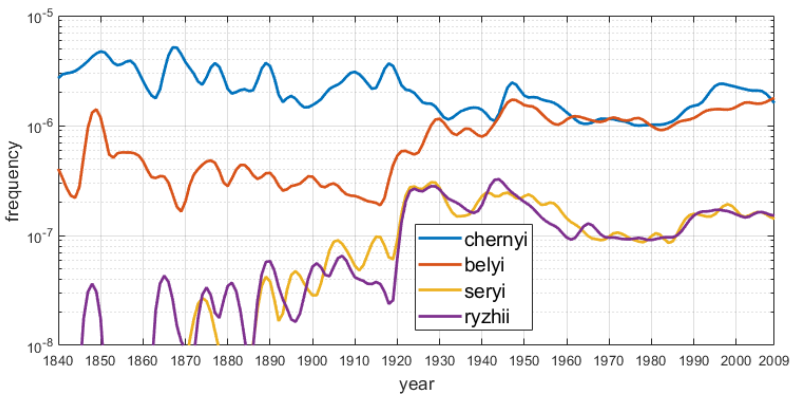

Fig. 1. Frequency dynamics of the colour adjectives

The dynamics of the most frequent colours combining with the studied words was analysed. It was shown that conventional colours are naturally appear at earlier time intervals. The black colour always dominate over the rest ones. However, a small decrease of its frequency is observed. Frequency of *belyi* (white) has been growing after 1920. A smooth growth of *seryi* (grey) with a peak in the 30s of the 20th century is observed, then it decreases slightly and stabilizes. The word *ryzhii* (ginger) that describe the nouns under study, appeared later than the other frequently used colours.

Its use grows until the 40s of the 20th century and only slightly decreases by modern time. Dynamics of the most frequent colour adjectives is shown in Figure 1.

2.2 Appearance and stature

The corpus contains approximately 400 adjectives that describe appearance and stature of *person*, *man* and *woman*. *Person* is essentially an object name. Therefore, it is not surprising that the list of 100 most frequent words of this group included adjectives that describe a person`s stature (*tolstyi, khudoi, sutulyi*), size and height (*nebol'shoi, prizemistyi*) and colour of body parts (*chernokozhii, chernoglazyi*).

The analysis of this group of words showed that it includes a list of adjectives that describe a person`s body parts or indirectly relate to them (for example, *lysyi* (bold) means a person who lacks hair).

One of the tasks was to reveal which part of the human body is the most productive one for the adjectives that describe the studied nouns. It was found that the most significant part of the human body in this case was hair, its color, condition and its presence/absence. *Sedoi* (grey-haired) and *lysyi* (bold) were the most frequently used adjectives. They are followed by the adjectives that include the word *boroda* (beard) and describe its presence/absence and colour.

It is interesting that the word *sedoi* (grey-haired) is used twice as frequently with *zhenshchina* (woman) than with *muzhchina* (man). *Lysaia zhenshchina* (bold woman) were not found in the corpus but the combination *borodataya zhenshchina* occurs in 540 cases. Human eyes are the second "productive" body part. The eye colour is the most relevant feature. The list of eye adjectives includes words that denote the number of eyes (*odnoglazyi, bezglazyi*), their form and shape (*pucheglazyi, uzkoglazyi, kosoglazyi*). It should be noted that the group of adjectives that has the word *eye* included some adjectives with figurative meaning such as *ostroglazy* and *bystroglazyi*.

There are also many adjectives describing a person`s skin and face. Complexion is described more often (*blednyi, rumianyi, zheltolitsyi*). The face itself is most often described in terms of form and shape (*shirokolitsyi, uzkolitsyi*). There are 11 adjectives including the word *nose*. The most frequent of them are *gorbonosyi, nosatyi, kurnosyi, dlinnonosyi*. The list also included the words describing moustache, shoulders, legs, head, cheeks, cheekbones, chest, forehead, eyebrows, teeth, stomach, arms, lips, color, hips, and even tail and balls.

A person`s appearance can be described using different language means and constructions. One can say *bol'sheglazyaya zhenshchina* (a big-eyed woman) or *u zhenshchina bol'shie glaza* (the woman has big eyes). However, the analysis of the given attributive structures showed that hair, eyes, skin, complexion and nose are the most significant when describing people.

