

A New Algorithm for Extracting Formulas from Poetic Texts and the Formulaic Density of Russian *Bylinas*

Dmitry Nikolayev

Ever since the promulgation of Milman Parry's formulaic theory of epic poetry in the works of Francis Magoun (1953) and Albert Lord (1960), the problem of estimating the formulaic density of a text or a group of texts has been frequently addressed in studies of oral and ancient poetry. Different estimates have been proposed for the formulaic densities of Homeric epics and hymns (Edwards 1988), Middle English alliterative poetry (Ward 1986), pre-Islamic Arabic poetry (Monroe 1972), and Russian *bylinas* (Uxov 1975; Arant 1990), to name but a few lines of enquiry. The early work in this field was characterized by widely divergent results (cf. 90% ratio of formulas in the *Iliad* proposed by Lord as compared to 57.5% of formulas in Homeric texts in Minton 1975), largely due to the methodology used by the scholars. Since they were not able to calculate the total number of formulas in a text of any considerable length, let alone a corpus of texts, the scholars resorted to sampling procedures. In most cases this meant choosing a representative body of verse lines and then calculating the ratio of lines in it found elsewhere in the same corpus. The advent of computer technologies providing easily searchable full-text corpora could have led to a new consensus in this field. However, the statistical study of formulas in poetic text fell into decline at the end of the 1970s, and the early computerized efforts in this field (Vikis-Freibergs and Freibergs 1978a; Strasser 1984) were not kept up.

In fact, as far as I am aware, Vikis-Freibergs and Freibergs' 1978 paper on the formulaic density of Latvian folk songs about the sun is the only published attempt at formulating a strict procedure for counting formulaic density of a poetic text amenable to implementation as a computer program. The scholars maintain that their method, "while developed on one particular corpus . . . can be readily adapted to other computer-accessible [corpora]" (Vikis-Freibergs & Freibergs 1978a:330). This is indeed true, but, as I will show in the following, their algorithm is not entirely satisfactory as far as the extraction of formulas itself is concerned. I instead propose a new algorithm, one which is both more flexible and more powerful, and report the results of its application to a corpus of Russian *bylina* epics.

Vikis-Freibergs and Freibergs' Algorithm

The algorithm used by Vikis-Freibergs and Freibergs to estimate the formulaic density of

Latvian sun songs is rather straightforward. All the texts from the Montreal corpus of Latvian Sun-songs are put into a computer and stored in the memory (Vikis-Freibergs and Freibergs 1978b). Frequency counts are then taken of overlapping word pairs and word triplets from each line (that is to say, $n-1$ word pairs and $n-2$ word triplets are taken from a line of n words) and of larger non-overlapping units: single lines and line couplets. The extracted units are then arranged in the form of two printouts: the first presents the units in order of decreasing frequencies of occurrence, and the second gives an alphabetical listing.

The data obtained in this way can be used to estimate the formulaic density of a text understood as the proportion of units occurring at least n times in the corpus. Applying the program to the Latvian sun-songs and using 2 as the lower threshold, Vikis-Freibergs and Freibergs obtained the following results: of all the word pairs in the corpus 63.1% are formulaic, and the same holds for 46.7% of word triplets, 43.5% of lines (the last two evidently constitute a single formulaic level), and 20.8% of couplets.

Unfortunately the results obtained in this way, though in clear numeric form, are hard to interpret. Firstly, formulaic word pairs and triplets can overlap since, by design, overlapping word pairs and triplets are all counted independently. In order to avoid over-counting one must carefully scan all the resulting formulaic units, which moreover must be indexed as to their position in the corpus. The data of the last kind are not provided by the scholars' algorithm, so when dealing with a sufficiently large corpus over-counting becomes unavoidable, which is a major drawback.

Secondly, the frequency counts of units of different level are hard to combine. It is not possible to simply sum up the ratios of formulaic word pairs, word triplets, and lines. Therefore, the formulaic density obtained in this way is manifold: formulaic densities for word pairs, word triplets, and lines are computed separately. To some extent this can be regarded as an advantage because the results are more nuanced. By looking at formulaic densities on different levels (word pairs, triplets, larger n -grams, lines) we can discern the basic building blocks used by singers to improvise the text. For instance, if we see a high formulaic density on the level of word pairs but a considerably lower one on the level of word triplets, we can conclude that it is word pairs that are memorized and recombined in the process of oral composition. A unified formulaic density measure does not provide such information. It is still desirable, however, to have a unified measure of formulaic density that can be applied to any poetic text regardless of its line lengths and couplet structure. The algorithm proposed below is meant to provide such a measure, while at the same time taking care of overlapping text units.

The New Algorithm

Before describing the algorithm, which was implemented in Python 3 for the sake of this study, it is useful to say a few words about the notion of *formula*. I follow Vikis-Freibergs and Freibergs' definition of a syntagmatic formula (as opposed to paradigmatic formulas, where elements can be substituted, and formulas as syntagmatic patterns) as "any syntagmatic string of units occurring at least twice in any given corpus" (1978a:331).

A case can be made against studying syntagmatic formulas in their own right since

formulas-as-patterns are no less prominent a feature of oral poetics. However, there is a good reason, at least at the moment, to do exactly this. Theoretically, it is possible to extract strings of words with identical syntactical features in order to identify variable-content formulas—that is, formulas-as-patterns if we do not demand a certain verbal overlap, or paradigmatic formulas if we do. However, this model/method is possible only for languages with good morphological analyzers, and these languages are precious few and do not include northern varieties of Russian in which *bylinas* are composed. Moreover, the application of this enhanced algorithm could not be universal since the definition of syntactically identical words will vary between languages, and the algorithm developed for this study is meant to be applicable to any text separated into words and lines.

I also follow Vikis-Freibergs and Freibergs in disregarding the position of a formula in a line. There are two main reasons for doing this. First, it is very hard to analyze metric structures automatically (moreover, many oral traditional texts, including Russian *bylinas*, are rather loose metrically and allow for additional words to be added at beginnings and endings of otherwise identical lines). Secondly, repeated text units found in different metrical positions undermine Parry and Lord's theory of oral poetry, which is predicated on the principle of maximum parsimony of formulaic elements used by oral poets, but there is no evident need to disqualify them as formulas *sensu latu*.

However, when applying formulas defined in this manner to large corpora the results are flooded with prepositional phrases (“to the house”), semantically impoverished pronoun-verb pairs (“he went”), and all sorts of combinations involving function words (“and then,” “you did,” and so forth). These clearly can not be regarded as formulas in any meaningful sense. Therefore, it is necessary to maintain a stop-list of the most frequent function words used in the texts under investigation and filter out the combinations with them.

The following stop-list was used for the analysis of *bylinas*:

а, а-й, ай, без, бы, в, вам, вас, ведь, во, вот, все, всё, вы, где, гой, да, дак, де, до, е, его, ей, ему, еси, есь, еще, же, за, и, из, им, их, ише, ишше, ище, й, к, ка, кабы, как, ко, ле, ли, м, меня, мне, мня, мой, моя, мы, на, нам, нас, не, нет, ни, о, ой, он, она, они, от, по, под, про, ро, с, сам, своей, свою, со, т, там, те, тебе, тебя, то, того, той, тому, тот, ту, тут, ты, тя, у, уж, х, что, я, ёго, ёму

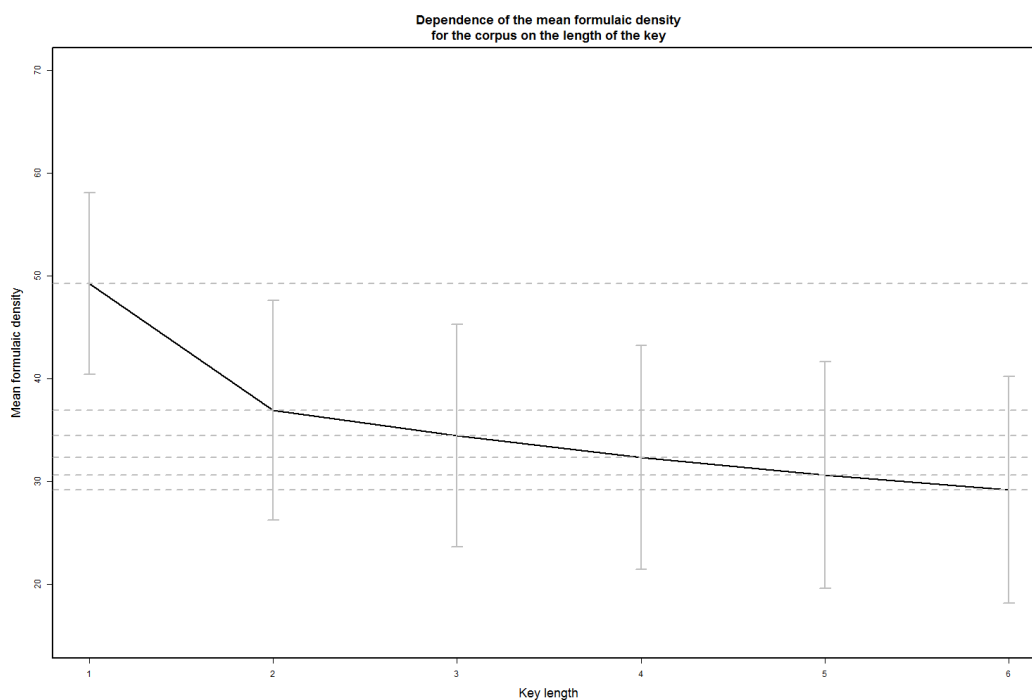
[a, a-], aj, bez, by, v, vam, vas, ved', vo, vot, vse, vsjo, vy, gde, goj, da, dak, de, do, e, ego, ej, emu, esi, es', eshhe, zhe, za, i, iz, im, ih, ishe, ishshe, ishhe, j, k, ka, kaby, kak, ko, le, li, m, menja, mne, mnja, moj, moja, my, na, nam, nas, ne, net, ni, o, oj, on, ona, oni, ot, po, pod, pro, ro, s, sam, svoej, svoju, so, t, tam, te, tebe, tebja, to, togo, toj, tomu, tot, tu, tut, ty, tja, u, uzh, h, chto, ja, jogo, jomu]

