Записки научных
семинаров ПОМИ
Том 529, 2023 г.

**V. V. Bochkarev, V. D. Solovyev, T. A. Nestik,
A. V. Shevlyakova**

VARIATIONS IN AVERAGE WORD VALENCE OF RUSSIAN BOOKS OVER A CENTURY AND SOCIAL CHANGE

Abstract. Valence of words in books reflects the situation in a society and allows one to assess the perception of life even in those countries and periods of time when direct research on well-being has not been conducted. We use the Google Books Ngram diachronic text corpus to analyze changes in the average valence of words in Russian books. We show that changes in the average valence correlate with the results of surveys on well-being. The average valence also responds to major historical events and social changes. Like other similar studies, quantitative data on the level of valence of words are based on calculations using dictionaries with word valence ratings. For the first time, we have carried out a comparative study based on a number of the most relevant Russian dictionaries. We have found that the obtained results depend on the applied meaning of such dictionaries and their lexical composition. This shows the need for careful selection of dictionaries for future research.

§1. Introduction

The level of national well-being is one of the most important sociological parameters. Its estimation, as well as estimation of the level of a closely related concept of happiness, has been performed for many decades both at national and international levels. Until recently, sociological surveys were the main research tool, which were used to determine the level of subjective perception of well-being (happiness).

The creation of large diachronic text corpora has given rise to a new method for studying the dynamics of percepted well-being. The method consists in revealing frequency dynamics of the use of emotive words and, more broadly, words with positive or negative valence. This line of research is based on the assumption that the degree of positivity of the used

Key words and phrases: word valence ratings and Google Books Ngram and word frequency and subjective well-being and sentiment dictionaries.

words directly reflects the authors' mood [31], i.e., their assessment of well-being. The main data source is the Google Books Ngram (GBN) corpus (https://books.google.com/ngrams/).

The Google Books Ngram corpus [24] contains data on frequencies of words and phrases from 8 languages over the past five centuries. The corpus is widely used in cultural and language evolution studies [14,44]. The third version of the Russian subcorpus is based on texts of more than one million books published between 1486–2019 with a total size of 89.4 billion words. It should be noted that GBN does not give access to texts but provides statistics on use of words and word combinations for a particular year.

The objective of our study is to also analyse how the method described in [15] and used in our work can be applied to valence studies based on Russian texts. To do this, we analyse the frequency dynamics of Russian words with positive/negative valence (including basic emotions) by using GBN and reveal how average valence responds to historical events, social changes, and economic factors such as the gross domestic product (GDP). We use words from the Russian sentiment dictionaries and obtain frequencies of these words from the GBN corpus. For the first time, we compare results obtained by employing several sentiment dictionaries.

§2. Related Work

Relevant studies on word valence and subjective well-being have been developing in two directions. The first one is aimed at studying pure dynamics of words with positive/negative valence and lexical emotiveness. Frequency dynamics of positively marked words has attracted the attention of researchers after the work by [10], who formulated the Pollyanna hypothesis (also known as Positive bias), according to which "people use positively toned words more often than negatively toned ones". In [18], the hypothesis was confirmed based on 4 large corpora of the English language. A diachronic study based on GBN [16] showed that the positive bias decreases with time, and the trend is well approximated by a linear law. This paper also shows that the linguistic positive bias correlates with the level of happiness in a country. Frequency dynamics of emotive words have been studied in many works [12, 28, 34, 41, 46].

When studying the frequency dynamics of emotive words, basic emotions are usually considered, and word lists are selected from special databases and dictionaries of synonyms based on psychological works on emotions. For example, the work [28] uses the digital resource “Linguistic inquiry and word count” introduced in [35] that contains over 900 emotive English words.

Another research direction is the study of how words with positive/negative valence respond to social factors and reflect the national mood and subjective well-being. Well-being dynamics in different countries have been studied in [6, 32, 37, 43].

The work [15] introduced the concept of the National Valence Index (NVI) — the average value of word valence considering word frequencies in a particular year of the studied interval. Affective ratings are taken from the well-known Affective Norms for English Words (ANEW) database [11] containing about 1000 words. The word frequencies are derived from GBN. This work shows that for 3 languages in 4 countries, namely United States, United Kingdom, Germany, and Italy, the NVI is highly correlated with survey-based measures of subjective well-being obtained by the Eurobarometer[1]. The response of the NVI to social upheavals is discussed. It is shown that in the long term, the growth of gross domestic product (according to the data from the Maddison Project) has no correlation with the NVI. However, significant local changes in GDP are accompanied by an increase in the NVI, although the correlation coefficient is not that significant.

The work [9] does not consider individual words but entire language constructs that, according to expert assessment, can indicate cognitive distortion. It is shown that an increase in the number of such constructions in the GBN books responds to negative social events, such as World War I and World War II. In this work, an unexpected effect of a sharp increase in the number of such constructions in three different languages over the last 3 decades was discovered. The authors consider that this effect is probably caused by a growing gap between increased productivity and slower wage growth, rising social inequality, as well as the polarization of societies that was further exacerbated by the rise of social networks.