If one takes into account the frequency of the given attributive constructions, the prototypical man and woman will look as follows in the 19th and 20th centuries (see Table 1). The numbers in Table 1 denote the percentage from the total number of word combinations that include the words *man* and *woman* and one of the adjectives from the “appearance and statute” group.

Table 1. Frequency portraits of *man* and *woman* in the 19th and 20th centuries

Man	XIX, %	XX, %	Woman	XIX, %	XX, %
Sedoi	1,04	4,39	Sedaja	0,65	9,26
Tolstyi	7,80	3,59	Golaja	11,18	8,58
Khudoi	2,30	3,04	Tolstaja	14,92	8,58
Krepkii	0,15	6,87	Hudaja	6,60	4,73
Borodatyi	9,51	8,83	Krupnaja	0,42	2,61
Krupnyi	0,97	5,20	Hudoshhavaja	6,60	2,63
Khudoshchavyi	8,77	4,51	Smuglaja	2,01	2,45
Plotnyi	20,80	6,78	Blednaja	6,76	2,90
Smuglyi	4,23	1,56	Nizkaja	2,73	0,32
Lysyi	0,37	2,65	Chernovolosaja	1,82	2,67
Tuchnyi	2,15	1,73	Belokuraja	3,31	3,12
Chernovolosyi	2,97	1,49			
Usatyi	0,30	3,00			
Plechistyi	3,79	2,29			
Zdorovennyi	4,53	2,50			

2.3 Age and "temporal relevance"

Age and time are important categories for human beings since perception of the world is limited to certain time periods. The third group included adjectives denoting specific age of a person (*deviatiletnii* (nine years old)) and adjectives that indicate a person's position on the timescale (*drevnii, sovremennyi, molodoi, pozhiloi, vzroslyi* - ancient, modern, young, old, adult). The most frequent adjectives were *molodoi, pozhiloi, vzroslyi, sovremennyi, staryi, drevnii, nemolodoi, prestarelyi.* If not considering the words *molodoi/pozhiloi* (young/elderly), the scale of a person`s perception in Russian culture is shifted to an older age (according to this list).

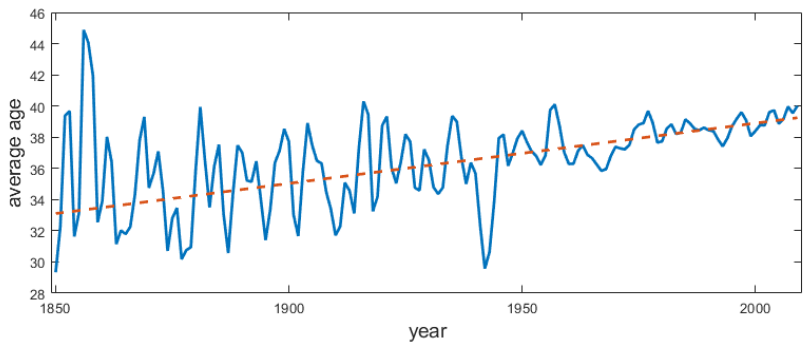

Fig. 2. Dynamics of the average age described by the adjectives

The most frequent and, therefore, most significant age adjectives attributed to the nouns under study were *sorokaletnii* (40 years), 30 *tridtsatiletnii* (30 years), shesti-

desiatiletnii (60 years), *piatidesiatiletnii* (50 years) and *dvadtsatiletnii* (20 years). The following ages are mentioned in the frequency descending order 70, 45, 25, 10, 80, 26, 22, 27, 7, 5, 19, 12, 18, 90, 23, 42, 24, 28, 4, 3, 15 and 9.

Age adjectives dynamics was analysed from quantitative and qualitative point of view. The frequency use of adjectives describing a specific age of a person was quantified (see Fig.2). It was shown that a person`s average age increases in the texts. This can be due to the increasing life expectancy and general changes in perception of human age.

