Probing the “Temperature” Approach on Ukrainian Texts: Long-prose Fiction by Ivan Franko

Solomija Buk, Andrij Rovenchak (Lviv)

Abstract. This paper will analyze Ivan Franko’s long prose fiction, using the previously developed approach involving the “temperature” parameter defined from the frequency spectrum of texts. The scaling of this parameter with text size is studied and certain groupings of texts with respect to the scaling exponent and coefficient are observed. A new quantity is proposed to analyze relations between different sub-systems in texts.

Keywords: *“temperature” of text; frequency spectrum; Ivan Franko; long prose fiction; author’s speech; direct speech.*

1. Introduction

At the current state of scientific development there is an increasing interest in intercultural and interdisciplinary fields. Researchers are able to observe common laws working in exact and natural sciences as well as in social sciences and humanities. Techniques from the domain of statistical physics are applied in mathematics (Tran et al. 2004), biology (Jin et al.2009), humanities and social sciences (Bohorquez et al. 2009; Palchykov et al. 2013; Colaiori et al. 2015; Mryglod et al. 2015). Linguistic studies are also well represented here (Eroglu 2014; Rodriguez et al. 2014), including approaches from the point of view of complex networks (Ferrer-i-Cancho et al. 2004; Čech et al. 2011; Mac Carrona, Kenna 2013).

The “temperature” concept with respect to languages was proposed in several ways (Mandelbrot 1953; Kosmidis et al. 2006), in particular, by analyzing the high-frequency words using the Boltzmann distribution (Miyazima, Yamamoto 2008; Rêgo et al. 2014).We introduced the “temperature” parameter in a different way, considering the low-frequency vocabulary (Rovenchak, Buk 2011a).

The idea of the present study is to extend the “temperature” approach to texts in Ukrainian written by Ivan Franko. The paper is organized as follows. Section 2 contains a brief description of Ivan Franko’s works of fiction. In Section 3, the approach applied for the quantitative text analysis is explained. Results and discussion are given in Section 4 followed by conclusions in Section 5.

2. Ivan Franko's prose

Ivan Franko (1856–1916) was a famous Ukrainian writer, scholar, and public figure (Stech, Zhukovsky 2007). His long prose works are well described from the quantitative and statistical point of view (Buk, Rovenchak 2007; Buk 2010; 2011; 2012; 2013; Kelih et al. 2014). Note that there are also quantitative studies of Franko's letters and poetry (Best, Zinenko 1998; 1999) and fables (Holovatch, Palchykov 2017). We have an ongoing project of compiling a tagged corpus of Ivan Franko's prose (Buk 2007); this corpus could be a good base for further quantitative studies.

Ivan Franko authored over one hundred works of fiction in prose of different sizes, from short stories of some two pages to rather lengthy novels. There is no strict quantitative definition separating the types of narratives (namely, short stories, novellas, and novels) but both theorist and practitioners of literature generally agree on the upper limit of 20,000 words for a short story (cf. Thrall, Hibbard 1960: 457–458; Kotter 2008: 248–249; King 2010; Waas 2012: ix). However, this number is based on short stories written in English, and it should be somewhat lowered when dealing with a more synthetic language such as Ukrainian.

Within Franko's œuvre, ten works are usually referred to as long prose fiction (cf. Pastukh 1996; Denysiuk 2008):

1. *Boa constrictor* (1st edition: 1878–84; 2nd edition: 1905–07);
2. *Boryslav smijetsja* (Boryslav Laughs) (1880–81);
3. *Zakhar Berkut* (1883);
4. *Ne spytavšy brodu* (Without Asking a Wade) (1885–86);
5. *Dlja domašnjoho ohnyšča* (For the Hearth) (1892);
6. *Osnovy suspil'nosty* (Pillars of Society) (1894–95);
7. *Perekhresni stežky* (The Cross-paths) (1900);
8. *Velykyj šum* (The Great Noise) (1907);
9. *Petriji j Dovbuščuky* (2nd edition: 1909–12).
10. *Lelum i Polelum* [in Polish].

Figure 1. Title pages of some Ivan Franko's long prose works.

We will henceforth refer to the titles using the first letters of the Ukrainian transliteration, with this possibly followed by the edition number, i.e.: BC1, BC2, BSm, ZBe, NSB, DDO, OSu, PSt, VSh, and PD2. See Figure 1 for some title pages.