[1]`https://europa.eu/eurobarometer/screen/home`

§3. Data and Methods

In the above-mentioned works, indices of word valence were taken from dictionaries created on the basis of questionnaires. A dictionary [27] of almost 20,000 words is currently used for English studies. An overview of English dictionaries can be found in [20]. There also exist dictionaries for other languages such as Russian, Danish, Spanish, German, Finnish, and Chinese. In our work, estimations of word valences are taken from several dictionaries (databases) listed below.

(1) *KFU Sentiment Dictionary* of the 1000 most frequent words created at Kazan Federal University using the questionnaire method based on the dictionary by O. Lyashevskaya, S. Sharova (http://dict.ruslang.ru/freq.php). It contains equal proportions of nouns, adjectives, and verbs. Each word received at least 50 ratings on a 9–point scale.
(2) *KFU Sentiment BERT Dictionary* of more than 25,000 words created at Kazan Federal University using the machine extrapolation method (the BERT neural network). Both KFU dictionaries are freely available at [2] and are described in detail in [42].
(3) *Russian NRC VAD Lexicon* [27] which contains estimates of 20,000 words obtained by using the estimates of the corresponding (synonymic) English words. The English words were the most commonly used ones extracted from various sources.
(4) The article [13] describes dictionaries of 10 languages, including the *Russian Dictionary* of 10 thousand words available at [1]. The most commonly used words were selected for estimation from a number of sources: Google Books, New York Times articles, Music Lyrics, Twitter messages, etc. The survey was conducted via Amazon Mechanical Turk on a 9-point scale.
(5) *LinisCrowd Dictionary*[2] of sentiment tokens created via crowdsourcing [19] that includes 7.5 thousand words. Each word was estimated by 3 experts on a scale of -2, -1, 0, 1, 2.
(6) *RuSentiLex* that contains 12 thousand words extracted automatically from the news corpus. Evaluative words and words with positive or negative connotation were assessed by experts on a 3–point scale [25].

[2] http://linis-crowd.org/

(7) *KartaSlovSent Dictionary* (https://kartaslov.ru) created by a questionnaire method using the positive, neutral, negative scale and with at least 25 ratings for each word. It contains more than 46 thousand words [21].

We note that these resources are created using different techniques and contain different sets of words. The applied methods include interviewing native speakers (dictionaries 1, 4, 5, 7), machine approximation of a small number of human estimates (2), translation of English words into Russian and then using the ratings of English words (3), automatic extraction from the text corpus (6). The latter case uses an approach combined with post-editing of the dictionaries.

Word sets are determined by the applied meaning of a dictionary and are also different. The most frequent words are used in 1, 2, 4, the most commonly understood words are included in 7, words selected by a combined approach with an emphasis on emotive words are in 3, while 5 and 6 contain mostly evaluative words.

As stated in [20], most dictionaries have a certain percentage of stop words (*Hedonometer* stands out in this regard). In our opinion, the interpretation of sentiment ratings of prepositions, conjunctions and other functional words is unclear. In our calculations, we tried to exclude stop words from the lists. The lists included nouns, adjectives, verbs, and adverbs (content words). Relation to a particular part of speech was determined using the electronic morphological dictionary OpenCorpora (http://opencorpora.org/dict.php) [8]. It should also be noted that some words in the dictionaries are duplicated. In the dictionaries obtained by the translation method, such duplication occurs because different English words can be translated by the same Russian synonym. In such cases, we averaged ratings for different entries of the word.

After the filtering described above, the word list included 55,590 words that occur in at least one of the 7 compared dictionaries. Statistics on pairwise identical part of the vocabulary in dictionaries is shown in Table 1. For example, the number of words in the *NRC VAD* and *KartaSlovSent* dictionaries (after filtering) is 13,694 and 42,559 respectively, and the number of identical words is 9,499. This is 22.3% of the *KartaSlovSent* dictionary or 69.4% of the *NRC VAD* dictionary. Data on the correlation of ratings of identical words in different dictionaries are shown in Table 2. The correlation coefficient of the ratings of the *KFU Sentiment* and *KFU Sentiment BERT* dictionaries equals 1 because the first one includes the second one.

Table 1. The ratio of the identical part of the vocabulary in the dictionaries to the size of the dictionaries in a line (%).

	NRC VAD	Hedonometer	KFU Sentimen	KFU Sent. BERT	RuSentiLex	LinisCrowd	KartaSlovSent
NRC VAD	100.0	12.9	4.7	58.4	19.2	22.8	69.4
Hedonometer	21.4	100.0	7.9	23.8	3.9	15.1	23.6
KFU Sentiment	64.4	65.9	100.0	99.8	17.1	68.3	95.1
KFU Sent. BERT	33.1	8.1	4.1	100.0	39.6	24.0	85.2
RuSentiLex	27.5	3.3	1.8	100.0	100.0	29.4	82.6
LinisCrowd	46.7	18.5	10.1	86.3	42.0	100.0	83.9
KartaSlovSent	22.3	4.6	2.2	48.4	18.6	13.2	100.0

Table 2. Spearman's correlation coefficient of ratings of identical words in different dictionaries.