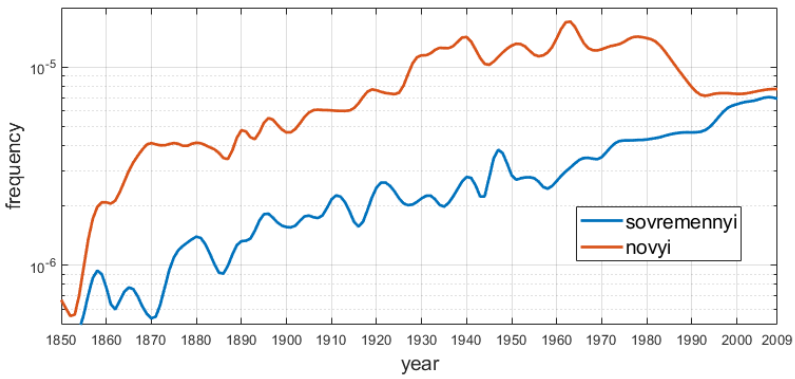

Fig. 3. Frequency dynamics of the time adjectives

If talking about qualitative dynamics of the adjectives, the graphs of the adjectives *sovremennyi*, *novyi*, *griadushchyi* seem to be interesting (see Fig.3). The frequency of the adjective *sovremennyi* has been increasing since 1850. This may indicate that modern people who live in the time of constant modernisation do not evaluate the progress as something unconditional and clearly contrapose themselves to the past generations on the timescale. The adjective *novyi* (new) has a general tendency to increase and shows a peak in the mid-1960s, slightly decreasing since that time and reaching a plateau in the early 1990s. Such dynamics can probably be due to political trends in the country. The idea of a new man (person) is widely discussed in the time of the abolition of serfage, the October revolution, the Great Patriotic War and the Reconstruction. However, this idea becomes less acute after the Reconstruction. The adjective *griadushchii* (upcoming) came into use after the revolution. It is less frequent than the two previously discussed adjectives of this group. However, its dynamics clearly responds to political events in the country. The peaks are observed after the revolution and after the end of the Great Patriotic War. Its smooth growth has been observed since 1980s, the time of new political trends in the USSR.

2.4 Opinion adjectives

Opinion adjectives is the largest group of words that combine with the nouns under study. Opinion adjectives can have either positive or negative connotation and de-

scribe different aspects of a person. For example, his mental abilities, ability to percept new information, inner world, attitude to other people and others. The group of opinion adjectives included adjectives reflecting subjective opinion about a person and describing a person`s traits of character.

The most frequent opinion adjectives are *dobryi, khoroshii, umnyi, zamechatel'nyi, poriadochnyi, sil'nyi*. All of them denote positive traits of character. The word *dobryi* (kind) heads the list of a person`s traits of character. It combines with the word person approximately 260 thousand times. It seems interesting to mention that the combination *dobraia zhenshchina* (kind woman) was found in 17 thousand cases but the word *dobryi* combines with the word *muzhchina* only in 843 cases. In other words, kindness as a trait of character is typical of Russian people (in general sense). However, kind man is not a typical word combination in Russian. The same situation is observed with the adjective *khoroshii* (good) which is also at the top of the list.

The most frequent adjectives with negative connotations are *zloi* (malicious), *plokhoi* (bad) and *glupyi* (silly). Moreover, these adjectives combine with the word *zhenshchina* several times more often than with the word *muzhchina*.

In general, opinion adjectives combine with the word *zhenshchina* more frequently. However, a short list of opinion adjectives which combine only (or more often) with the word *muzhchina* were extracted from the corpus. They are *sil'nyi, chestnyi, trudosposobnyi, vidnyi, rabotosposobnyi, strastnyi, stepennyi, respektabel'nyi, zhelchnyi, impozantnyi, bravyi, razviaznyi, galantnyi, frantovatyi* etc.

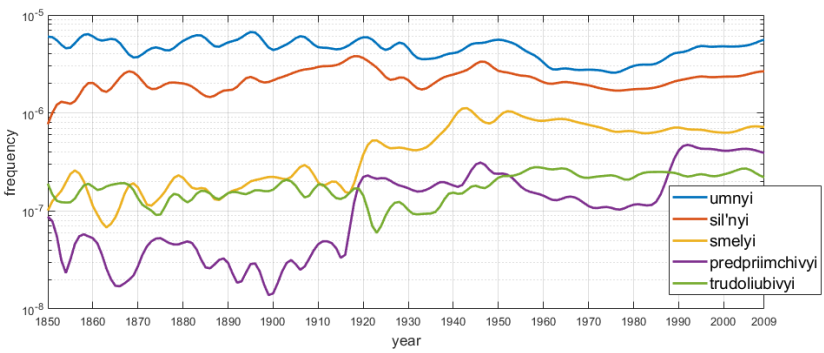

Fig. 4. Frequency dynamics of the opinion adjectives

The following results were obtained during the diachronic analysis of the opinion adjectives (see Fig. 4, 5). The general trend of the words *umnyi* (clever) and *sil'nyi* (strong) does not show high peaks. However slight growth of their use is observed in the war and post-war times when strength and ingenuity are relevant.