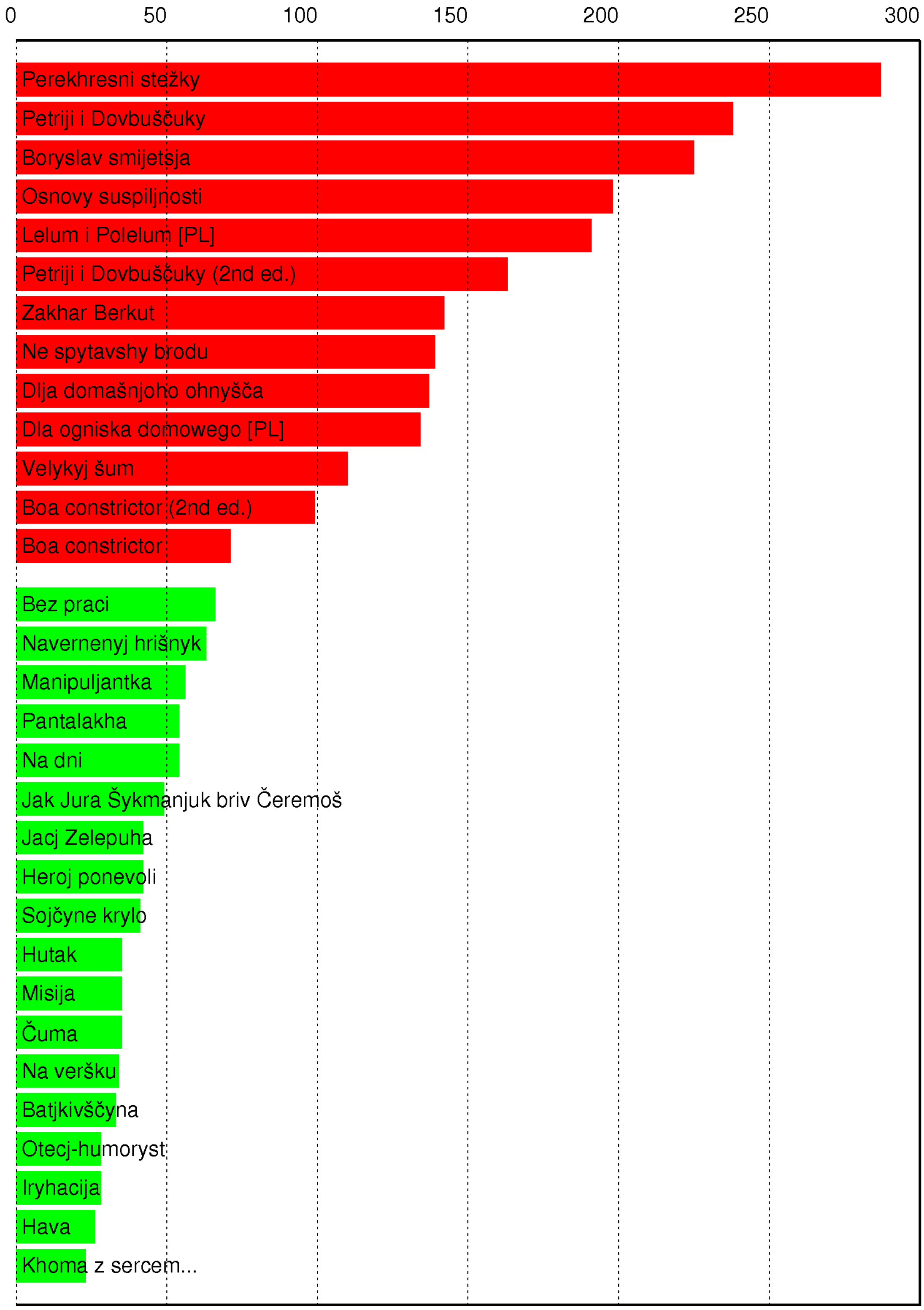

Figure 2. Ivan Franko's prose fiction, by size (in pages).
The mark [PL] denotes texts written in Polish

Interestingly enough, such a division of Franko's prose is confirmed by quantitative data. In Figure 2 the size of each piece of prose is shown, measured in number of pages (Franko 1976–86). On average, a printed page corresponds to

340 words. So, the first edition of *Boa constrictor* (71 pages, ca. 25 thousand words) can be attributed as a novel whereas its nearest lower neighbor *Bez praci* (Without work, 66 pages, ca. 22 thousand words) can be attributed as a short story with some reservations if considering solely quantitative data. A qualitatively significant jump is observed only for the second edition of *Boa constrictor* (99 pages, about 34 thousand words).

3. The “temperature” approach

The approach used for the analysis in the present work was previously developed by the authors (Rovenchak, Buk 2011a; 2011b) and then applied with several modifications to obtain some hints on language evolution (Rovenchak 2014) as well as to some contrastive studies (Rovenchak 2015a; 2015b).

The idea of this approach is based on the analogy between the frequency spectrum of texts and the so-called Bose-distribution in statistical physics (Huang 1987: 183). The frequency spectrum N_j is the number of tokens having absolute frequency exactly equal to j (Tuldava 1996; Popescu et al. 2009). The value of N_1 thus corresponds to the number of *hapax legomena*. The set of N_j is obtained for each text and subsequently fitted to the model of the Bose-distribution:

$$N_j = \frac{1}{z^{-1}e^{(j-1)^\alpha/T}-1}\,. \qquad (1)$$

Equation (1) contains three parameters: z, α, and T. The first parameter is fixed by the number of hapaxes,

$$z = \frac{N_1}{N_1+1}\,. \qquad (2)$$

The remaining parameters, α and T, are calculated by fitting the observed frequency spectrum to model (1) using the least-squares method. The relation $\tau = \ln T / \ln N_{\text{tot}}$, where N_{tot} is the total number of words (tokens) in the given text, proved to be a parameter suitable for text classification alongside the exponent α (Rovenchak, Buk 2011a; 2011b; Rovenchak 2014).

Table 1
Fitting parameters of model (1) for Ivan Franko's long prose fiction

Title	**Abbr.**	N_{tot}	T	α	$\ln T / \ln N_{tot}$
Boa Constrictor (1st ed.)	BC1	25427	1554	1.66	0.724
Boa Constrictor (2nd ed.)	BC2	34215	1899	1.61	0.723
Boryslav smijetsja	BSm	77456	3192	1.54	0.717
Dlja domašnjoho ohnyšča	DDO	44841	2241	1.59	0.720
Ne spytavšy brodu	NSB	49170	2422	1.61	0.721
Osnovy suspiljnosty	OSu	67174	3038	1.58	0.721
Perekhresni stežky	PSt	93888	4099	1.56	0.727
Petriji i Dovbuščuky (2nded.)	PD2	52751	2730	1.58	0.728
Velykyj šum	VSh	37005	1923	1.66	0.719
Zakhar Berkut	ZBe	50206	2521	1.54	0.724
			Average:	**1.59**	**0.722**

We further study the behavior of the "temperature" parameter T in the course of text production, i.e. how T changes for the first 1,000 of words in each text, for the first 2,000 of words, and so on. For simplicity, we fix the parameter α at 1.6. It appears that the dependence of T on N is, to a good accuracy, described by a simple power law:

$$T = tN^{\beta}. \qquad (3)$$

Note that such dependences occur in linguistics in various contexts (Naranan, Balasubrahmanyan 1998; Köhler 2002; Kaniadakis 2009).