	NRC VAD	Hedonometer	KFU Sentimen	KFU Sent. BERT	RuSentiLex	LinisCrowd	KartaSlovSent
NRC VAD	-	0.773	0.720	0.671	0.762	0.706	0.746
Hedonometer	0.773	-	0.855	0.735	0.875	0.614	0.797
KFU Sentiment	0.720	0.855	-	1.000	0.808	0.526	0.752
KFU Sent. BERT	0.671	0.735	1.000	-	0.609	0.600	0.636
RuSentiLex	0.762	0.875	0.808	0.609	-	0.762	0.797
LinisCrowd	0.706	0.614	0.526	0.600	0.762	-	0.779
KartaSlovSent	0.746	0.797	0.752	0.636	0.797	0.779	-

Table 3. Average well-being index in Russia according to the RSS surveys.

Year	2006	2008	2010	2012	2014	2016	2018
Percentage	43	48	52	54	60	54	53

Table 4. Average well-being index in Russia according to [17].

Year	1990	1995	1999	2006	2008	2011	2017
well-being index	5.37	4.45	4.65	6.15	6.50	6.13	6.45

We considered both dictionaries because they were created using different methods, differ in size, and behave differently when compared to other dictionaries.

Frequency data on the use of Russian words were taken from the Russian subcorpus of the 3rd version of GBN. For the i-th word, we calculated its relative frequency $p_{i,t}$ in the year t. To do this, the absolute frequency $f_{i,t}$ extracted from the corpus was normalized to the total frequency of all rated words in year t. The average valence $\langle \nu_t \rangle$ (or NVI_t) was then calculated by the following formula [15]:

$$\langle \nu_t \rangle = NVI_t = \sum_i \nu_{i,t} p_{i,t} \tag{1}$$

Here ν_t is the valence rating of the i-th word in the year t.

Data on Russia's gross domestic product (GDP) are taken from [3, 26, 40]. Life satisfaction survey data comes from two sources. The first one is the Russian Social Survey (RSS)[3] [39]. The survey was conducted from 2006 to 2018 with an interval of 2 years. The survey used an 11-point scale from 0 to 10, where 10 means complete satisfaction with life. The question (b24 of the 2006 survey) was formulated as “How satisfied are you with your life in general?”. Table 3 shows the proportion of respondents who rated their well-being in the range from 6 to 10, i.e., positively. Satisfaction with life increases until 2014, then it decreases. The work [17] provides data for a longer time interval from 1990 to 2017 (see Table 4). A scale from 1 to 10 was used for the ranking. Low values of the well-being index at the end of the last century have been replaced by significantly higher ones,

[3] `http://www.ess-ru.ru/`

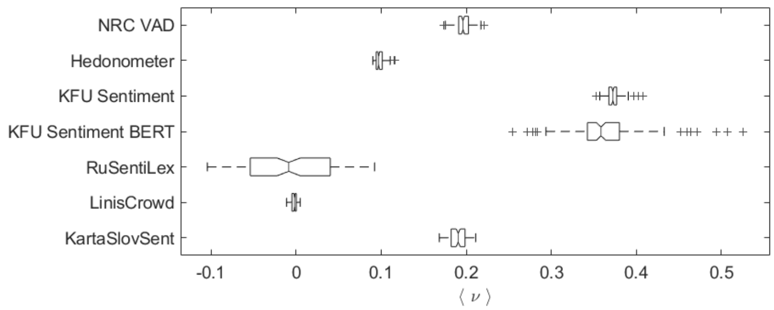

Figure 1. Ranges of change in the average valence for 7 dictionaries over 1900–2019.

reflecting the rapid economic growth in Russia in the early 2000s. After the economic crisis of 2008, the situation has gotten worse.

§4. Results

It is interesting to find out to how the words presented in the above-mentioned Russian dictionaries respond to social changes and perform a comparison of the dictionaries. Such systematic comparative analysis has never been carried out, although there exist various sentiment dictionaries created for other languages by different methods.

Time series of the average valence $\langle \nu_t \rangle$ were calculated for each of the 7 dictionaries. To make the comparisons more convenient, the ratings presented in each dictionary were linearly transformed so that the ratings ranged from ˘1 to +1. In different dictionaries, the ratio of the number of words with positive and negative ratings is different, and therefore the characteristic ranges of change turn out to be different (see Figure 1).

Since the typical values of $\langle \nu_t \rangle$ for different dictionaries vary significantly, it is reasonable to normalize them for a better visual representation. To do this, the average value (for 1900–2019) was subtracted from the values $\langle \nu_t \rangle$ and further divided by the standard deviation (for the same time interval). The resulting normalized series are shown in Figure 2.