Additionally, when dealing with highly inflected languages, such as Russian, it is desirable to be able to conflate into a single formula the same prepositional or noun phrases with different case endings, like *в занней горнице* (“in the back room”) and *в занню горницу* (“to the back room”). The preferable strategy would be to lemmatize each word in the text and to use strings of lemmas instead of the original token strings. However, the orthography of the written texts of *bylinas* rather faithfully reflects the features of the singers' varying dialects, and no

existing lemmatizing algorithm would be able to sort out the forms. Fortunately, Russian words are generally rather long, and it is possible to simply disregard all the letters from the fifth on. The above mentioned prepositional phrases then become equivalent: *в занней горнице / в занню горницу* [v zannej gornitse / v zannu gornitsu].

It turns out that the probability of different non-trivial phrases having the same 4-symbol word-beginnings is rather slight. To measure it I randomly took five texts from the corpus and investigated all the formulas extracted from them by the algorithm. Only one text of five produced spurious formulas: there were two of them, together comprising 10 words—out of 597 words in all the formulas in this text. Consequently we should expect our algorithm to exaggerate the formulaic density of texts by approximately 1.5%.

It can also be noted that there is no truly significant difference between using 4-symbol keys and full words. The following chart shows that the difference in the mean formulaic density for the corpus computed with 4-symbol keys and 5-symbol keys (full words, essentially) does not exceed 2.5% (vertical bars indicate standard deviation):



A change from 4- to 3-symbol keys also yields the difference of less than 3%. Only off-kilter 2- and 1-symbol keys change the mean formulaic density dramatically.

Therefore, the following definition was used for this study:

A **formula** is a non-trivial syntagmatic string of words occurring at least n times in a given corpus, where *non-trivial* means that

- (1) if a formula consists of two words both of them should be absent from the stop list;
- (2) if a formula consists of more than two words, not more than half of them (rounding down) should be present in the stop list.

The occurrences of a formula are counted based on the occurrences of its *formulaic key*, where a *formulaic key* is a concatenation of the m -letter beginnings of all the words in the string. I use concatenation in the sense of linking/chaining character strings together. For instance, concatenation of “one” and “two” yields “onetwo.”

n was taken to be 2, and m to be 4 in the present study.

Finally, I take **formulaic density** to be the *ratio of the words* of the text used in the formulas in it.

Based on these definitions the algorithm proceeds as follows:

1. A text, read from a file and stripped of all punctuation signs, is first converted to a 2-dimensional array where each subarray corresponds to a line in the original text and every element in these subarrays is an instance of a special data structure, PoeticWord. A PoeticWord consists of:

- i) the original word string,
- ii) a Boolean value serving as an indicator as to whether this word was used in a formula, and
- iii) a list of text units (N-grams) to which it belongs and which track its state.

When a PoeticWord is used, it sends a signal to the N-grams containing it, which then become blocked to prevent overlap.

2. Overlapping N-grams of different lengths (a lower threshold of 2 and an upper threshold of 14 were used in this study) are extracted from subarrays corresponding to the lines of the original text. Each N-gram observes the state of all the PoeticWords it contains.

At the same time these N-grams are added to a dictionary where their formulaic keys (as defined above) are used as the keys and the entries are lists of N-grams having the same key. The keys are also added to the special key list where they are stored in the decreasing order of length. Inside the same-length runs they are stored in the order of occurrence in the original text.

A random 5-element snippet from the N-gram dictionary extracted from one of the *bylinas* in the corpus looks as follows:

подазлатсереведь: *Подарил-то злато серебро ведь*
 наширодвор: *на широкой двор; на широкой двор; на широкой двор; на широкой двор; На широком двори*
 уживадамыведькожусьдибуде: *У жива да мы ведь кожу-ту съдирать будем*
 тутзаплнашода: *Тут заплакало-то нашо-то да*
 взелелужк: *в зелены лужка; В зелены лужка*

[podazlatsereved': *Podaril-to zlato srebro ved'*
 nashirodvor: *na shirokoj dvor; na shirokoj dvor; na shirokoj dvor; na shirokoj dvor; Na shirokom dvori*

uzhivadamyved'kozhus'dibude: *U zhiva da my ved' kozhu-tu s'dirat' budem*
 tutzaplnashoda: *Tut zaplakalo-to nasho-to da*
 vzeleluzhk: *v zeleny luzhka; V zeleny luzhka]*

The second and fifth items in this sample contain N-grams which occur more than once in the text.

3. The keys from the key list are taken one by one. If the corresponding N-gram list from the N-gram dictionary is of length one, it means that this N-gram is unique in the corpus and the next key is taken.

If the length of the N-gram list in the corresponding dictionary entry is greater than one, the first non-blocked N-gram from it is added to a temporary list and all the words in this N-gram are marked as used (which blocks all other N-grams containing these words). If there are other unblocked N-grams in the list in this dictionary entry, they are also added to the temporary list, which is then added as an entry to the resulting formula dictionary with the same formulaic key used as the key. If there is only one unblocked N-gram in the list in the original-dictionary entry, it is eventually unblocked since the words in it are not yet used in any formula. If there are no unblocked N-grams at all, nothing happens and the next key is taken.

When the list is exhausted all the formulas from the text are contained in the formula dictionary and can be retrieved. A random snippet of the formula dictionary for the same *bylina* looks as follows:

всезелевина: *все зелена вина; все зелена вина*
 красзолло: *красно золото; красна золота; красным золотом*
 тутильядаильямура: *тут Илья да Илья Мурамецъ; тут Илья да Илья Мурамецъ*
 ильюмура: *Илью Мурамця; Илью Мурамця; Илью Мурамця; Илью Мурамця; Илью Мурамця*

vsezelevina: *vse zelena vina; vse zelena vina*
 [kraszolo: *krasno zoloto; krasna zolota; krasnym zolotom*
 tutil'jadail'jamura: *tut Il'ja da Il'ja Muramec'; tut Il'ja da Il'ja Muramec'*
 il'jumura: *Il'ju Muramecja; Il'ju Muramecja; Il'ju Muramecja; Il'ju Muramecja; Il'ju Muramecjaa]*

The formulaic density is then computed by counting the used PoeticWords in the 2-dimensional array and dividing this number by the total number of words.¹

Following this procedure we obtain a list of non-overlapping and non-embedded formulas. Repeated occurrences of a word pair, for instance, are counted only if they do not form part of some repeated N-gram of a greater length. The algorithm is flexible: all the parameters (the number of occurrences of an N-gram sufficient to prove that it is a formula, the definition of an equivalent pair of N-grams) can be tweaked independently. The algorithm can process poetic texts in any language provided that it is divided into words and lines and that the algorithm

¹ The program presented in the appendix also maintains a list of first and last words of all formulas, which makes it possible to provide a printout of the original texts with formulas in brackets and cross references between lines with the same formulas. Below I use this feature to present one fully analyzed *bylina*.

knows the relevant alphabet and the stop list (which can be made empty). Finally, it should be pointed out that the algorithm is sufficiently fast to take on really large corpora: a huge document compiled from 609 Russian *bylina* texts and containing more than 140,000 lines was analyzed in less than a minute.

In the following section I report the results of applying this algorithm to the a corpus of Russian *bylina* epics, but first I provide an example of full analysis of a text from the corpus. Every formula is accompanied by references to other lines in the text where it appears; semicolons separate references for different formulas.