Figure 3 demonstrates the results of fitting for full texts (with no division on the author's and direct speech). Detailed results of calculations are given in the next section. The analysis was completed for texts as a whole as well as for the author's and direct speech separately. These types of speech differ in the values of some text parameters, e.g., number of hapax legomena, dictionary and text richness, etc. (Buk 2011).

We can observe minor oscillations around the fitting curve of model (3). The nature of such behavior could itself become the subject of more detailed studies in the future (cf. Zörnig et al. 1990).

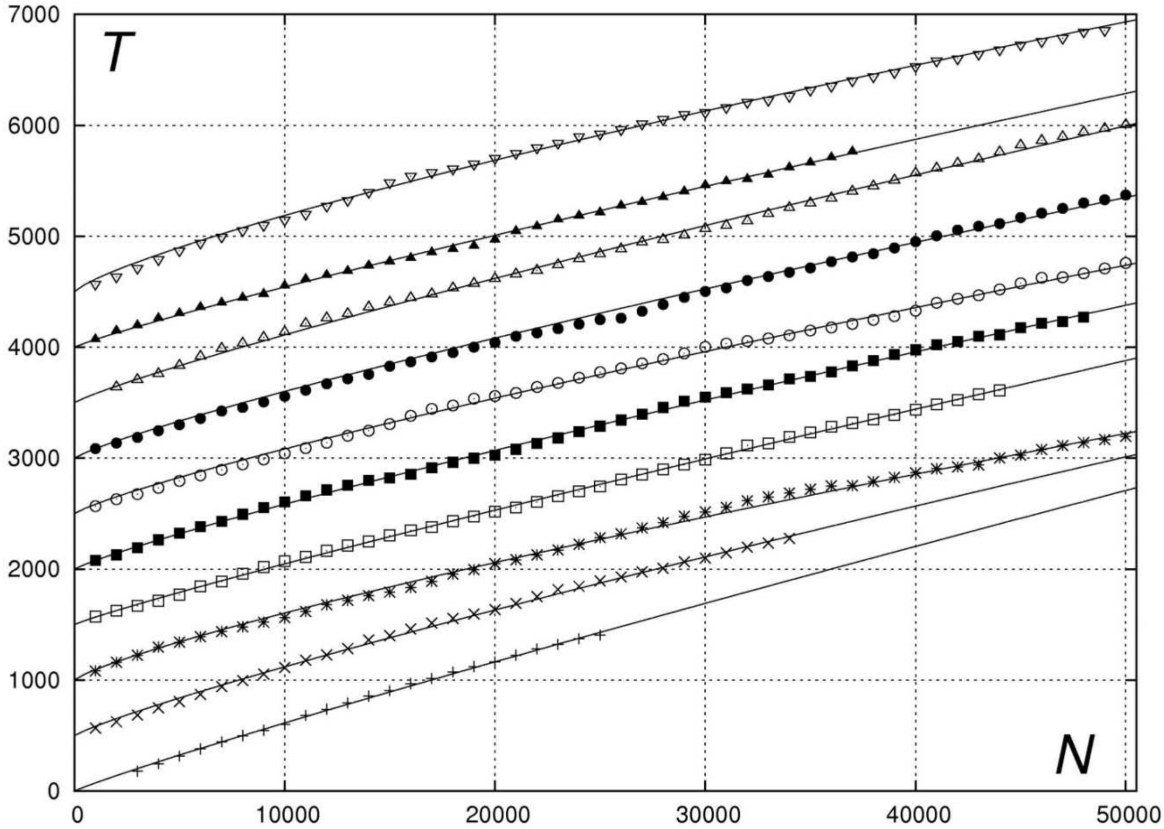

Figure 3. "Temperature" T as a function on text length N in Ivan Franko's prose fiction. For convenience, the values are shifted vertically by 500 for ten novels listed in Table 1, bottom to top: BC1 (+), BC2 (×), BSm (✳), DDO (□), NSB (■), OSu (○), PSt (●), PD2 (△), VSh (▲), and ZBe (▽).

4. Results and their interpretation

The following figures demonstrate the values of the t and β parameters of Ivan Franko's long prose fiction for full texts, author's speech, and direct speech of all the ten Ukrainian texts listed in Table 1. In Figure 4, the same scale is used for all the three panels while different scales (enlarged regions) are shown in Figure 5 for better visualization. Numerical values of the parameters and respective errors are given in Table 2.

In the direct speech data one can observe the highest concentration for PSt, PD2, and BC2; ZBe and VSh are located close together, as are BSm and OSu.

As expected, rather compact data are observedin author's speech. For full texts, the highest concentration is found for NSB, PD2, and BC2. A separate location of PSt and OSu is also clearly seen. A somewhat dispersed group containing DDO, BC1, and VSh can be distinguished as well.

It would be interesting to find some common features of texts reflecting the abovementioned grouping. Below, we will analyze the correlation of the calculated parameters with some other numerical text data.