One of the main results of our study is that for different vocabularies we obtain different valence dynamics curves. Visually, all the considered dictionaries are clearly divided into two clusters that differ significantly. The first one includes *NRC VAD*, *RuSentiLex*, *LinisCrowd*, and *KartaSlovSent*,

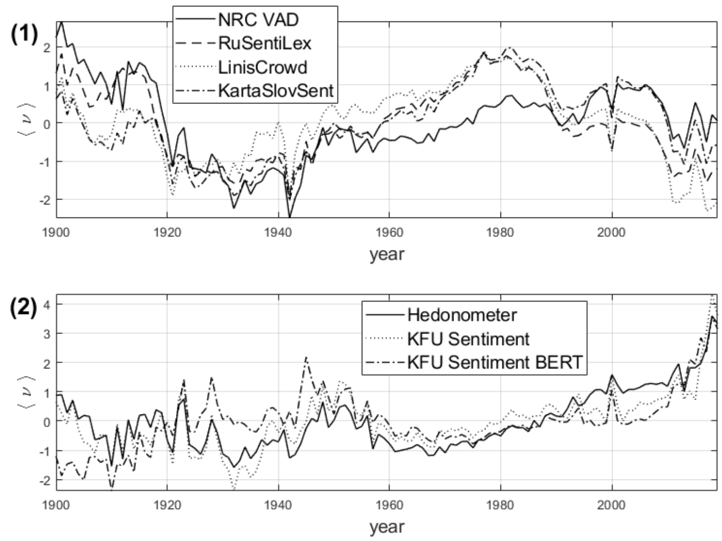

Figure 2. Change in the normalized mean valence over time for 7 dictionaries.

and the second one includes *Hedonometer*, *KFU Sentiment*, and *KFU Sentiment BERT* with very similar curve shapes (see Figure 2). Further we will refer to them as the 1st and 2nd clusters.

As expected, all dictionaries show a decrease in positivity during World War II. A similar effect was previously found for other languages. Note that the response of the curves calculated for the words from the dictionaries of the second cluster is significantly weaker.

Let us pay attention to historical periods characterized by significant social events in Russia. During the Revolution and Civil War of 1917–1922 dictionaries of the 1st cluster record a drop in positivity. The response of the words from the dictionaries of the second cluster to these events is not significant.

Words from the dictionaries of the 1st cluster show a global trend towards an increase in positivity approximately from the 1930s to the 1980s (except for the period of World War II) with a subsequent decrease. It is natural to associate this with the period of development of socialism in the country, including an increase in the standard of living, the peak of which falls precisely on the year 1980 followed by a decline. The positive shift, apparently, is also influenced by the high ideological content of society and state control over printed products, which in particular prohibited to depict many negative aspects of life in the country.

Dictionaries from the 2nd cluster paint a different picture. They also show an increase in positivity from 1930 (until 1950); then they show a decrease until the mid-1960s followed by an increase. This decrease is possibly associated with the criticism of Stalin's personality cult at that time and the so-called Thaw which was characterized by relatively greater freedom. The most interesting is almost a mirrored behavior of the graphs in both clusters after 1980.

To compare the two obtained clusters, we select one most relevant dictionary from each cluster. They are the *KartaSlovSent Dictionary* which is the largest one (42,559 words; here and below we show the numbers of words after filtering out functional words) and the well-known dictionary *Hedonometer* (8,238 words). In total, 48,885 words appear in at least one of the two selected dictionaries, which we will divide into three groups: (a) 1,942 words found in both dictionaries, (b) 40,617 words found only in the *KartaSlovSent Dictionary*, and (c) 6,296 words found only in the *Hedonometer* dictionary.

The Spearman's correlation coefficient between time series of average sentiment $\langle \nu_t \rangle$ obtained using *Hedonometer* and *KartaSlovSent* dictionaries is 0.274 (Pearson's coefficient is 0.158). If we calculate $\langle \nu_t \rangle$ only for the identical part of the vocabulary, then the Spearman's coefficient increases to 0.791. This is almost equal to the correlation coefficient of sentiment ratings in the overlapping part of the dictionaries (0.797, see Table 2). If, on the contrary, each of the series $\langle \nu_t \rangle$ is counted only by the non-identical part of the vocabulary, the Spearman coefficient will be -0.007.

Figure 3 shows boxplots of the distribution of ratings in the non-identical part of the two compared dictionaries. As for the *KartaSlovSent Dictionary*, the median of ratings for this part of the vocabulary is close to zero, and the distribution for the *Hedonometer* dictionary has a significant shift towards positive values. More precisely, the ratio of the number of words with positive ratings to the number of words with negative ratings is 3.713 for the *Hedonometer* dictionary and 1.417 for the *KartaSlovSent Dictionary*. Thus, one of the reasons for the observed differences in graphs $\langle \nu_t \rangle$ may be a different proportion of words with positive and negative ratings

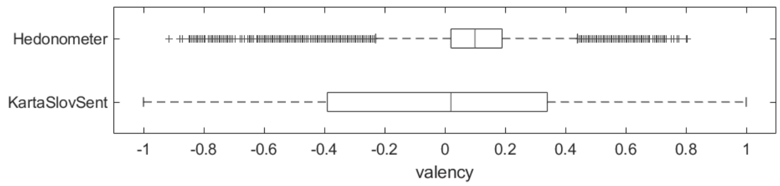

Figure 3. The range of valence ratings in the non-identical part of the *KartaSlovSent* and *Hedonometer* dictionaries.