Dobrynya and the dragon (Selivanov 1988)

1	[Dobrynjushke-to matushka govarivala]	24
2	Da Nikitichu-to matushka nakazyvala	
3	Ty [ne ezdi-ka daleche vo chisto pole]	26, 69, 100
4	Na tu na goru da Sorochinskiju	
5	Ne topchi-ka ty [mladyh zmeenyshej]	28, 71, 101, 166, 168, 182, 199
6	Ty [ne vyruchaj-ka polonov da russkiih]	29, 72
7	[Ne kuplis' Dobrynja vo Puchaj-reke]	30
8	[Puchaj-reka ochen' svirepaja]	31
9	[Serednjaja-to strujka kak ogon' sechet]	32
10	Dobrynja svoej matushki ne slushalsja	
11	Kak on edet [daleche vo chisto pole]	164, 179
12	Na tu [na goru na Sorochinskiju]	27, 70, 90
13	[Potoptal on mladyih zmeenyshej]	181
14	Povyruchal on [polonov da russkiih]	76, 167, 200
15	Bogatyrsko ego serdce raspotelosja	
16	Raspotelosja serdce nazhadelosja	
17	On pripravil svoego [dobra konja]	18, 19, 53, 144
18	On [dobra konja] da ko Puchaj-reke	17, 19, 53, 144
19	On slezal Dobrynja so [dobra konja]	17, 18, 53, 144
20	Da snimal [Dobrynja plat'e cvetnoe]	54
21	On [zabrel za struechku za] pervuju	22
22	Da [zabrel za struechku za] srednjuju	21
23	Govoril sam da takovo slovo	
24	Mne [Dobrynjushke matushka govarivala]	1
25	Mne Nikitichu mamen'ka nakazyvala	
26	Chto [ne ezdi-ka daleche vo chisto pole]	3, 69, 100
27	Na tu [na goru na Sorochinskiju]	12, 70, 90
28	Ne topchi-ka [mladyh zmeenyshej]	5, 71, 101, 166, 168, 182, 199
29	[Ne vyruchaj polonov da russkiih]	6, 72
30	I [ne kuplis' Dobrynja vo Puchaj-reke]	7
31	[Puchaj-reka ochen' svirepaja]	8
32	[Serednjaja strujka kak ogon' sechet]	9

33	A Puchaj-reka ona krotka-smirna	
34	Ona budto luzha-to dozhdevaja	
35	Ne uspel Dobrynja slovca smolviti	
36	Vetra net da tuchu nadneslo	
37	Tuchi net da budto dozhd' dozhdit	
38	A dozhdja-to net da tol'ko [grom gremit]	39
39	[Grom gremit] da svishhet molnija	38
40	Kak letit zmeishhe Gorynishhe	
41	O tyeih dvenadcati o hobotah	
42	Dobrynja toj Zmei ne priuzhahnetsja	
43	Govorit Zmeja emu prokljataja	
44	Ty teper' Dobrynja vo moih rukah	
45	[Zahochu tebjja Dobrynju teper'] potoplju	46
46	[Zahochu tebjja Dobrynju teper'] s'em-sozhru	45
47	Zahochu tebjja Dobrynju [v hobota voz'mu]	48
48	[V hobota voz'mu] [Dobrynju vo noru] snesu	47; 239
49	[Pripadaet Zmeja] ko bystroj reke	83
50	A Dobrynjushka plavat' gorazd ved' byl	
51	[On nyrnet na berezhok] na tamoshnij	52
52	[On nyrnet na berezhok] na zdeshnij	51
53	Net u Dobrynjushki [dobra konja]	17, 18, 19, 144
54	Da net u [Dobryni plat'ev cvetnyih]	20
55	Tol'ko lezhit odin [puhov kolpak]	56
56	[Puhov kolpak] da [zemli Grecheskoj]	55; 58
57	Po vesu tot kolpak da celyh tri puda	
58	Kak uhvatil on kolpak [zemli Grecheskoj]	56
59	Da shibnet vo Zmeju vo prokljatuju	
60	On otshib Zmee dvenadcat' hobotov	
61	Tut upala Zmeja da vo kovyl'-travu	
62	Dobrynjushka na nozhku povertok byl	
63	Skochit on na zmeinye da [grudi belye]	65
64	Na kreste u Dobryni byl bulatnyj nozh	
65	Hochet on rasplastat' ej [grudi belye]	63
66	A Zmeja emu Dobryne vzmolitsja	
67	Oj ty [Dobrynja syn Nikitinich]	161
68	My polozhim s toboj zapoved' velikuju	
69	Tebe [ne ezditi daleche vo chisto pole]	3, 26, 100
70	Na tu [na goru na Sorochinskuju]	12, 27, 90
71	Ne topat' bol'she [mladyh zmeenyshej]	5, 28, 101, 166, 168, 182, 199
72	[Ne vyruchat' polonov da russkiih]	6, 29
73	Ne kupat'sja tebe Dobrynja vo Puchaj-reke	
74	I mne ne letat' da [na Svjatuju Rus']	99
75	Ne nosit' ljudej mne bol'she russkiih	

76	Ne kopit' mne [polonov da russkiih]	14, 167, 200
77	On povypustil Zmeju kak s-pod kolen svoih	
78	Podnjalas' Zmeja da vverh pod oblaku	
79	Sluchilos' ej letet' da mimo Kiev-grada	
80	Uvidala [on Knjazevu plemjannicu]	103
81	[Molodu Zabavu doch' Putjatichnu]	93, 104, 111, 134, 205, 248
82	Iduchis' po ulice po širokoj	
83	Tut [pripala Zmeja] da ko syroj zemle	49
84	Zahvatila ona [Knjazevu plemjannicu]	92, 110, 133, 204, 206, 210
85	Unesla [vo noru vo glubokuju]	132
86	Togda [solnyshko Vladimir stol'nokievskij]	95, 106, 129
87	Po tri dnja da tut [bilich klikal]	88
88	A [bilich klikal] da slavnyh rycarej	87
89	Kto by mog [s"ezdit' daleche vo chisto pole]	109, 130
90	Na tu [na goru na Sorochinskiju]	12, 27, 70
91	Shodit' vo noru da vo glubokuju	
92	Dostat' ego [knjazevu plemjannicu]	84, 110, 133, 204, 206, 210
93	[Molodu Zabavu doch' Putjatichnu]	81, 104, 111, 134, 205, 248
94	Govoril Aleshen'ka Levont'evich	
95	Ah ty [solnyshko Vladimir stol'nokievskij]	86, 106, 129
96	Ty [nakin'-ka jetu sluzhbu da velikuju]	107
97	Na [togo Dobrynju na Nikiticha]	108
98	U nego ved' so Zmeeju zapoved' polozhena	
99	Chto ej ne letat' [na Svjatuju Rus']	74
100	A emu [ne ezdit' daleche vo chisto pole]	3, 26, 69
101	Ne toptat'-to [mladyih zmeenyshej]	5, 28, 71, 166, 168, 182, 199
102	Da ne vyruchat' polonov russkiih	
103	Tak voz'met [on knjazevu plemjannicu]	80
104	[Molodu Zabavu doch' Putjatichnu]	81, 93, 111, 134, 205, 248
105	[Bez boju bez draki-krovolitija]	207, 211
106	Tut [solnyshko Vladimir stol'nokievskij]	86, 95, 129
107	Kak [nakinul jetu sluzhbu da velikuju]	96
108	Na [togo Dobrynju na Nikiticha]	97
109	Emu [s"ezdit' daleche vo chisto pole]	89, 130
110	I dostat' emu [Knjazevu plemjannicu]	84, 92, 133, 204, 206, 210
111	[Molodu Zabavu doch' Putjatichnu]	81, 93, 104, 134, 205, 248
112	On poshel domoj Dobrynja zakruchinilsja	
113	Zakruchinilsja Dobrynja zapechalilsja	
114	Vstrechaet ego da [rodna matushka]	123, 135
115	[Chestna vdova Efim'ja Aleksandrovna]	124, 136
116	Oj ty rozhono moe ditjatko	
117	[Molodoj Dobrynja syn Nikitinich]	186, 217
118	Ty chto s piru nevesel idesh'	