The frequency of the word *smelyi* shows sharp increase after the October revolution, reaches the peak in the time of the Great Patriotic war, slightly decreases and reaches the plateau. In other words, its frequency grew during significant political events and is still high today. The adjective *trudoliubivyi* (hard-working) is most frequently used in the war and post-war times when the destroyed country had to be

restored by all means. This word is still frequent at the beginning of the 21st century because the idea of personal success becomes very important at this time and no success can be achieved without one's own efforts.

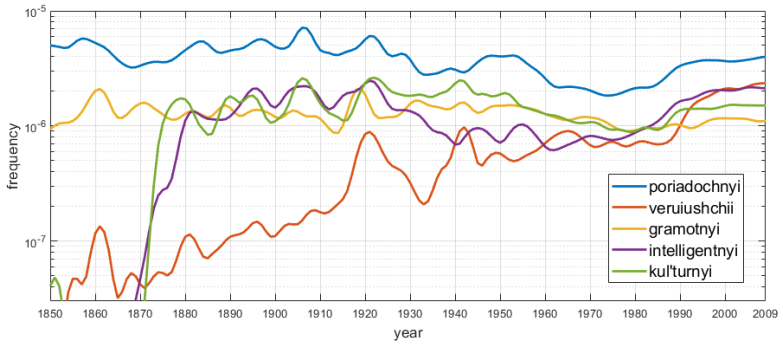

Fig. 5. Frequency dynamics of the opinion adjectives

The rapid growth of the word *predpriimchivyi* (enterprising) is observed after the revolution and has two large peaks - in the post-war period and the beginning of the 1990s, responding to social events in the country.

The graph of the word *veruiushchii* (believer) is also interesting. In general, its frequency grows. The peaks are observed in the post-revolutionary period and during the Great Patriotic War. Its frequency has been growing rapidly since the 1990s of the 20th century. In other words, the issues of belief are very acute in the period of social cataclysms and after the collapse of the USSR.

According to the corpus, the words *intelligentnyi* (genteel) and *kul'turnyi* (cultured) appeared in the Russian texts at the end of the 19th century. After the 30s of the 20th century, there is some decrease in the frequency of the adjective *intelligenthyi* which can probably be due to bourgeois perception of intelligence. However, its frequency starts growing after the 1980s. The word *kul'turnyi* has a similar graph. The adjective *gramotnyi* (literate) has a peak after the revolution since this issue becomes relevant in connection with the fight against illiteracy. The frequent word *poriadochnyi* (trustworthy) shows increase in the years of the October revolution, post-revolution and the wartime. Then, there has been some decrease since the 60s and the frequency has been growing since the 90s of the 20th century.

Summing up, it should be said that significant social and political events can influence the presentation of a person`s traits of character and his perception in written texts.

2.5 Emotional and physical condition/social status

Adjectives that describe emotional and physical condition of a person and his social status were included in this group. The most frequent adjectives were *zdorovyi, svobodnyi, bol'noi, odinokii, semeinyi, mertvyi, p'ianyi, bezdetnyi* etc.

The adjectives *schastlivyi/neschastnyi* (happy/unhappy) are also among them. At that, *schastlivyi* is used almost twice as much as *neschastnyi*. It is also interesting to mention *that schastlivyi/neschastnyi* combine with *woman* several folds more often than with *man*. There are 8 derivatives of the word *schastlivyi*, three with positive and five with negative connotations. However, total frequency of the words with positive connotation is higher.