Table 5. Spearman's correlation coefficients of the average valence of words compared with the results of sociological surveys and changes in GDP.

	Well-being, RSS 2006–2018	Well-being, [17] 1990–2017	GDP 1913–2019
NRC VAD	**-0.811**	-0.107	**0.421**
Hedonometer	0.559	0.714	**0.450**
KFU Sentiment	0.667	**0.786**	**0.522**
KFU Sent. BERT	0.721	0.750	0.134
RuSentiLex	-0.775	-0.429	**0.281**
LinisCrowd	-0.721	-0.679	**0.271**
KartaSlovSent	-0.775	-0.143	**0.589**

in the compared dictionaries. Another possible reason is that even after filtering, a large number of non-connotated words (pronominal adjectives, existential verbs etc.) remain in the *Hedonometer* dictionary. As a result, a change in the frequency of such words that actually have no positive or negative connotation will lead to a change in $\langle \nu_t \rangle$.

Finally, let us consider how changes in the average valence correlate with the results of sociological surveys on the degree of satisfaction with life [17,39], as well as with GDP. Table 5 shows the values of the Spearman's correlation coefficients between these values and the average valence calculated using the data from different dictionaries. Significant (at the level of 0.05) correlation coefficients are in bold type.

It should be noted that the comparison with the results of sociological surveys was carried out only for 7 points and cannot be considered statistically significant (in most cases) in spite of high absolute values. More reliable data on correlation with well-being can be obtained when longer series of sociological surveys are accumulated. Compared with GDP, a moderate positive correlation is observed. The highest correlation coefficient with GDP was obtained for the average valence, calculated according to the data of the *KartaSlovSent Dictionary*. Pearson's coefficient for this case is 0.595, which is noticeably higher than the value obtained in [15] (0.36). In contrast to the United Kingdom, the United States, Germany, and Italy (the data considered in [15]), the average tonality in Russian has a long-term upward trend throughout much of the XX century, which affects the value of the correlation coefficient with GDP.

The correlation of the average valence with the GDP growth rate (which is defined as the difference derivative of the logarithm of GDP, $\Delta \log \mathrm{GDP}$) was also calculated in [15]. In our case, the correlation of these values is low for short and moderate lags (less than 15 years). Figure 2 shows that there is also a drop in the average valence during periods when GDP was falling (the revolution and civil war, World War II, early 1990s), but these jumps occur each time from a different initial level. This observation explains a low correlation between the mean valence and GDP growth rates.

Therefore, we also estimated the correlation between GDP growth rates and average valence increments calculated by the following formula:

$$\Delta\langle\nu_t\rangle = \langle\nu_t\rangle - \langle\nu_{t-T}\rangle \tag{2}$$

Figure 4,a shows cross-correlation functions (Spearman's correlation coefficient was used) of average valence increments and GDP growth rates for the *KartaSlovSent*, *LinisCrowd*, and *Hedonometer* dictionaries. There, positive lag values correspond to the delay in changes in the average valence in relation to changes in GDP. As we can see, the highest level of correlation is observed with a lag of 11–13 years. This is larger than the delay value of 5 years obtained for the English language in [15].

Figure 4,b shows the values of the maximum correlation level for various dictionaries. The level of correlation for the dictionaries from the 1st cluster is slightly higher than that for the dictionaries from the 2nd cluster. The highest correlation level (0.335) is observed for the *RuSentiLex* dictionary.

A decrease in the average valence of Russian texts observed since about 1980 (for the first cluster, see Figure 2) is in good agreement with the results of corpus studies indicating a decrease in linguistic markers of