To extend the previously applied apparatus, one can also define the quantity

$$\mu = T \ln z \qquad (4)$$

known in physics as the chemical potential. For sub-systems in equilibrium, the chemical potentials are equal. We can estimate to what extent such a claim is applicable to our model by considering the direct and author's speech as separate sub-systems within each text. To make the comparison, the following relative errors were calculated in each case:

$$\delta\mu = 2\,|\mu_a - \mu_d| \,/\, (\mu_a + \mu_d). \qquad (5)$$

The subscripts "a" and "d" correspond to the author's and direct speech, respectively.

The results are presented in Table 2. While there is quite strong inverse correlation between $\delta\mu$ and text size N (with the coefficient $R = -0.61$, see Table 3), we can still observe that the direct and author's speech is best balanced in *Perekhresni stežky* and worst in *Petriji i Dovbuščuky* (2nd ed.).

Curiously enough, such an estimation contradicts opinions of literary reviewers in the case of *Perekhresni stežky* (cf. Batsevych et al. 2007: 7-8).

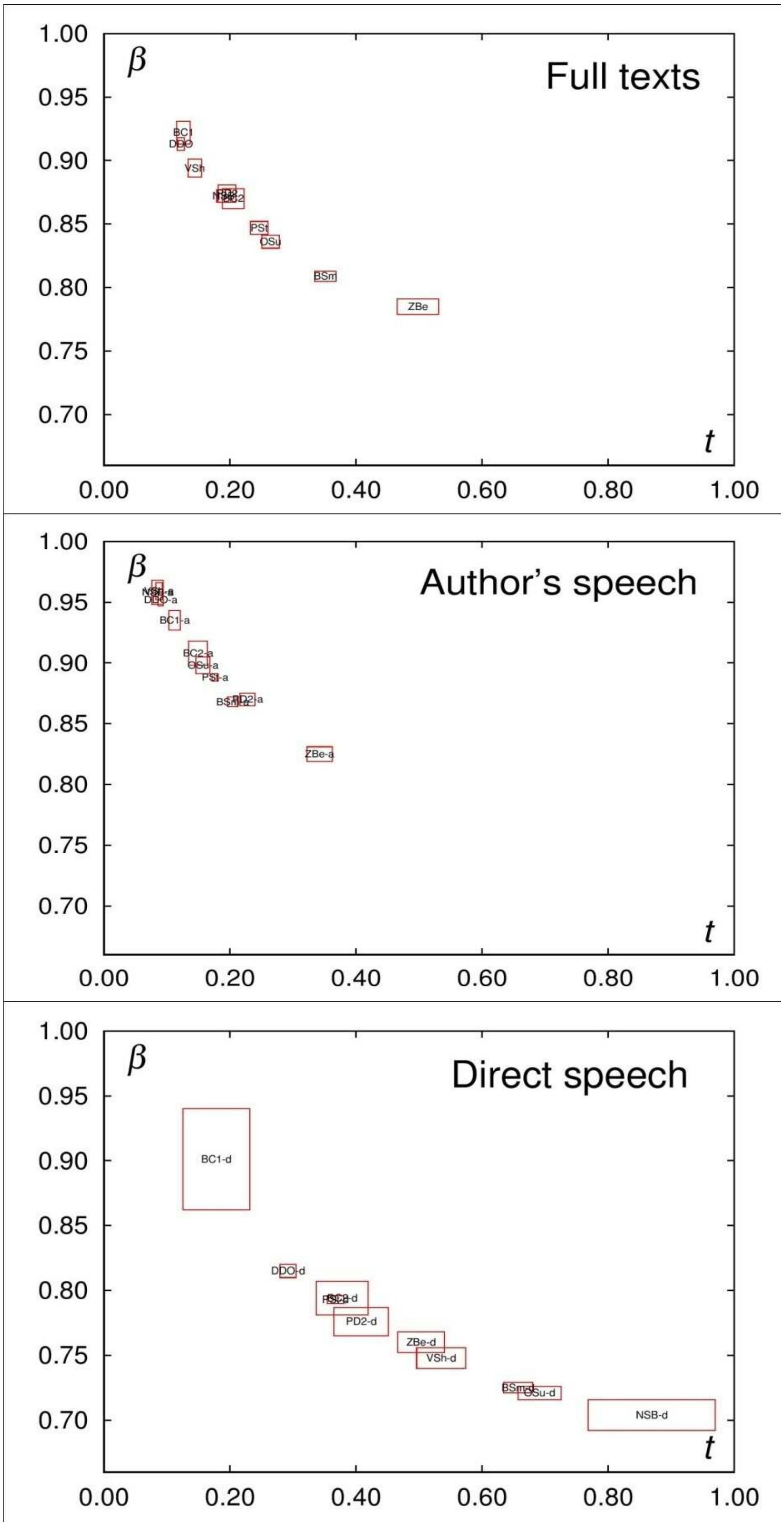

Figure 4. Parameters of Ivan Franko's long prose fiction for full texts, author's speech, and direct speech