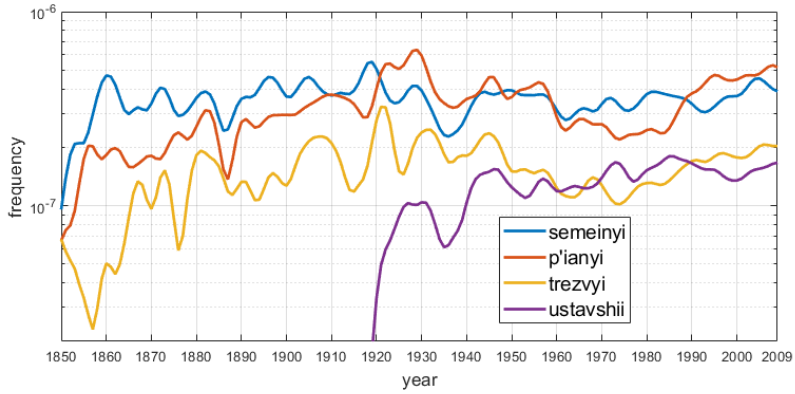

Fig. 6. Frequency dynamic of the emotional/physical condition and social status adjectives

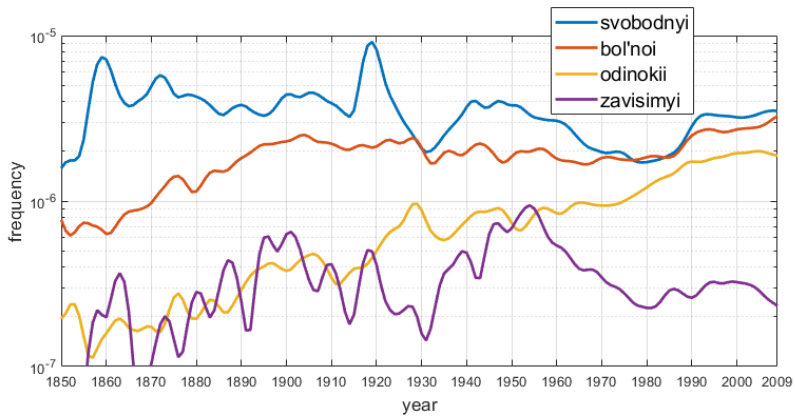

Fig. 7. Frequency dynamics of the emotional/physical condition and social status adjectives

The diachronic analysis of the most frequent and interesting (in our opinion) adjectives (see Fig. 6, 7) allowed us to obtain the following results. The issue of drunkenness has always been acute in Russia. Since 1980, there has been a consistently high use of the word *p'ianyi* (drunken). One frequency peak of this adjective is observed in the late 20s of the 20th century and two peaks are observed in the post-war period. The frequency of *p'ianyi* decreases between the 60s and 80s of the 20th century when the government struggles against alcoholism in the USSR. However, the frequency tends to grow after this period.

The graph shape of the antonym *trezvyi* (sober) resembles the one of *p'ianyi* but it differs at the end of the 1920s. In other words, the amount of *trezvyi* becomes less with the growth of *p'ianyi*. Besides, it should be noted that *trezvyi* can have figurative meaning. The adjective *semeinyi* (married) has always been acute for Russia. However, there is a gap in the mid-30s of the 20th century. The adjective *ustavshii* (tired) describes a person`s physical state. It has appeared in the corpus since 1920 and its frequency tends to grow. Thus, men and women become more tired in the texts.

The adjective *svobodnyi* (free) often combines with the given nouns. Its frequency peaks are observed in the late 50s and early 60s of the 19th century (when the issue of the abolition of serfage was acute), during the revolution of 1917 and the Great Patriotic War, as well as in the early 1990s when human rights and freedoms were widely discussed in the USSR. The frequency of the adjective *bol'noi* (ill) tends to grow in the combination with the given nouns. After the 1990s, it increases even more. It is also interesting to mention that frequency of the word *odinokii* (lonely) has a general trend to grow. The frequency peak is observed at the end of the 1920s and a its steady growth has begun since 1970s. The graph of the adjective *zavisimyi* (dependent) is heterogeneous and shows peaks in different years. The highest of them are observed in the period of the October revolution and the mid-1950s.

3 Discussion

The article is based on a computer method of processing big data which has been widely used for the analysis of text corpus data. The problem of combinability was studied in [4] and [7]. Studies of collocability of specific words was performed in [17]. This article studies a combination of the *Adj.+Noun* (man, woman, person) type.

Adjectives have always been interesting for linguists. Different semantic classifications of adjectives have been introduced in [14, 16]. The obtained adjectives were divided into 5 groups in this work. However, it should be noted that any classifications can`t be strict as the language itself.

The obtained data show us one of the facets of linguistic perception of a person in Russian culture. Moreover, this perception is associated with gender differences. Some adjectives combine only with man and (or) woman. Reflecting the realty, the use of the studied constructions can correlate with significant social or political events which happen in the country.