119	Znat' [mesto bylo] tebe ne po chinu	125
120	Znat' [charoj na piru] tebja priobnesli	126
121	Al' durak nad toboj nasmejalsja-de	
122	[Govoril Dobrynja syn Nikitinich]	201
123	Oj ty gosudarynja [rodna matushka]	114, 135
124	Ty [chestna vdova Efim'ja Aleksandrovna]	115, 136
125	[Mesto bylo] mne da po chinu	119
126	[Charoj na piru] menja ne obnesli	120
127	Durak-to nado mnoj ne nasmejalsja ved'	
128	A nakinul sluzhbu da velikuju	
129	[Solnyshko Vladimir stol'nokievskij]	86, 95, 106
130	Chto [s"ezdit' daleche vo chisto pole]	89, 109
131	Na tu na goru da na vysokuju	
132	Mne shodit' [vo noru vo glubokuju]	85
133	Mne dostat'-to [Knjazevu plemjannicu]	84, 92, 110, 204, 206, 210
134	[Molodu Zabavu doch' Putjatichnu]	81, 93, 104, 111, 205, 248
135	[Govorit Dobryne] [rodna matushka]	209; 114, 123
136	[Chestna vdova Efim'ja Aleksandrovna]	115, 124
137	Lozhis'-ka spat' da rano s vechera	
138	Mudrenee utro budet vechera	
139	On vstavaj po utrechku raneshen'ko	
140	Umyvalsja da on beleshen'ko	
141	Snarjazhalsja on horoshohon'ko	
142	Da idet na konjushnju na stojaluju	
143	A beret v ruki uzdu on da tesmjanuju	
144	A beret on dedushkova da ved' [dobra konja]	17, 18, 19, 53
145	On poil Burka pit'em medvjanyim	
146	On kormil pshenoj da belojarovoj	
147	Sedlal Burka v sedlyshko cherkasskoe	
148	On [potnichki da klal] na potnichki	149
149	On na [potnichki da klal] vojlochki	148
150	Klal na vojlochki cherkasskoe sedlyshko	
151	Vse podtjagival dvenadcat' tugih podprugov	
152	On trinadcatyj klal da radi kreposti	
153	Chtoby dobryj kon' s-pod sedla ne vyskochil	
154	Dobra molodca v chistom pole ne vyrutil	
155	Podprugi byli shelkovye	
156	A shpen'ki u podprug vse bulatnye	
157	Prjazhki u sedla da [krasna zolota]	159
158	Tot shelk ne rvetsja bulat ne tretsja	
159	[Krasno zoloto] ne rzhaveet	157
160	Molodec na kone sidit da sam ne stareet	
161	Poezhal [Dobrynja syn Nikitinich]	67

162	Na proshhan'e emu matushka pletku podala	
163	[Sama govorila takovo slovo]	196
164	Kak budesh' [daleche vo chistom pole]	11, 179
165	Na toj na [gore da na vysokija]	180
166	Potopchesh' [mladyih zmeenyshej]	5, 28, 71, 101, 168, 182, 199
167	Povyruichish' [polonov da russkiih]	14, 76, 200
168	Kak tyi-to [mladye zmeenyshej]	5, 28, 71, 101, 166, 182, 199
169	[Podtochat u Burka oni shhetochki]	183
170	[Chto ne mozhet bol'she Burushko poskakivat']	184
171	A [zmeenyshej ot nog da on otrjahivat']	177, 185, 192
172	Ty voz'mi-ka jetu [pletochku shelkovuju]	187
173	A ty bej [Burka da promezhu] nogi	188
174	[Promezhu nogi] da [promezhu ushi]	189; 189
175	[Promezhu nogi da mezhu zadnie]	190
176	Stanet tvoj [Burushko poskakivat']	191
177	[Zmeenyshej ot nog da on otrjahivat']	171, 185, 192
178	Ty pritopchesh' [vseh do edinogo]	193
179	Kak budet on [daleche vo chistom pole]	11, 164
180	Na toj na [gore da na vysokoi]	165
181	[Potoptal on mladyh zmeenyshej]	13
182	Kak te li [mladye zmeenyshej]	5, 28, 71, 101, 166, 168, 199
183	[Podtochili u Burka oni shhetochki]	169
184	[Chto ne mozhet bol'she Burushko poskakivat']	170
185	[Zmeenyshej ot nog da on otrjahivat']	171, 177, 192
186	Tut [molodoj Dobrynja syn Nikitinich]	117, 217
187	Beret on [pletochku shelkovuju]	172
188	On b'et [Burka da promezhu] ushi	173
189	[Promezhu ushi] da [promezhu nogi]	174; 174
190	[Promezhu nogi da mezhu zadnie]	175
191	Tut stal ego [Burushko poskakivat']	176
192	A [zmeenyshej ot nog da on otrjahivat']	171, 177, 185
193	Pritoptal on [vseh do edinogo]	178
194	Vyhodila [Zmeja ona prokljataja]	208
195	Iz toj iz nory iz glubokoi	
196	[Sama govorila takovo slovo]	163
197	[Ah ty jej Dobrynjushka] Nikitinich	228
198	Ty znat' porushil svoju zapoved'	
199	Zachem stoptal [mladyih zmeenyshej]	5, 28, 71, 101, 166, 168, 182
200	Pochto vyruchal [polony da russkie]	14, 76, 167
201	[Govoril Dobrynja syn Nikitinich]	122
202	Ah ty jej Zmeja da ty prokljataja	
203	Chert li tja nes da cherez Kiev-grad	
204	Ty zachem vzjala [knjazevu plemjannicu]	84, 92, 110, 133, 206, 210

205	[Molodu Zabavu doch' Putjatichnu]	81, 93, 104, 111, 134, 248
206	Ty otdaj zhe mne [knjazevu plemjannicu]	84, 92, 110, 133, 204, 210
207	[Bez boju bez draki-krovolitija]	105, 211
208	Togda [Zmeja ona prokljataja]	194
209	[Govorila-to Dobryne] da Nikitichu	135
210	Ne otdam ja tebe [knjazevoj plemjannicy]	84, 92, 110, 133, 204, 206
211	[Bez boju bez draki-krovolitija]	105, 207
212	Zavodila ona boj-draku velikuju	
213	Oni dralis' [troi sutochki]	224
214	No [ne mog Dobrynja] Zmeju perebit'	225
215	Hochet tut Dobrynja ot Zmei otstat'	
216	Kak [s nebes Dobryne] [glas glasit]	227; 227
217	[Molodoj Dobrynja syn Nikitinich]	117, 186
218	Dralsja so Zmeej [ty troi sutochki]	229
219	Poderis' [so Zmeeju eshhe tri chasa]	221
220	Ty pob'esh' [Zmeju da tu prokljatuju]	222
221	On podralsja [so Zmeeju eshhe tri chasa]	219
222	On pobil [Zmeju da tu prokljatuju]	220
223	Ta Zmeja ona krov'ju poshla	
224	Stojal u Zmei on [troi sutochki]	213
225	[Ne mog Dobrynja] krovi perezhdat'	214
226	Hotel Dobrynja ot krovi otstat'	
227	[S nebes Dobryne] opjat' [glas glasit]	216; 216
228	[Ah ty jej Dobrynja] syn Nikitinich	197
229	Stojal u krovi [ty troi sutochki]	218
230	Postoj u krovi da eshhe tri chasa	
231	Beri svoe kop'e da burzameckoe	
232	I bej kop'em da vo syru zemlju	
233	Sam kop'ju da progovarivaj	
234	Rasstupis'-ka [matushka syra zemlja]	237
235	Na chetyre rasstupis' da ty na chetverti	
236	Ty pozhri-ka jetu [krov' da vsju zmeinuju]	238
237	Rasstupilas' togda [matushka syra zemlja]	234
238	Pozhrala ona [krov' da vsju zmeinuju]	236
239	Togda [Dobrynja vo noru] poshel	48
240	Vo te vo nory da vo glubokie	
241	Tam sidjat [sorok carej] [sorok carevichej]	; ²
242	Sorok [korolej da korolevichej]	246
243	A prostoj-to sily toj i smetu net	
244	Togda Dobrynjushka Nikitinich	

² These marks indicate the second instance of this formula (which is actually a false positive discussed later in the text) appears on the same line as first one (and vice versa—hence two marks).

245	Govoril-to on carjam da on carevicham	
246	I tem [koroljam da korolevicham]	242
247	Vy idite nyn' tuda otkel' prineseny	
248	A ty [moloda Zabava doch' Putjatichna]	81, 93, 104, 111, 134, 205
249	Dlja tebja ja jedak teper' stranstvoval	
250	Ty poedem-ka ko gradu ko Kievu	
251	A j ko laskovomu knjazu ko Vladimiru	

In line 241, we see a false positive: *sorok carej* and *sorok carevichej* are regarded as a formula. *Sorok carej* means “forty tsars,” and *sorok carevichej* means “forty princes.” *Prince* is derived from *tsar* by means of a suffix, which misleads the algorithm. On the whole, however, the method seems to be remarkably robust.