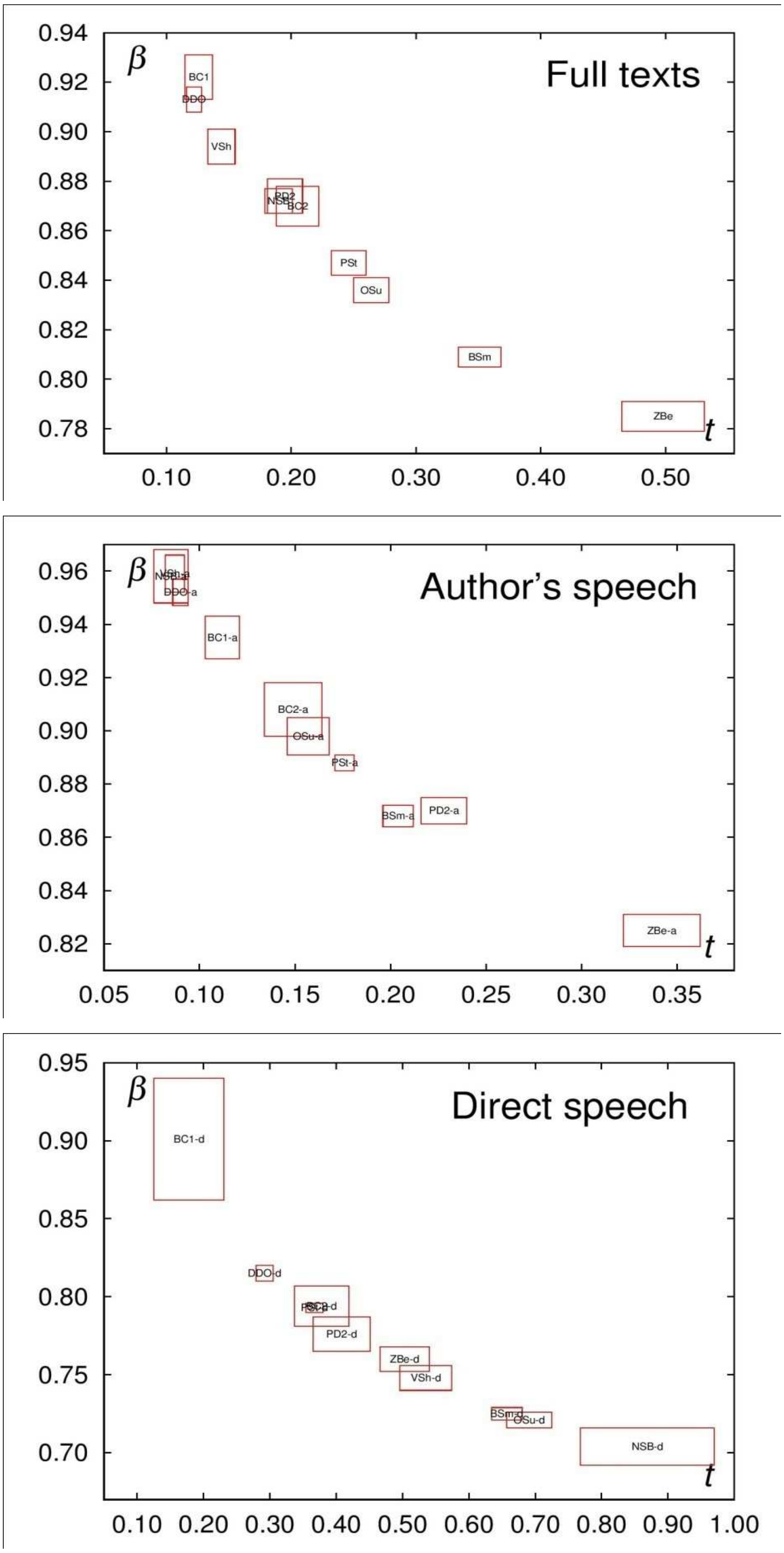

Figure 5. Parameters of Ivan Franko's long prose fiction for full texts, author's speech, and direct speech, enlarged regions compared to Figure 4.

Table 2
The values of parameters for Ivan Franko's long prose.
The extensions "-a" and "-d" correspond to the author's and direct speech, respectively. The "size" column lists the total number of tokens in thousands (mathematical floor is applied).

Text	t	Δt	β	$\Delta\beta$	T	N_1	μ	$\delta\mu$	size
BC1	0.126	0.011	0.922	0.009					
BC1-a	0.112	0.009	0.935	0.008	1270	5254	–0.242	13.4%	25
BC1-d	0.178	0.053	0.901	0.039	245	886	–0.276		
BC2	0.205	0.017	0.870	0.008					
BC2-a	0.149	0.015	0.908	0.010	1527	6522	–0.234	8.0%	34
BC2-d	0.378	0.041	0.794	0.013	443	1747	–0.254		
BSm	0.351	0.017	0.809	0.004					
BSm-a	0.204	0.008	0.868	0.004	2179	7930	–0.275	5.0%	76
BSm-d	0.657	0.023	0.725	0.004	1223	4231	–0.289		
DDO	0.122	0.006	0.913	0.005					
DDO-a	0.090	0.004	0.952	0.005	1393	5785	–0.241	8.7%	44
DDO-d	0.292	0.013	0.815	0.005	885	3368	–0.263		
NSB	0.190	0.011	0.872	0.005					
NSB-a	0.085	0.009	0.958	0.010	1479	6293	–0.235	4.9%	48
NSB-d	0.869	0.101	0.704	0.012	922	3738	–0.247		
OSu	0.264	0.014	0.836	0.005					
OSu-a	0.157	0.011	0.898	0.007	2051	7499	–0.274	8.8%	67
OSu-d	0.691	0.034	0.721	0.005	1121	4475	–0.251		
PSt	0.246	0.014	0.847	0.005					
PSt-a	0.176	0.005	0.888	0.003	2481	8640	–0.287	0.8%	93
PSt-d	0.367	0.013	0.793	0.003	1858	6524	–0.285		
PD2	0.195	0.014	0.874	0.007					
PD2-a	0.228	0.012	0.870	0.005	2035	6991	–0.291	18.3%	52
PD2-d	0.408	0.043	0.776	0.011	761	3142	–0.242		
VSh	0.144	0.011	0.894	0.007					
VSh-a	0.087	0.005	0.959	0.007	1176	5784	–0.203	13.2%	37
VSh-d	0.535	0.039	0.748	0.008	773	3332	–0.232		
ZBe	0.498	0.033	0.785	0.006					
ZBe-a	0.342	0.020	0.825	0.006	1769	6541	–0.270	6.6%	49
ZBe-d	0.503	0.037	0.760	0.008	829	2870	–0.289		

As there were no immediately apparent text properties reflecting the behavior of t, β, and μ, we analyzed their correlation with text size (N_{tot}), year of publication, epithetization index (EI), as well as the so-called vocabulary richness (TTR). The epithetization index is the number of adjectives divided by the number of nouns in a given text (Ruszkowski 2004; Buk 2012). The vocabulary richness is equal to the ratio of the number of different lemmas and the total number of words in a given text (type-token ratio or TTR).