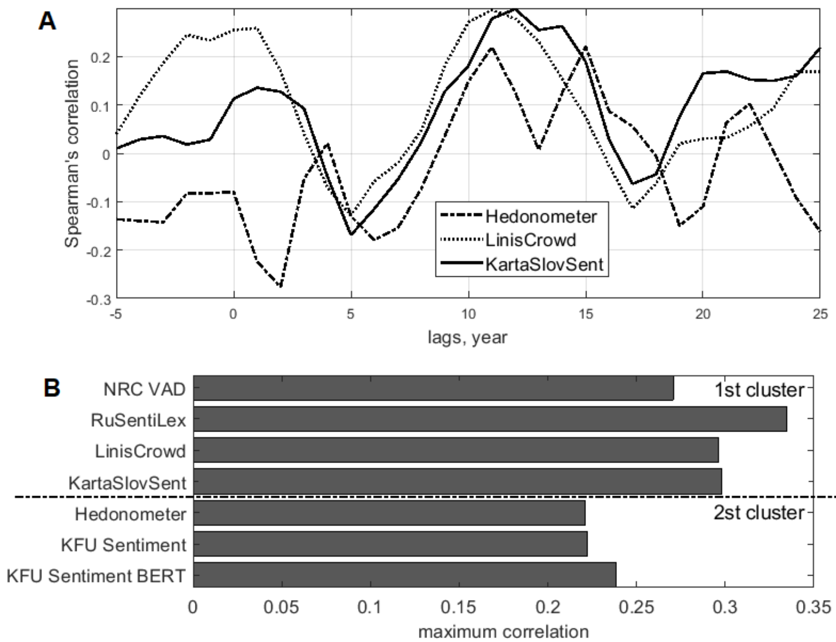

Figure 4. A) Spearman's correlation coefficient between the series of the GDP growth rates and the series of average valence increments at different lags for the *KartaSlovSent*, *LinisCrowd*, and *Hedonometer* dictionaries; B) The maximum level of the Spearman's correlation coefficient between the series of GDP growth rates and the series of the mean valence increments for various dictionaries.

psychological well-being in English, Spanish and German texts, especially pronounced since the mid-1970s [9, 15, 16]. The analysis of the valence of lyrical songs also reveals the same negative trend [29]. Further comprehensive studies are needed to explain these changes. We can put forward two assumptions that need to be tested.

The first one is associated with a decrease in the level of self-efficacy, that is, the individual's confidence in their ability to influence their present and future [5]. In developed countries, this was due to the gradual destruction of the social contract that was formed after World War II and assumed that the system of social relations and the state provide a gradual improvement in the quality of life of citizens. Disparities in economic growth and income levels for most of the population led to the collapse of the American Dream [22] and rising inequality [36].

In the 1980s, the Soviet economy ceases to outstrip the US economy in terms of growth, it finally enters stagnation with growth of about 1% per year, low efficiency, high dependence on world oil prices, as well as a demotivating reward system [4, 23]. Another factor that reduces an individual's belief in the ability to change the situation with their own or joint efforts was a transition to a risk society, awareness of limited resources, and growing exposure to global difficult-to-control threats [30].

The second possible reason could be urbanization accompanied by a decrease in prosocial attitudes and trust that, as psychological studies show, make it easier to satisfy the need for affiliation and received emotional support [33, 38]. Corpus studies point to the relationship of urbanization with the growing orientation towards individualistic values [14], the strengthening of which was observed in the USSR and accelerated in the post-Soviet period [45].

A meta-analysis of 20 studies shows that levels of intragroup trust, solidarity, and prosocial attitudes increase significantly after wars [7]. This may be one of the explanations for the sharp increase in positivity in Russian texts in the post–war years.

§5. Conclusion

National subjective well-being reflected in natural language, including published books, has long been a focal issue among researchers. The main results obtained in the present work are as follows.

(1) For the first time, we provide detailed data on the change in the average valence of words in Russian texts over a hundred years.

(2) The influence of social upheavals on text average valence is revealed. The average valence change is neither random nor determined by purely linguistic mechanisms and general cognitive laws such as the positive bias. It is definitely determined by changes in society. Such events as wars and revolutions increase the frequency of words with negative valence.

(3) We found a correlation of the average valence of words with the results of sociological surveys on the level of well-being, as well as a moderate correlation with changes in GDP. The presented data should be considered preliminary because data available from sociological surveys is insufficient, and no statistically reliable conclusions can be drawn now.
(4) For the first time in this kind of research, we used a number of dictionaries with valence ratings and compared the results obtained using these dictionaries. It was found that each dictionary provides its own result. However, according to the general data obtained, all the dictionaries can be divided into two clusters.
(5) A preliminary explanation of this effect is that dictionaries differ significantly by the vocabulary included and the proportion of words with positive and negative valence. It seems that the observed effect is very interesting and requires further detailed studies.
(6) The method offered by Hills et al [15] and used in our work requires a thorough selection of words with valence ratings. At that, balance of words with positive and negative valence is needed, otherwise this method will not provide reproducible results.

The obtained results for the Russian language based on the dictionaries from the 1st cluster are in good agreement with previously published results for other languages. A decrease in the frequency of words with positive valence has been observed not only in Russian texts but also in English, German and Italian since the last quarter of the XX century (caused by wars and revolutions).

An important final conclusion is that the above-described studies should not be based on a single dictionary because they provide different data. Joint use of several dictionaries is reasonable until a common methodology for creating dictionaries with word valence ratings is established.

Acknowledgments. This research was financially supported by Russian Science Foundation, grant No. 20-18-00206.