This type of representation can be used to analyze different types of verbal repetitions found in the texts. It also can be used to investigate the locality patterns in formula repetitions. The following diagram was produced by connecting all the lines that have the same formulae (lines are arranged from left to right):

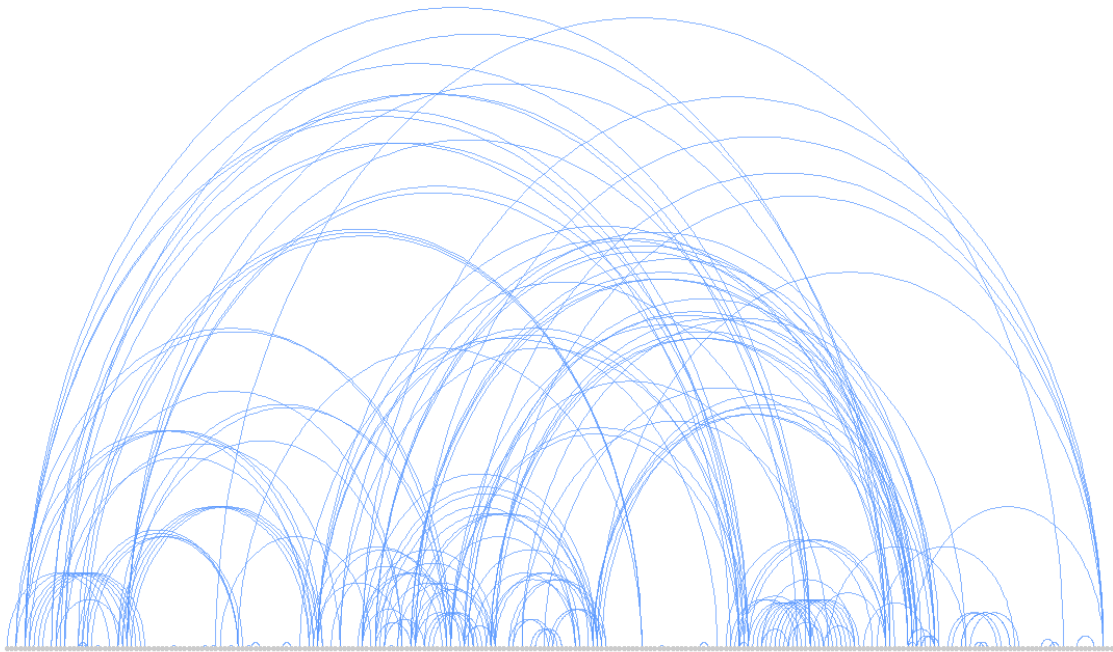


Fig 1: An arc diagram of formulaic connections between the lines of the bylina *Dobrynya and the dragon*. Points on the horizontal axis represent the lines of the text arranged from left to right, and the arcs connect the lines having the same formulae. It was made in R using package “*arcDiagram*” by Gaston Sanchez available at <https://github.com/gastonstat/arcDiagram>

We see that formulaic repetitions tend to be nonlocal—that is, they connect distant lines—and also to some degree symmetric. Instead of simply helping the oral poet to improvise the text, they also help him structure *bylina* as a whole.

Formulaic Density of Russian *bylinas*

For the sake of this study a corpus was assembled containing the texts from major editions of Russian *bylinas* (Azbeleva and Marčenko 2002; Astahova 1958, 1961; Gorelov 2001a, 2001b, 2003, 2004; Selivanov 1988; Smirnov 1978; Smirnov and Smolickij 1974). After removing the duplicates and the texts shorter than 100 lines, which were mostly incomplete, the resulting corpus contained 609 texts and over 140,000 lines.

There are two possible ways to compute the formulaic density of the texts in this corpus:

- (1) Since the texts are sufficiently long it is feasible to compute their formulaic density based exclusively on the internal repetitions (this measure can be tentatively called *internal formulaic density* of a given text) and then take the average or investigate the quantiles.
- (2) It is also possible to compute all the recurrent lines in all the texts; in this case it seems reasonable to raise the minimum number of repetitions that constitutes a formula since, on the one hand, corpus contains very close variants of the same *bylinas*, and, on the other hand, there is a greater probability of a false match and of extracting meaningless word groups.

Since both these procedures are easily performed using the new algorithm, both the internal formulaic densities and the overall density were computed.

The internal formulaic densities of the texts in the corpus have the following histogram:

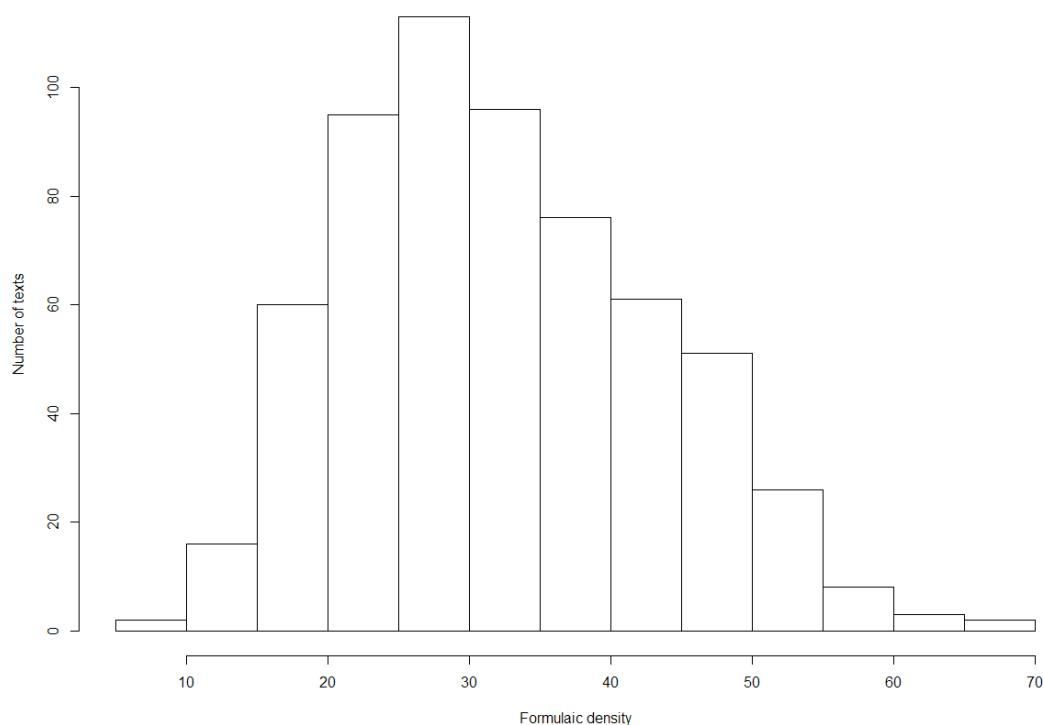


Fig 2: A histogram of formulaic densities of the texts of *bylinas* in the corpus. The full list of values can be found in Appendix I.

This sample has a mean of 32% and a standard deviation of 11, which indicates, roughly, that an average Russian *bylina* is 22-42% internally formulaic (minimum and maximum values are 7.3% and 70% respectively). Also it can be noted that 75% of *bylinas* are more than 24.3% internally formulaic and 25% of *bylinas* are more than 40% internally formulaic.

To compute the overall density of the corpus, a mega-document containing all the 609 texts was compiled. Setting the minimum number of repetitions of an N-gram to 5 generates a formulaic density of 34.7%, which is predictably (though very slightly) higher than the average internal formulaic density of singleton texts. Counting any duplicate items, the overall formulaic density rises to 56%.

Finally, to illustrate the applicability of the algorithm to texts in different languages using different orthographies, the formulaic density of the *Iliad* was calculated based on the edition by Allen and Munro (1920; 1922). Formulaic keys were not used; only exact N-gram matches were taken into account. With the provisional stop list comprising the most frequent function words of the *Iliad* we get the formulaic density of 44.8%. With an empty stop list this number rises to 56.6%.

Conclusion

The algorithm presented in this paper provides an adequate means of estimating the formulaic density of a poetic text or a corpus of poetic texts in any language provided that the texts are divided into lines and words. The algorithm is fast, it is immune to over-counting, and, with certain modifications, it can be used to extract different types of formulas: simple verbatim repetitions, strings of identical lemmas, provided an additional lemmatizing module is used, or even sets of words if the order of the elements in the formula is not crucial.

The implementation of the algorithm used in this study is available at the GitHub repository (<https://github.com/macleginn/fast-formulaic-analysis>); its usage is described in Appendix II. It requires a working distribution of Python 3, but no knowledge of Python itself is needed. All the analysis parameters—alphabet, stop-list, number of symbols used in the N-gram key, and required number of occurrences of an N-gram—are specified in a simple configuration file.