Table 3
Correlation coefficients between various text parameters

	t	β	N_{tot}	μ	year	EI	TTR
t	—	–0.96	+0.42	–0.41	–0.42	–0.07	–0.57
β	–0.96	—	–0.59	+0.49	+0.31	–0.13	+0.70
N_{tot}	+0.42	–0.59	—	–0.61	–0.08	+0.63	–0.92

The highest correlation of β with TTR is in agreement with previously discovered relation between the “temperature”-related parameters and the level of language analyticity (Rovenchak, Buk 2011a; 2011b; Rovenchak 2014). Another interesting result is the absence of correlation between t, β and the epithetization index. The latter could signify, in particular, some complications in substantiating the argument that the groupings observed in Figures 4 and 5 were literary-based.

In view of a high correlation between t and β, it is interesting to find the fitting function as a simple power law:

$$\beta = bt^{-a} = 0.722t^{-0.114}. \qquad (6)$$

We thus obtain a very simple link between $\tau = \ln T / \ln N_{tot}$ and β:

$$\tau = \beta + \frac{1}{a}\frac{\ln b - \ln \beta}{\ln N_{tot}}, \qquad (7)$$

where $b = 0.722\ldots$ and $a = 0.114\ldots$. Obviously, the second item vanishes as the text size becomes large (cf. Rovenchak 2015b).

5. Prospects and discussion

In this paper we analyzed Ivan Franko’s long prose fiction using a previously developed method inspired by a physical model. We were able to reconfirm previously revealed relations of the “temperature” parameter with the level of language analyticity. This time it is made indirectly, from the correlation between the “temperature” scaling exponent β and the vocabulary richness (type-token ratio with lemmas considered as types).

Moreover, some new results were also obtained. The observed small oscillations of the $T(N)$ dependence around the simple power-law fitting function might signify some supra-sentence peculiarities in the frequency structure of text (cf. Altmann 2014; Chen, Altmann 2015). In future studies it would be interesting to apply the analysis based on the introduced μ parameter for different sub-systems in texts, e.g. part-of-speech distribution. We also hope to obtain similar data for other languages and authors, to verify the approach developed in this study. It seems tempting to ascribe some unmeasurable properties to the analyzed texts on the basis of the calculated parameters. Whether it is possible – at least to a certain extent – remains so far an open issue. We expect that additional studies involving experts in the domain of literary criticism will clarify this problem.

Acknowledgements

A. R. acknowledges partial support by Project FF-30F (No. 0116U001539) from the Ministry of Educationand Science of Ukraine and by Grant F64/37-2016 (No. 0116U005055) from the State Fund For Fundamental Research of Ukraine.

References

Altmann, G. (2014). Supra-sentence levels. *Glottotheory 5: 25–39*.

Batsevych, F.S.; Buk, S. N.; Protsak, L.M.; Rovenchak; A.A.; Svarychevska, L. Yu.; Tsikhotskyj, I. L. (eds.) (2007). *Stežkamy Frankovoho tekstu (komunikatyvni, stylistyčni ta leksyčni vymiry romanu "Perekhresni stežky")*. Lviv: Lviv University Press.

Best, K.-H.; Zinenko S. (1998). Wortkomplexität im Ukrainischen und ihre linguistische Bedeutung. *Zeitschrift für Slavische Philologie 58: 107–123*.

Best, K.-H.; Zinenko S. (1999). Wortlängen in Gedichten des ukrainischen Autors Ivan Franko. In: Genzor, J.; Ondrejovič S. (eds.) *Pange lingua. Zborník na počest' Viktora Krupu: 201–213*. Bratislava: Veda.

Bohorquez, J.C., Gourley, S., Dixon, A.R., Spagat, M., Johnson, N.F. (2009). Common ecology quantifies human insurgency. *Nature 462: 911–914.*

Buk, S. (2007). Korpus tekstiv Ivana Franka: sproba vyznačennja osnovnykh parametriv. In: Shyrokov, V.A. (ed.), *Prykladna linhvistyka ta linhvistyčni tekhnolohiji: MegaLing 2006: 72–82*. Kyiv: Dovira.

Buk, S. (2010). Statystyčna structura romanu Ivana Franka "Boryslav smijetsja". *Učenye zapiski Tavričeskogo natsionalnogo universiteta im. V.I. Vernadskogo. Serija Filologija 23(62), № 3, 114–118*.

Buk, S. (2011). Prjama j avtorska mova velykoji prozy Ivana Franka: linhvostatystyčne doslidžennja u konteksti korpusnoji linhvistyky. *Visnyk Lvivskoho universytetu. Serija filolohična 52, 199–209*.

Buk, S. (2012). The epithetization index in a work of fiction (on the basis of the text corpus of Ivan Franko's long prose fiction). In: Obrębska, A. (ed.)

Practical Applications of Linguistic Research: 73–85. Łódź: Primum Verbum.

Buk, S. (2013). Kvantytatyvna parametryzacija tekstiv Ivana Franka: proekt ta joho realizacija. *Visnyk Lvivskoho universytetu. Serija filolohična 58, 290–307*.