References

1. *Hedonometer word list: labmt-ru-v2*, https://hedonometer.org/words/labMT-ru-v2/, last accessed 2023/03/15.
2. *The technology of creation of semantic electronic dictionaries*, https://kpfu.ru/tehnologiya-sozdaniya-semanticheskih-elektronnyh.html, last accessed 2023/03/15.
3. *World economic outlook*, https://www.imf.org/en/Publications/WEO, last accessed 2023/03/15.
4. A. D. Aseev and V. V. Shishkov, *Sovetskii soiuz v lovushke "imperskogo perenapriazheniia": voprosy i problemy analiza sovetskoi politiki.* — Vestnik Moskovskogo Gosudarstvennogo Lingvisticheskogo Universiteta. Obshchestvennye nauki (2022), No. 3, 9–17.
5. A. Bandura, C. Pastorelli, C. Barbaranelli, G. Caprara, *Self-efficacy pathways to child depression.* — J. Personality and Social Psychology **76** (1999), 258–69.
6. H.-W.-S. Bao, H. Cai, Z. Huang, *Discerning cultural shifts in china? commentary on hamamura et al.* (2021)., American Psychologist **77**, No. 6 (2022), 786–788.
7. M. Bauer, C. Blattman, J. Chytilova, J. Henrich, E. Miguel, T. Mitts, *Can war foster cooperation?.* — J. Economic Perspectives **30**, No. 3 (2016), 249–74.
8. V. V. Bocharov, S. V. Alexeeva, D. V. Granovsky, E. V. Protopopova, M. E. Stepanova, A. V. Surikov, *Crowdsourcing morphological annotation.* — Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue", vol. 12, RGGU, 2013, pp. 109–115.
9. J. Bollen, M. ten Thij, F. Breithaupt, A. Barron, L. Rutter, L. Lorenzo-Luaces, M. Scheffer, *Historical language records reveal a surge of cognitive distortions in recent decades.* — Proceedings of the National Academy of Sciences **118** (2021), e2102061118.
10. J. Boucher, C. E. Osgood, *The pollyanna hypothesis.* — J. Verbal Learning and Verbal Behavior **8**, No. 1 (1969), 1–8.
11. M. M. Bradley, P. J. Lang, *Affective norms for english words (anew): Instruction manual and affective ratings*, 1999.
12. C. O. Brand, A. Acerbi, A. Mesoudi, *Cultural evolution of emotional expression in 50 years of song lyrics.* — Evolutionary Human Sciences **1** (2019), e11.
13. P. Dodds, E. Clark, S. Desu, M. Frank, A. Reagan, J. Williams, L. Mitchell, K. Harris, I. Kloumann, J. Bagrow, K. Megerdoomian, M. Mcmahon, B. Tivnan, Ch. Danforth, *Human language reveals a universal positivity bias.* — Proceedings of the National Academy of Sciences **112** (2014).
14. P. M. Greenfield, *The changing psychology of culture from 1800 through 2000.* — Psychological Science **24, No. 9** (2013), 1722–1731.
15. T. T. Hills, E. Proto, D. Sgroi, C. I. Seresinhe, *Historical analysis of national subjective wellbeing using millions of digitized books.* — Nature Human Behaviour **3**, No. 12 (2019), 1271–1275.
16. R. Iliev, J. Hoover, M. Dehghani, R. Axelrod, *Linguistic positivity in historical texts reflects dynamic environmental and psychological factors.* — Proceedings of the National Academy of Sciences **113**, No. 49 (2016), E7871–E7879.

17. E. A. Kamalov, E. D. Ponarin, *National pride and subjective well-being of russians.* — Monitoring of Public Opinion: Economic and Social Changes (2020), no. 1, 177–205.
18. I. M. Kloumann, C. M. Danforth, K. D. Harris, C. A. Bliss, P. S. Dodds, *Positivity of the english language.* — PLOS ONE **7**, No. 1 (2012), 1–7.
19. O. Y. Koltsova, S. V. Alexeeva, S. N. Kolcov, *An opinion word lexicon and a training dataset for russian sentiment analysis of social media.* — Komp'yuternaia Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoj Konferentsii "Dialog", 2016, pp. 277–287.
20. E. V. Kotel'nikov, E. V. Razova, A. V. Kotel'nikova, S. V. Vychegzhanin, *Sovremennye slovari ocenochnoj leksiki dlya analiza mnenij na russkom i anglijskom yazykah (analitich-eskij obzor).* — Nauchno-tekhnicheskaya informaciya. Seriya 2: Informacionnye processy i sistemy (2020), no. 12, 16–33.
21. D. I. Kulagin, *Publicly available sentiment dictionary for the russian language kartaslovsent.* — Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue", Suppl. volume, vol. 20, 2021, pp. 1106–1119.
22. M. Lamont, *From 'having' to 'being': self-worth and the current crisis of american society.* — The British Journal of Sociology **70** (2019), 660–707.
23. C. S. Leonard, *The Soviet Economy (1918–1991).* — The Contemporary Russian Economy (Marek Dabrowski, ed.), Springer Books, Springer, 2023, pp. 61–77.
24. Y. Lin, J.-B. Michel, E. A. Lieberman, J. Orwant, W. Brockman, S. Petrov, *Syntactic annotations for the Google Books NGram corpus.* — Proceedings of the ACL 2012 System Demonstrations (Jeju Island, Korea), Association for Computational Linguistics, July 2012, pp. 169–174.
25. N. Loukachevitch, A. Levchik, *Creating a general Russian sentiment lexicon.* — Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16) (Portorož, Slovenia), European Language Resources Association (ELRA), May 2016, pp. 1171–1176.
26. A. Markevich, M. Harrison, *Great war, civil war, and recovery: Russia's national income, 1913 to 1928.* — J. Economic History **71**, No. 3 (2011), 672–703.
27. S. Mohammad, *Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words.* — Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (Melbourne, Australia), Association for Computational Linguistics, 2018, pp. 174–184.
28. O. Morin, A. Acerbi, *Birth of the cool: a two-centuries decline in emotional expression in anglophone fiction.* — Cognition and Emotion **31**, No. 8 (2017), 1663–1675.
29. K. Napier, L. Shamir, *Quantitative sentiment analysis of lyrics in popular music.* — J. Popular Music Studies **30** (2018), 161–176.
30. T. A. Nestik, A. L. Zhuravlev, *Psikhologiia global'nykh riskov*, IP RAS, Moskow, 2018.
31. T. Nguyen, D. Phung, B. Adams, T. Tran, S. Venkatesh, *Classification and pattern discovery of mood in weblogs.* — Advances in Knowledge Discovery and Data Mining (Berlin, Heidelberg) (Mohammed J. Zaki, Jeffrey Xu Yu, B. Ravindran, and Vikram Pudi, eds.), Springer Berlin Heidelberg, 2010, pp. 283–290.