*Russian State University for Humanities /
The Russian Presidential Academy of National Economy
and Public Administration (Moscow)*

APPENDIX I

Internal formulaic densities of the *bylinas* from the corpus used in the study:

17.8, 20.6, 20.7, 18.6, 17.3, 22.1, 13.3, 23.5, 26.6, 19.2, 15.1, 16.1, 27.9,
 28.0, 19.4, 25.1, 16.7, 17.1, 22.5, 13.5, 21.4, 22.9, 29.9, 25.9, 21.3, 13.0,
 17.6, 20.4, 12.2, 11.5, 24.2, 17.4, 23.9, 13.0, 20.1, 18.9, 15.6, 18.6, 19.8,
 23.1, 22.8, 22.7, 12.1, 14.0, 24.7, 15.6, 22.8, 20.0, 18.9, 16.1, 12.2, 11.6,
 15.3, 25.0, 26.9, 26.9, 29.8, 23.9, 23.0, 37.1, 36.3, 27.2, 27.5, 34.1, 30.0,
 33.0, 30.6, 25.4, 28.2, 35.5, 23.2, 21.2, 19.8, 17.4, 24.8, 17.8, 37.5, 53.4,
 33.3, 24.2, 31.2, 40.8, 35.0, 39.6, 30.4, 29.8, 32.2, 40.3, 46.2, 32.8, 35.9,
 23.6, 38.1, 27.0, 24.1, 28.8, 13.6, 18.5, 36.2, 46.9, 53.4, 24.4, 38.6, 25.9,
 34.7, 34.5, 33.2, 32.4, 42.2, 33.0, 30.5, 35.8, 46.4, 29.6, 50.3, 35.8, 56.1,
 30.1, 16.3, 36.7, 49.4, 45.7, 48.7, 33.4, 46.0, 49.4, 49.3, 40.1, 62.0, 30.4,
 34.7, 50.5, 42.8, 51.8, 30.3, 26.9, 45.3, 47.9, 8.1, 16.6, 29.0, 36.8, 35.1,
 20.7, 26.2, 27.2, 30.7, 25.8, 23.3, 21.0, 22.5, 24.5, 16.9, 29.0, 21.0, 22.2,
 19.4, 28.3, 27.7, 26.3, 27.0, 37.4, 20.2, 25.9, 21.2, 25.2, 38.0, 32.4, 22.5,
 19.6, 26.3, 30.7, 25.4, 44.9, 24.8, 15.0, 30.5, 44.6, 44.9, 29.6, 21.4, 27.2,
 30.6, 43.7, 42.3, 25.6, 43.3, 55.5, 50.5, 42.3, 37.8, 41.2, 26.8, 46.3, 37.8,
 26.7, 22.8, 34.6, 38.3, 28.5, 33.9, 44.1, 43.8, 30.0, 39.1, 36.7, 30.7, 32.1,
 37.8, 31.3, 25.2, 35.4, 34.5, 23.5, 42.6, 34.8, 44.2, 27.1, 22.6, 25.0, 35.5,
 36.8, 19.8, 26.4, 47.3, 55.0, 51.5, 28.3, 32.1, 49.7, 43.0, 41.5, 25.9, 29.2,
 19.9, 35.3, 31.4, 25.1, 50.9, 39.6, 36.9, 17.3, 35.9, 31.5, 32.6, 16.2, 23.8,
 36.4, 27.4, 28.4, 34.9, 26.9, 48.7, 41.5, 25.7, 23.8, 21.7, 15.6, 28.0, 21.7,
 28.4, 20.1, 20.2, 22.2, 28.2, 24.6, 17.8, 24.7, 29.8, 23.4, 21.2, 25.4, 17.6,
 40.4, 30.7, 47.9, 37.6, 43.3, 40.9, 33.1, 34.6, 37.0, 40.7, 34.7, 47.0, 44.0,
 28.5, 33.4, 40.7, 40.8, 48.2, 38.0, 52.1, 25.5, 15.7, 23.6, 53.7, 49.5, 54.8,
 22.7, 23.1, 30.7, 41.6, 40.3, 39.1, 43.9, 29.5, 34.5, 44.9, 31.8, 42.1, 37.1,
 23.4, 34.2, 26.6, 27.5, 32.3, 26.4, 14.6, 19.8, 18.7, 26.7, 26.7, 18.0, 48.2,
 52.7, 49.5, 42.2, 31.4, 35.9, 23.0, 52.4, 41.5, 21.6, 29.4, 30.6, 37.3, 30.4,
 37.8, 17.7, 47.1, 32.5, 70.0, 47.5, 31.4, 41.3, 35.3, 33.1, 32.7, 25.3, 32.7,
 38.7, 40.9, 33.2, 29.5, 42.3, 38.1, 29.7, 39.5, 40.1, 31.8, 39.2, 31.6, 31.0,
 31.1, 20.1, 27.7, 37.5, 29.6, 38.9, 47.5, 42.2, 26.2, 38.9, 50.4, 38.8, 27.7,
 26.7, 47.7, 28.0, 24.9, 44.1, 33.2, 37.9, 42.0, 38.0, 41.2, 58.1, 34.5, 35.5,
 13.8, 47.0, 52.7, 16.3, 40.3, 32.9, 32.5, 20.3, 30.8, 34.7, 41.5, 28.7, 52.7,
 53.3, 51.8, 42.7, 46.7, 50.1, 35.4, 66.0, 32.6, 18.8, 28.4, 31.6, 26.1, 47.6,
 29.7, 33.4, 45.8, 26.3, 47.7, 20.8, 24.4, 28.7, 21.9, 23.8, 30.5, 28.6, 30.5,
 50.7, 27.4, 29.6, 31.0, 29.9, 24.3, 36.0, 23.6, 34.8, 26.4, 24.3, 29.7, 36.2,
 24.3, 37.3, 44.6, 48.9, 32.2, 43.8, 47.9, 32.7, 24.7, 32.4, 23.4, 24.7, 26.0,
 58.8, 32.6, 34.8, 7.3, 28.5, 22.2, 28.4, 47.5, 46.5, 38.6, 28.0, 16.6, 33.1,
 17.7, 24.5, 30.7, 39.6, 23.2, 40.5, 29.7, 10.9, 26.5, 33.6, 24.3, 37.0, 32.5,
 47.8, 29.4, 24.4, 26.7, 19.0, 25.0, 16.9, 22.9, 36.6, 43.0, 44.9, 28.9, 33.9,
 39.9, 58.2, 29.2, 31.1, 31.5, 41.5, 40.6, 23.3, 25.5, 17.3, 32.9, 39.8, 30.2,

53.5, 46.9, 44.0, 52.9, 45.6, 22.5, 40.5, 26.1, 36.0, 28.9, 23.0, 22.0, 21.7,
 46.5, 25.0, 37.7, 24.0, 14.0, 27.1, 52.7, 16.6, 56.9, 19.2, 25.0, 40.4, 45.2,
 19.0, 27.0, 24.5, 39.7, 18.4, 20.5, 27.7, 32.2, 38.0, 21.2, 19.8, 25.4, 16.3,
 19.0, 47.5, 15.6, 17.6, 31.6, 27.8, 28.6, 27.0, 40.2, 43.1, 43.2, 39.9, 25.5,
 36.9, 28.7, 30.0, 34.4, 25.4, 37.6, 48.9, 60.9, 29.7, 19.4, 31.5, 43.3, 48.5,
 34.1, 29.8, 24.2, 46.2, 26.2, 45.6, 46.5, 48.1, 37.3, 31.1, 39.9, 36.6, 59.6,
 49.4, 44.4, 35.3, 42.7, 38.5, 45.3, 52.9, 34.1, 37.7, 19.6, 32.6, 23.7, 46.2,
 56.2, 39.2, 35.9, 38.1, 47.7, 60.2, 48.0, 34.2, 21.3, 54.9, 19.4

APPENDIX II

Usage guide for the Python program:

The program consists of four files, which should reside in the same directory:

FormulaicAnalysisLib.py is a library containing classes and functions implementing the algorithm described in this paper. Please refer to it in order to gain a deeper understanding of how the program works.

alpha_stop.conf is a configuration file containing the alphabet, the stop list, the formulaic key length, and the number of occurrences, which make a repeated N-gram qualify as a formula. All these parameters, except stop-list, must be non-empty.

compute_formulaic_density.py is a script that computes formulaic density of a single text based on the parameters in the configuration file. The usage is:

```
python3 compute_formulaic_density.py [options] pathToFileWithText
```

Options include `-s` (show formulas on the screen) and `-f` (create an html file with the text with formulas in square brackets and cross-references). These arguments can be used separately or combined as `-sf` or `-fs`. If the path is a directory, a random file from that directory will be selected.

compute_formulaic_density_batch.py is a script for computing formulaic density of all files in a given directory. The usage is:

```
python3 compute_formulaic_density_batch.py pathToDirectory
```

Text files should be in UTF-8 and their names should not start with a dot or period “.” as temporary or hidden files created by the operating system sometimes start with dots or periods, and this restriction prevents the algorithm from trying to process them. The script prints the number of texts (files), lines, and words on the screen, as well as the mean formulaic density, sample standard deviation, and quartiles. The full report is written to report.csv.