Conclusion

The article studied attributive constructions using computational methods. Namely, Russian adjectives that combine with the word forms of *person, man* and *woman* were considered. The Google Books text corpus served as a study material. The Open Corpora morphological dictionary was used to verify parts of speech.

There were approximately 17 thousand adjectives obtained during the analysis which were divided into 5 groups. The center and periphery of the colour palette of the studied colour adjectives were identified.

The analysis of adjectives describing appearance and physical parameters of a person showed that hair, eyes, complexion, body skin, face and nose are significant for the Russian culture while describing a person`s appearance. The obtained data allowed us to make a frequency portrait of a typical man and woman of the 19th and 20th centuries.

Dynamics of the age and time adjectives was analysed and showed increase in person`s average age, in other words the timeline of human perception shifts to the right in the Russian culture. The largest group was opinion adjectives. Adjectives with positive connotation were at the top of the frequency list. According to the corpus texts, the perception of a person in Russian culture is positive. The adjective *dobryi* (kind) is the most frequent one in this group. Some gender differences in the use of the adjectives were revealed.

Acknowledgements. This research was financially supported by the Russian Government Program of Competitive Growth of Kazan Federal University, and by RFBR, grants № 17-29-09163.

References

1. Wierzbicka, A.: Lexicography and conceptual analysis. Karoma, Ann Arbor (1985).
2. Wierzbicka, A.: The semantics of grammar. Benjamins, Amsterdam (1988).
3. Melchuk, I.A.: Kurs obshchei morfologii. Tom 1. Progress, Moskva (1997). [in Russian].
4. Rakhilina, E.V.: Kognitivnyi analiz predmetnykh imen: semantika i sochetaemost'. Russkie slovari, Moskva (2000). [in Russian].
5. Rakhilina E.V.: Lingvistika konstruktsii. Azbukovnik, Moskva (2010). [in Russian].
6. Wierzbicka, A.: Semantics of color terms: cultural and cognitive aspects. Cognitive linguistics 1(1), 99-150 (1990).
7. Rakhilina, E.V.: Linguistic construal of colors: the case of Russian. In: MacLaury, R.M., Paramei, G.V., Dedrick, D. (eds.). Anthropology of Colour: Interdisciplinary Multilevel Modelling, pp. 363-377. Benjamins, Amsterdam (2007).
8. Rosch, E.: Human categorization. In: Warren, N. (ed.) Advances in cross-cultural psychology, vol. 1, pp. 1-49. Academic press, N.Y. (1977).
9. Michel, J.-B., Shen, Y.K., Aiden, A.P., Veres, A., Gray, M.K. et al.: Quantitative Analysis of Culture Using Millions of Digitized Books. Science 331(6014), 176-82 (2011).
10. Bocharov, V.V., Alexeeva, S.V., Granovsky, D.V., Protopopova, E.V., Stepanova, M.E., Surikov, A.V.: Crowdsourcing morphological annotation. In: Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue", 12 (1), pp. 109-115. RGGU, Moskow (2013).
11. Khristoforov, S., Bochkarev, V., Shevlyakova, A.: Recognition of Parts of Speech Using the Vector of Bigram Frequencies. In: Aalst, W., et al (eds.) Analysis of Images, Social Net-works and Texts 2019, CCIS, vol. 1086, pp. 132-142. Springer, Heidelberg (2020). DOI: 10.1007/978-3-030-39575-9_13.
12. Bolinger, D.: Adjectives in English: Attribution and predication. Lingua 18, 1-34 (1967).
13. Givon, T: Notes on the Semantic structure of English Adjectives. Language 46(4), 816–837 (1970).
14. Dixon, R.M.W.: Where have all the adjectives gone? Studies in language 1, 19–80 (1977).

15. Shramm, A.N.: Ocherki po semantike kachestvennyh prilagatel'nyh: na materiale sovremennogo russkogo jazyka. Leningrad State University, Leningrad (1979). [in Russian].
16. Grashchenkov, P.V., Kobozeva, I.M.: Semantic Classes and Government of Adjectives. In: Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue", 16(2), pp. 134-149. RGGU, Moskow (2017).
17. Zakharov, V. P.: Set phrases: A view through corpora. In: Computational Linguistics and Intellectual Technologies. Papers from the Annual Inter-national Conference "Dialogue", 14(1), pp. 667-682. RGGU, Moskow (2015).