32. S. Oishi, J. Graham, S. Kesebir, I. C. Galinha, *Concepts of happiness across time and cultures.* — Personality and Social Psychology Bulletin **39**, No. 5 (2013), 559–577.
33. J. Oyanedel, D. Paez, *Editorial: Social belongingness and well-being: International perspectives.* — Frontiers in Psychology **12** (2021).
34. L. Peng, S. Luo, *Impact of social economic development on positive and negative affect among chinese college students: a cross-temporal meta-analysis*, 2001–2016, Culture and Brain **10**, No. 2 (2022), 95–112.
35. J. W. Pennebaker, C. K. Chung, M. Ireland, A. Gonzales, R. J. Booth, *The development and psychometric properties of liwc2007.* — Austin, TX, LIWC. Net (2007) (English).
36. T. Piketty, *Capital and ideology*, Harvard University Press, Cambridge, 2020.
37. L. Qiu, J. Lu, C. Chiu, *Detecting the needs for happiness and meaning in life from google books.* — 2014 International Conference on Orange Technologies, 2014, pp. 145–148.
38. D. G. Rand, G. Kraft-Todd, J. Gruber, *The collective benefits of feeling good and letting go: Positive emotion and (dis)inhibition interact to predict cooperative behavior.* — PLOS ONE **10**, No. 1 (2015), 1–12.
39. V. A. Sibirev, N. A. Golovin, *Intergenerational differences in life satisfaction and a feel-ing of happiness in russia (based on the european social survey data).* — Monitoring of Public Opinion: Economic and Social Changes (2020), no. 1, 296–315.
40. S. Smirnov, *Industrial output in the ussr and russia, 1861–2012. part i. reconstruction of basic time-series*, Voprosy Ekonomiki (2013), no. 6, 59–83.
41. V. Solovyev, V. Bochkarev, V. Bayrasheva, *Dynamics of emotions in european languages.* — The 7th international conference on cognitive science (Moscow), Institute of Psychology of RAS, 2016, pp. 71–72.
42. V. Solovyev, M. Islamov, V. Bayrasheva, *Dictionary with the evaluation of positivity/negativity degree of the russian words.* — Speech and Computer: 24th International Conference, SPECOM 2022, Gurugram, India, November 14–16, 2022, Proceedings (Berlin, Heidelberg), Springer-Verlag, 2022, pp. 651–664.
43. V. D. Solovyev, V. V. Bochkarev, A. D. Kaveeva, *Variations of social psychology of russian society in last 100 years.* — 2015 IEEE International Conference on Smart City/SocialCom/SustainCom/DataCom/SC2 2015, Chengdu, China, December 19-21, 2015, IEEE Computer Society, 2015, pp. 519–523.
44. J. M. Twenge, W. K. Campbell, B. Gentile, *Male and female pronoun use in u.s. books reflects women's status*, 1900–2008, Sex Roles **67**, No. 9-10 (2012), 488–493.

45. B. B. Velichkovsky, V. D. Solovyev, V. V. Bochkarev, F. F. Ishkineeva, *Transition to market economy promotes individualistic values: Analysing changes in frequencies of russian words from 1980 to 2008.* — International Journal of Psychology **54**, No. 1 (2019), 23–32.
46. P. F. Wodarz, I. G. Harris, *Historical and contemporary patterns of emotional expression in written texts*, Research Square (2022).

Kazan Federal University,
Kazan 420008, Russia
E-mail: vladimir.bochkarev@kpfu.ru

Поступило 6 сентября 2023 г.

Kazan Federal University,
Kazan 420008, Russia
E-mail: maki.soloviev@mail.ru

Institute of Psychology
of the Russian Academy of Sciences,
13 Yaroslavskaya str.,
Moscow, 129366, Russia
E-mail: nestik@gmail.com

Kazan Federal University,
Kazan 420008, Russia
E-mail: anna_ling@mail.ru