APPENDIX III

Dobrynya and the dragon in Cyrillic script:

	[Добрынюшке-то матушка говаривала]	24
2	Да Никитичу-то матушка наказывала	
3	Ты [не езд-ка далече во чисто поле]	26, 69, 100
4	На ту на гору да Сорочинскую	
5	Не топчи-ка ты [младых змеенышей]	28, 71, 101, 166, 168, 182, 199
6	Ты [не выручай-ка полонов да русскиих]	29, 72
7	[Не куплись Добрыня во Пучай-реке]	30
8	[Пучай-река очень свирепая]	31
9	[Средняя-то струйка как огонь сечет]	32
10	Добрыня своей матушки не слушался	
11	Как он едет [далече во чисто поле]	164, 179
12	На ту [на гору на Сорочинскую]	27, 70, 90
13	[Потоптал он младых змеенышей]	181
14	Повыручал он [полонов да русскиих]	76, 167, 200
15	Богатырско его сердце распотелося	
16	Распотелося сердце нажаделося	
17	Он приправил своего [добра коня]	18, 19, 53, 144
18	Он [добра коня] да ко Пучай-реке	17, 19, 53, 144
19	Он слезал Добрыня со [добра коня]	17, 18, 53, 144
20	Да снимал [Добрыня платье цветное]	54
21	Он [забрел за струечку за] первую	22
22	Да [забрел за струечку за] среднюю	21
23	Говорил сам да таково слово	
24	Мне [Добрынюшке матушка говаривала]	1
25	Мне Никитичу маменька наказывала	
26	Что [не езд-ка далече во чисто поле]	3, 69, 100
27	На ту [на гору на Сорочинскую]	12, 70, 90
28	Не топчи-ка [младых змеенышей]	5, 71, 101, 166, 168, 182, 199
29	[Не выручай полонов да русскиих]	6, 72
30	И [не куплись Добрыня во Пучай-реке]	7
31	[Пучай-река очень свирепая]	8
32	[Средняя струйка как огонь сечет]	9
33	А Пучай-река она кротка-смирна	
34	Она будто лужа-то дождевая	
35	Не успел Добрыня словца смолвити	
36	Ветра нет да тучу наднесло	
37	Тучи нет да будто дождь дождит	
38	А дождя-то нет да только [гром гремит]	39
39	[Гром гремит] да свищет молния	38

40	Как летит змеище Горынище	
41	О тых двенадцати о хоботах	
42	Добрыня той Змеи не приужахнется	
43	Говорит Змея ему проклятая	
44	Ты теперь Добрыня во моих руках	
45	[Захочу тебя Добрыню теперь] потоплю	46
46	[Захочу тебя Добрыню теперь] съем-сожру	45
47	Захочу тебя Добрыню [в хобота возьму]	48
48	[В хобота возьму] [Добрыню во нору] снесу	47; 239
49	[Припадает Змея] ко быстрой реке	83
50	А Добрынюшка плавать горазд ведь был	
51	[Он нырнет на бережок] на тамошний	52
52	[Он нырнет на бережок] на здешний	51
53	Нет у Добрынюшки [добра коня]	17, 18, 19, 144
54	Да нет у [Добрыни платьев цветных]	20
55	Только лежит один [пухов колпак]	56
56	[Пухов колпак] да [земли Греческой]	55; 58
57	По весу тот колпак да целых три пуда	
58	Как ухватил он колпак [земли Греческой]	56
59	Да шибнет во Змею во проклятую	
60	Он отшиб Змее двенадцать хоботов	
61	Тут упала Змея да во ковыль-траву	
62	Добрынюшка на ножку поверток был	
63	Скочит он на змеиные да [груды белые]	65
64	На кресте у Добрыни был булатный нож	
65	Хочет он распластать ей [груды белые]	63
66	А Змея ему Добрыне взмолился	
67	Ой ты [Добрыня сын Никитинич]	161
68	Мы положим с тобой заповедь великую	
69	Тебе [не ездити далече во чисто поле]	3, 26, 100
70	На ту [на гору на Сорочинскую]	12, 27, 90
71	Не топтать больше [младых змееньшей]	5, 28, 101, 166, 168, 182, 199
72	[Не выручать полонов да русских]	6, 29
73	Не купаться тебе Добрыня во Пучай-реке	
74	И мне не летать да [на Святую Русь]	99
75	Не носить людей мне больше русских	
76	Не копить мне [полонов да русских]	14, 167, 200
77	Он повыпустил Змею как с-под колен своих	
78	Поднялась Змея да вверх под облаку	
79	Случилось ей лететь да мимо Киев-града	
80	Увидала [он Князеву племянницу]	103
81	[Молоду Забаву дочь Путятичну]	93, 104, 111, 134, 205, 248
82	Идучись по улице по широкой	

83	Тут [припала Змея] да ко сырой земле	49
84	Захватила она [Князеву племянницу]	92, 110, 133, 204, 206, 210
85	Унесла [во нору во глубокую]	132
86	Тогда [солнышко Владимир стольнокиевский]	95, 106, 129
87	По три дня да тут [билич кликал]	88
88	А [билич кликал] да славных рыцарей	87
89	Кто бы мог [съездить далече во чисто поле]	109, 130
90	На ту [на гору на Сорочинскую]	12, 27, 70
91	Сходить во нору да во глубокую	
92	Достать его [князеву племянницу]	84, 110, 133, 204, 206, 210
93	[Молоду Забаву дочь Путятичну]	81, 104, 111, 134, 205, 248
94	Говорил Алешенька Левонтьевич	
95	Ах ты [солнышко Владимир стольнокиевский]	86, 106, 129
96	Ты [накинь-ка эту службу да великую]	107
97	На [того Добрыню на Никитича]	108
98	У него ведь со Змеею заповедь положена	
99	Что ей не летать [на Святую Русь]	74
100	А ему [не ездить далече во чисто поле]	3, 26, 69
101	Не топтать-то [младых змеенышей]	5, 28, 71, 166, 168, 182, 199
102	Да не выручать полонов русских	
103	Так возьмет [он князеву племянницу]	80
104	[Молоду Забаву дочь Путятичну]	81, 93, 111, 134, 205, 248
105	[Без бою без драки-кроволития]	207, 211
106	Тут [солнышко Владимир стольнокиевский]	86, 95, 129
107	Как [накинул эту службу да великую]	96
108	На [того Добрыню на Никитича]	97
109	Ему [съездить далече во чисто поле]	89, 130
110	И достать ему [Князеву племянницу]	84, 92, 133, 204, 206, 210
111	[Молоду Забаву дочь Путятичну]	81, 93, 104, 134, 205, 248
112	Он пошел домой Добрыня закручинился	
113	Закручинился Добрыня запечалился	
114	Встречает его да [родна матушка]	123, 135
115	[Честна вдова Ефимья Александровна]	124, 136
116	Ой ты рожно мое дитятко	
117	[Молодой Добрыня сын Никитинич]	186, 217
118	Ты что с пиру невесел идешь	
119	Знать [место было] тебе не по чину	125
120	Знать [чарой на пиру] тебя приобнесли	126
121	Аль дурак над тобой насмеялся-де	
122	[Говорил Добрыня сын Никитинич]	201
123	Ой ты государыня [родна матушка]	114, 135
124	Ты [честна вдова Ефимья Александровна]	115, 136
125	[Место было] мне да по чину	119

126	[Чарой на пиру] меня не обнесли	120
127	Дурак-то надо мной не насмеялся ведь	
128	А накинул службу да великую	
129	[Солнышко Владимир стольнокиевский]	86, 95, 106
130	Что [съездить далече во чисто поле]	89, 109
131	На ту на гору да на высокую	
132	Мне сходить [во нору во глубокую]	85
133	Мне достать-то [Князеву племянницу]	84, 92, 110, 204, 206, 210
134	[Молоду Забаву дочь Путятичну]	81, 93, 104, 111, 205, 248
135	[Говорит Добрыне] [родна матушка]	209; 114, 123
136	[Честна вдова Ефимья Александровна]	115, 124
137	Ложись-ка спать да рано с вечера	
138	Мудренее утро будет вечера	
139	Он вставал по утречку ранешенько	
140	Умывался да он белешенько	
141	Снаряжался он хорошоохонько	
142	Да идет на конюшню на стоялую	
143	А берет в руки узду он да тесмяную	
144	А берет он дедушкова да ведь [добра коня]	17, 18, 19, 53
145	Он поил Бурка питьем медвяным	
146	Он кормил пшеной да белояровой	
147	Седлал Бурка в седлышко черкасское	
148	Он [потнички да клал] на потнички	149
149	Он на [потнички да клал] войлочки	148
150	Клал на войлочки черкасское седлышко	
151	Все подтягивал двенадцать тугих подпругов	
152	Он тринадцатый клал да ради крепости	
153	Чтобы добрый конь с-под седла не выскочил	
154	Добра молодца в чистом поле не вырутил	
155	Подпруги были шелковые	
156	А шпеньки у подпруг все булатные	
157	Пряжки у седла да [красна золота]	159
158	Тот шелк не рвется булаг не трется	
159	[Красно золото] не ржавеет	157
160	Молодец на коне сидит да сам не стареет	
161	Поезжал [Добрыня сын Никитинич]	67
162	На прощанье ему матушка плетку подала	
163	[Сама говорила таково слово]	196
164	Как будешь [далече во чистом поле]	11, 179
165	На той на [горе да на высокия]	180
166	Потопчешь [младых змеенышей]	5, 28, 71, 101, 168, 182, 199
167	Повыручишь [полонов да русских]	14, 76, 200
168	Как тыи-то [младые змееныши]	5, 28, 71, 101, 166, 182, 199