Buk, S., Rovenchak, A. (2007). Statistical parameters of Ivan Franko's novel *Perekhresni stežky (The Cross-Paths)***.** In: Grzybek, P., Köhler, R. (eds.) *Quantitative Linguistics 62: Exact methods in the study of language and text: dedicated to Professor Gabriel Altmann on the occasion of his 75th birthday*: *39-48*. Berlin–New York: Mouton de Gruyter.

Chen, R., Altmann, G. (2015). Conceptual inertia in texts. *Glottometrics 30, 73-85*.

Colaiori, F., Castellano, C., Cuskley, Ch. F., Loreto, V., Pugliese, M., Tria, F. (2015). General three-state model with biased population replacement: Analytical solution and application to language dynamics. *Physical Review E 91: 012808 [12 pages]*.

Čech, R., Mačutek, J., Žabokrtský, Z. (2011).The role of syntax in complex networks: Local and global importance of verbs in a syntactic dependency network. *Physica A 390: 3614–3623*.

Denysiuk, I. (2008). Novatorstvo Franka-prozajika. *Ukrajinske literatureznav-stvo 70: 138–152*.

Eroglu, S. (2014). Menzerath–Altmann law: Statistical mechanical interpretation as applied to a linguistic organization. *Journal of Statistical Physics 157: 392–405.*

Ferrer-i-Cancho, R., Solé, R.V., Köhler, R. (2004). Patterns in syntactic dependency networks. *Physical Review A 69: 051915* .

Franko, I. (1976–86). *Zibrannja tvoriv u 50-ty tomakh*. Kyiv: Naukova Dumka.

Holovatch, Yu., Palchykov, V. (2017). Complex networks of words in fables. In: Kenna, R.; Mac Carron, M.; Mac Carron, P. (eds.) *Maths Meets Myths: Quantitative Approaches to Ancient Narratives*: *159–175.* Springer.

Huang, K. (1987). *Statistical Mechanics.* 2nd edition. New York: Wiley.

Jin, Neng-zhi, Liu, Zi-xian, Qiu, Wen-yuan (2009). Frequency and correlation of nearest neighboring nucleotides in human genome. *Chinese Journal of Chemical Physics 22, 27–33.*

Kaniadakis, G. (2009). Maximum entropy principle and power-law tailed distributions. *European Physical Journal B 70, 3–13*.

Kelih, E., Rovenchak, A., Buk S. (2014). Analysing h-point in lemmatised and non-lemmatised texts. In: Altmann, G., Čech, R., Mačutek, J., Uhlířová, L. (eds.) *Studies in Quantitative Linguistics 17: Empirical Approaches to Text and Language Analysis; dedicated to Luděk Hřebiček on the occasion of his 80th birthday: 81–93*. Lüdenscheid: RAM-Verlag.

King, S. (2010). Afterword. In: *Different Seasons*. London: Hachette Livre UK Co.

Köhler, R. (2002). Power law models in linguistics: Hungarian. *Glottometrics 5, 51–61*.

Kosmidis, K., Kalampokis, A., Argyrakis, P. (2006). Statistical mechanical approach of human language. *Physica A 366, 495–502*.

Kotter, M.P. (2008). *English Literature: Modern World View*. New Delhi: Cyber Tech Publications.

Mac Carrona, P., Kenna, R. (2013). Network analysis of the Íslendinga sögur – the Sagas of Icelanders. *European Physical Journal B 86, 407*.

Mandelbrot, B. (1953). An informational theory of the statistical structure of languages. In: Jackson, W. (ed.) *Communication Theory 486–502*. New York: Academic.

Miyazima, S., Yamamoto, K. (2008). Measuring the temperature of texts. *Fractals 16: 25–32*.

Mryglod, O., Fuchs, B., Szell, M., Holovatch, Yu., Thurner, S. (2015). Interevent time distributions of human multi-level activity in a virtual world. *Physica A 419, 681–690.*

Naranan, S., Balasubrahmanyan, V. K. (1998). Models for power law relations in linguistics and information science. *Journal of Quantitative Linguistics 5, 35–61.*

Palchykov, V., Kaski, K., Kertész, J., Barabási, A.-L., Dunbar, R.I.M. (2012). Sex differences in intimate relationships. *Scientific Reports 2, 370.*

Pastukh, T. (1996). Roman u systemi prozovykh tvoriv Ivana Franka. *Ukrajins'ke literaturoznavstvo* 6*2, 100–108.*

Popescu, I.-I., Altmann, G., Grzybek, P., Jayaram, B.D., Köhler, R., Krupa, V., Mačutek, J., Pustet, R., Uhlířová, L., Vidya, M.N. (2009). *Word Frequency Studies*. Berlin-New York: Mouton de Gruyter.

Rêgo, H.H.A., Braunstein, L.A., D'Agostino, G., Stanley, H.E., Miyazima, S. (2014). When a text is translated does the complexity of its vocabulary change? Translations and target readerships. *PLoS ONE 9(10): e110213.*

Rodriguez, E., Aguilar-Cornejo, M., Femat, R., Alvarez-Ramirez, J. (2014). Scale and time dependence of serial correlations in word-length time series of written texts. *Physica A 414: 378–386*.

Rovenchak, A. (2014). Trends in language evolution found from the frequency structure of texts mapped against the Bose-distribution. *Journal of Quantitative Linguistics 21(3), 281–294*.

Rovenchak, A. (2015a). Models of frequency spectrum in texts based on quantum distributions in fractional space dimensions. In: Dumitrache, I., Florea, A. M., Pop, F., Dumitraşcu, A. (eds.) *20th International Conference on Control Systems and Computer Science CSCS 2015: Proceedings, 27-29 May 2015, Bucharest, Romania, Vol. 2: 645–649*. Los Alamitos, CA: IEEE Computer Society.