169	[Подточат у Бурка они щеточки]	183
170	[Что не может больше Бурушко поскакивать]	184
171	А [змеенышей от ног да он отряхивать]	177, 185, 192
172	Ты возьми-ка эту [плеточку шелковую]	187
173	А ты бей [Бурка да промежу] ноги	188
174	[Промежу ноги] да [промежу уши]	189; 189
175	[Промежу ноги да между задние]	190
176	Станет твой [Бурушко поскакивать]	191
177	[Змеенышей от ног да он отряхивать]	171, 185, 192
178	Ты притопчешь [всех до единого]	193
179	Как будет он [далече во чистом поле]	11, 164
180	На той на [горе да на высокои]	165
181	[Потоптал он младых змеенышей]	13
182	Как те ли [младые змееныши]	5, 28, 71, 101, 166, 168, 199
183	[Подточили у Бурка они щеточки]	169
184	[Что не может больше Бурушко поскакивать]	170
185	[Змеенышей от ног да он отряхивать]	171, 177, 192
186	Тут [молодой Добрыня сын Никитинич]	117, 217
187	Берет он [плеточку шелковую]	172
188	Он бьет [Бурка да промежу] уши	173
189	[Промежу уши] да [промежу ноги]	174; 174
190	[Промежу ноги да между задние]	175
191	Тут стал его [Бурушко поскакивать]	176
192	А [змеенышей от ног да он отряхивать]	171, 177, 185
193	Притоптал он [всех до единого]	178
194	Выходила [Змея она проклятая]	208
195	Из той из норы из глубокои	
196	[Сама говорила таково слово]	163
197	[Ах ты эй Добрынюшка] Никитинич	228
198	Ты знать порушил свою заповедь	
199	Зачем стоптал [младых змеенышей]	5, 28, 71, 101, 166, 168, 182
200	Почто выручал [полоны да русские]	14, 76, 167
201	[Говорил Добрыня сын Никитинич]	122
202	Ах ты эй Змея да ты проклятая	
203	Черт ли ты нес да через Киев-град	
204	Ты зачем взяла [князеву племянницу]	84, 92, 110, 133, 206, 210
205	[Молоду Забаву дочь Путятичну]	81, 93, 104, 111, 134, 248
206	Ты отдай же мне [князеву племянницу]	84, 92, 110, 133, 204, 210
207	[Без бою без драки-кроволития]	105, 211
208	Тогда [Змея она проклятая]	194
209	[Говорила-то Добрыне] да Никитичу	135
210	Не отдам я тебе [князевой племянницы]	84, 92, 110, 133, 204, 206
211	[Без бою без драки-кроволития]	105, 207

212	Заводила она бой-драку великую	
213	Они дрались [трои суток]	224
214	Но [не мог Добрыня] Змею перебить	225
215	Хочет тут Добрыня от Змеи отстать	
216	Как [с небес Добрыне] [глас гласит]	227; 227
217	[Молодой Добрыня сын Никитинич]	117, 186
218	Дрался со Змеей [ты трои суток]	229
219	Подерись [со Змеєю еще три часа]	221
220	Ты побьешь [Змею да ту проклятую]	222
221	Он подрался [со Змеєю еще три часа]	219
222	Он побил [Змею да ту проклятую]	220
223	Та Змея она кровью пошла	
224	Стоял у Змеи он [трои суток]	213
225	[Не мог Добрыня] крови переждать	214
226	Хотел Добрыня от крови отстать	
227	[С небес Добрыне] опять [глас гласит]	216; 216
228	[Ах ты эй Добрыня] сын Никитинич	197
229	Стоял у крови [ты трои суток]	218
230	Постой у крови да еще три часа	
231	Бери свое копьё да бурзамецкое	
232	И бей копьём да во сыру землю	
233	Сам копьё да проговаривай	
234	Расступись-ка [матушка сыра земля]	237
235	На четыре расступись да ты на четверти	
236	Ты пожри-ка эту [кровь да всю змеиную]	238
237	Расступилась тогда [матушка сыра земля]	234
238	Пожрала она [кровь да всю змеиную]	236
239	Тогда [Добрыня во нору] пошел	48
240	Во те во норы да во глубокие	
241	Там сидят [сорок царей] [сорок царевичей]	; ;
242	Сорок [королей да королевичей]	246
243	А простой-то силы той и смету нет	
244	Тогда Добрынюшка Никитинич	
245	Говорил-то он царям да он царевичам	
246	И тем [королям да королевичам]	242
247	Вы идите нынъ туда откель принесены	
248	А ты [молода Забава дочь Путятична]	81, 93, 104, 111, 134, 205
249	Для тебя я эдак теперь странствовал	
250	Ты поедем-ка ко граду ко Киеву	
251	А й ко ласковому князю ко Владимиру	

References

- Allen and Munro 1920 T. W. Allen and D. B. Munro. *Iliad, Books 1-12*. Oxford Classical Texts: Homeri Opera 1. Oxford: Oxford University Press.
- Allen and Munro 1922 _____. *Iliad, Books 13-24*. Oxford Classical Texts: Homeri Opera 2. Oxford: Oxford University Press.
- Arant 1990 Patricia Arant. *Compositional Techniques of the Russian Oral Epic, the Bylina*. New York: Garland.
- Astahova 1958 Anna M. Astahova, ed. *Il'ja Muromec*. Moscow, Leningrad: Izdatel'stvo AN SSSR.
- Astahova 1961 _____, ed. *Byliny Pečory i Zimnego Berega*. Pamjatniki russkogo fol'klora. Moscow, Leningrad: Academy of Science of the USSR.
- Azbelev and Marčenko 2002 Sergej N. Azbelev and Jurij Marčenko, eds. *Belomorskie and stariny i duhovnye stihi: Sobranie A. V. Markova* Saint-Petersburg: Dmitrij Bulanin.
- Edwards 1988 Mark W. Edwards. "Homer and Oral Tradition: The Formula, Part II." *Oral Tradition*, 3.1-2:11-60.
- Gorelov 2001a Aleksandr A. Gorelov, ed. *Byliny: v 25 tomah. Tom pervyj: byliny Pečory*. Saint Petersburg, Moscow: Nauka; Izdatel'skij centr «Klassika».
- Gorelov 2001b _____, ed. *v 25 tomah. Tom vtoroj: byliny Pečory*. Saint Petersburg; Moscow: Nauka; Izdatel' skij centr «Klassika».
- Gorelov 2003 _____, ed. *Byliny: v 25 tomah. Tom tretij: byliny Mezeni*. Saint Petersburg; Moscow: Nauka; Izdatel' skij centr «Klassika».
- Gorelov 2004 _____, ed. *Byliny: v 25 tomah. Tom četvertyj: byliny Mezeni*. Sankt-Peterburg; Moscow: Nauka; Izdatel' skij centr «Klassika».
- Lord 1960 Albert B. Lord. *The Singer of Tales*. Cambridge, MA: Harvard University Press.
- Magoun 1953 Francis P. Magoun. "Oral-formulaic character of Anglo-Saxon narrative poetry." *Speculum*, 28:446-67.
- Minton 1975 William W. Minton. "The frequency and structuring of traditional formulas in Hesiod's *Theogony*." *Harvard Studies in Classical Philology*, 79:25-54.
- Monroe 1972 James T. Monroe. "Oral composition in pre-Islamic poetry." *Journal of Arabic Literature*, 3:1-53.
- Selivanov 1988 Fëdor M. Selivanov, ed. *Byliny. Biblioteka russkogo fol'klora*. Moscow: Sovetskaja Rossija.
- Smirnov 1978 Jurij I. Smirnov, ed. 1978. *Novgorodskie byliny*. Moscow: Nauka.
- Smirnov and Smolickij 1974 Jurij I. Smirnov and Viktor G. Smolickij, eds. *Dobrynja Nikitič i Aleša Popovič*. Moscow: Nauka.

- Strasser 1984 Franz Xaver Strasser. *Zu den Iterata der frühgriechischen Epik*. Königstein: Hain.
- Uxov 1975 Pyotr D. Uxov. "Fixed Epithets in the Byliny as a Means of Creating and Typifying Images." In *The Study of Russian Folklore*. Ed. by Felix J. Oinas and Stephen Soudakoff. The Hague: Mouton. pp. 219-31.
- Vikis-Freibergs and Freibergs 1978a Vaira Vikis-Freibergs and Imants Freibergs. "Formulaic analysis of the computer-accessible corpus of Latvian sun songs." *Computers and the Humanities*, 12:329-39.
- Vikis-Freibergs and Freibergs 1978b Vaira Vikis-Freibergs and Imants Freibergs. "The Sun-songs of Latvian Folklore: A Computer-Accessible Corpus." *Journal of Baltic Studies*, 9:20-31.

This page is intentionally left blank.