Rovenchak, A. (2015b). Where Alice meets Little Prince: Another approach to study language relationships. In: Mikros, G. K., Mačutek, J. (eds.), *Sequences in Language and Text: 217–230*. Berlin–Boston: Mouton de Gruyter.

Rovenchak, A., Buk, S. (2011a). Application of a quantum ensemble model to linguistic analysis. *PhysicaA 390(7), 1326–1331*.

Rovenchak, A., Buk, S. (2011b). Defining thermodynamic parameters for texts from word rank-frequency distributions. *Journal of Physical Studies 15(1), 1005 [6 pages]*.

Ruszkowski, M. (2004). Wskaźnik epitetyzacji w badaniach stylistycznych. *Respectus Philologicus 5(10), 48–53.*

Stech, M. R., Zhukovsky, A. (2007). Franko, Ivan. In: Kubijovyč, V. (ed.) *Encyclopedia of Ukraine*, http://www.encyclopediaofukraine.com/pages/F/R/FrankoIvan.htm

Thrall, W.F., Hibbard, C.A. (1960). *A Handbook to Literature* (2nd ed, revised and enlarged by C. Hugh Holman). New York: Odyssey Press.

Tran, M.N., Murthy, M.V.N., Bhaduri, R.J. (2004). On the quantum density of states and partitioning an integer. *Annals of Physics 311, 204–219*.

Tuldava, J. (1996). The frequency spectrum of text and vocabulary. *Journal of Quantitative Linguistics 3(1), 38–50*.

Waas, G. (2012). Introduction. In: *The Great American Short Story*. Bloomington, IN: AuthorHouse.

Zörnig, P., Köhler, R., Brinkmöller, R. (1990). Differential equation models for the oscillation of the word length as a function of frequency. *Glottometrika 12, 25–40.*

Issues in Quantitative Linguistics

4

edited by

Emmerich Kelih
Róisín Knight
Ján Mačutek
Andrew Wilson

Dedicated to Reinhard Köhler on the occasion of his 65th birthday

2016

RAM-Verlag

Studies in Quantitative Linguistics

1. U. Strauss, F. Fan, G. Altmann, *Problems in quantitative linguistics 1*. 2008, VIII + 134 pp.
2. V. Altmann, G. Altmann, *Anleitung zu quantitativen Textanalysen. Methoden und Anwendungen.* 2008, IV+193 pp.
3. I.-I. Popescu, J. Mačutek, G. Altmann, *Aspects of word frequencies*. 2009, IV +198 pp.
4. R. Köhler, G. Altmann, *Problems in quantitative linguistics 2*. 2009, VII + 142 pp.
5. R. Köhler (ed.), *Issues in Quantitative Linguistics*. 2009, VI + 205 pp.
6. A. Tuzzi, I.-I. Popescu, G.Altmann, *Quantitative aspects of Italian texts*. 2010, IV+161 pp.
7. F. Fan, Y. Deng, *Quantitative linguistic computing with Perl*. 2010, VIII + 205 pp.
8. I.-I. Popescu et al., *Vectors and codes of text*. 2010, III + 162 pp.
9. F. Fan, *Data processing and management for quantitative linguistics with Foxpro*. 2010, V + 233 pp.
10. I.-I. Popescu, R. Čech, G. Altmann, *The lambda-structure of texts*. 2011, II + 181 pp
11. E. Kelih et al. (eds.), *Issues in Quantitative Linguistics Vol. 2*. 2011, IV + 188 pp.
12. R. Čech, G. Altmann, *Problems in quantitative linguistics 3*. 2011, VI + 168 pp.
13. R. Köhler, G. Altmann (eds.), *Issues in Quantitative Linguistics Vol 3*. 2013, IV + 403 pp.
14. R. Köhler, G. Altmann, *Problems in Quantitative Linguistics Vol. 4*. 2014, VIII+148 pp.
15. Best, K.-H., Kelih, E. (eds.), *Entlehnungen und Fremdwörter: Quantitative Aspekte*. 2014. VI + 163 pp.
16. I.-I. Popescu, K.-H. Best, G. Altmann, G. *Unified Modeling of Length in Language*. 2014, VIII + 123 pp.

17. G. Altmann, R. Čech, J. Mačutek, L. Uhlířová (eds.), *Empirical Approaches to Text and Language Analysis dedicated to Luděk Hřebíček on the occasion of his 80th birthday*. 2014. VI + 231 pp.
18. M. Kubát, V. Matlach, R. Čech, *QUITA Quantitative Index Text Analyzer*. 2014, VII + 106 pp.
19. K.-H. Best, *Studien zur Geschichte der Quantitativen Linguistik.* 2015. III + 158 pp.
20. P. Zörnig, K. Stachowski, I.-I. Popescu, T. Mosavi Miangah, P. Mohanty, E. Kelih, R. Chen, G. Altmann, *Descriptiveness, Activity and Nominality in Formalized Text Sequences.* 2015. IV+120 pp.
21. G. Altmann, *Problems in Quantitative Linguistics Vol. 5*. 2015. III+146 pp.
22. P. Zörnig, K. Stachowski, I.-I. Popescu, T. Mosavi Miangah, R. Chen, G. Altmann, *Positional occurrences in texts: Weighted Consensus Strings.* 2015. II+178 pp.

ISBN: 978-3-942303-44-6

RAM-Verlag
Stüttinghauser Ringstr. 44
D-58515 Lüdenscheid
Germany
RAM-Verlag@t-online.de
http://ram-verlag.eu

Contents