Glottometrics 20, 2010, 70-109

Dynamics of Word Length in Sentence

Fengxiang Fan, Dalian
Peter Grzybek, Graz
Gabriel Altmann, Lüdenscheid

Abstract. In the present article the length of words in individual positions in sentence is being scrutinized. The conspicuous increase of word length with increasing position in short sentences very regularly slows down with increasing sentence length.

Keywords: *word length, sentence length, position in sentence, English, Hungarian, Indonesian, Latin, Russian, Slovak*

1. Introduction

Linguistic laws may arise in two ways: either they are derived directly from an existing theory, or they represent an isolated hypothesis to be tested. If accepted, the hypothesis must be systematized, i.e. incorporated into a system of similar statements. This systematization can consist in showing that either the given hypothesis is a consequence of higher level laws, theories, or axioms, or that it can itself be used to draw consequences about other phenomena; in the latter case, the hypothesis serves as a kind of axiom or as the starting-point of a new theory. If none of these two procedures is possible, we are concerned with a local phenomenon, i.e. with testing an inductive hypothesis on restricted data. If we are lucky and the hypothesis can be maintained, this does not mean that it can simply be declared to be a law.

In searching for linguistic laws, there is a tendency to jump to generalizations, not sufficiently taking into account important boundary conditions. That is, we tend to tacitly accept the *ceteris paribus* condition because our argumentation seems to be linguistically well substantiated. Fortunately, occasional rejections of a hypothesis in some languages need not weaken the status of the law-candidate – on the contrary, they force us to refine it by includeing some necessary boundary conditions. Natural laws, too, hold only if some boundary conditions are fulfilled, and even mathematical theorems begin with famous sentences like "Let be given…, then …" or "Let hold that …, then…". With regard to linguistic studies on word length, Grzybek, Kelih and Stadlober (2008) pointed out a number of boundary conditions whose consideration may lead to different parameters in models, or even to different functions: such conditions are, among others, data scarcity, data homogeneity, language type, intra- vs. inter-textual approach, etc.

The present study focuses on word length distributions according to word positions within a sentence and according to the length of the sentence: We concentrate on the question if word length is stable in the course of a sentence, from its beginning to its end, or if there is a particular change or development of word length, depending on the length of the sentence. Since we proceed inductively, the first step consists of scrutinizing texts in several languages in order to achieve more corroboration so as to reduce language-specific boundary conditions possibly coming into play.

Here we want to study a hypothesis concerning the interrelation between word length and its position in sentence. In this context, word length is determined in terms of the number of syllables per word, sentence length in terms of the number of words per sentence – though one should better measure it in terms of the number of clauses (or phrases), since these are the direct constituents of sentences. This means that we search for some specific form of self-regulation – if there is any – and for the boundary conditions under which it holds.

Most probably it was A. Niemikorpi (1991) who first observed that, in Finnish sentences, the longest words tend to occur at the end of a sentence, or a clause.[1] Niemikorpi's study was based on Finnish corpus material from the 1960s, including oral and written sources. From five basic text categories (non-fiction, press reportage, fiction, and informal standard language), which were further subdivided into 58 subcategories, 100 delimited, classified text samples consisting of five sentences or approximately 60 textual words were selected by random sampling. As a result, a corpus of ca. 430,000 running words was obtained and submitted to detailed analysis.

Niemikorpi's first calculation showed that, except for very short sentences, the average length of running words (measured in the number of graphemes per word) is greater at the end of a sentence than in the middle or at the beginning; it was also found that in a sentence the word in initial position is normally longer than the word that follows it.

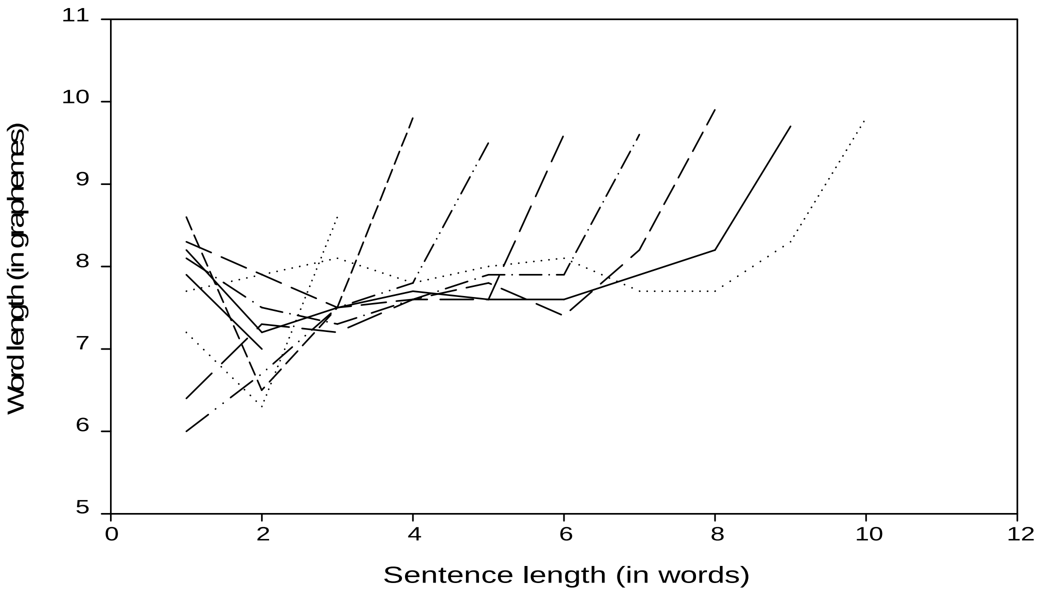

Figure 1. Average word length in different positions in sentences of different length (Niemikorpi 1991: 287)

Basing the same kind of calculation on clauses, not on sentences, Niemikorpi found that in this case, too, word length almost invariably increases towards the end. According to Niemikorpi, this regularity applies equally well to the length of running words in both main and subordinate clauses; Figure 3 shows the tendency for main clauses.

Given these observations, Niemikorpi suggested the English term "scoop law" to denote this phenomenon, considering this term as an equivalent for the Finnish term 'viskurilaki' an adaptation from the term "Wannmühlegesetz", previously used in Finnish folklor-

[1] Unfortunately, in Saukonen's (1994) summary of Niemikorpi's work, only the term 'clause' is used, ignoring Niermikorpi's detailed and comparative study of main and subordinate clauses as well as complete sentences.

istics, to describe the fact that in the verses of the Finnish folk epic Kaleva, the heaviest elements tend do occur at the end of stanzas.[2]

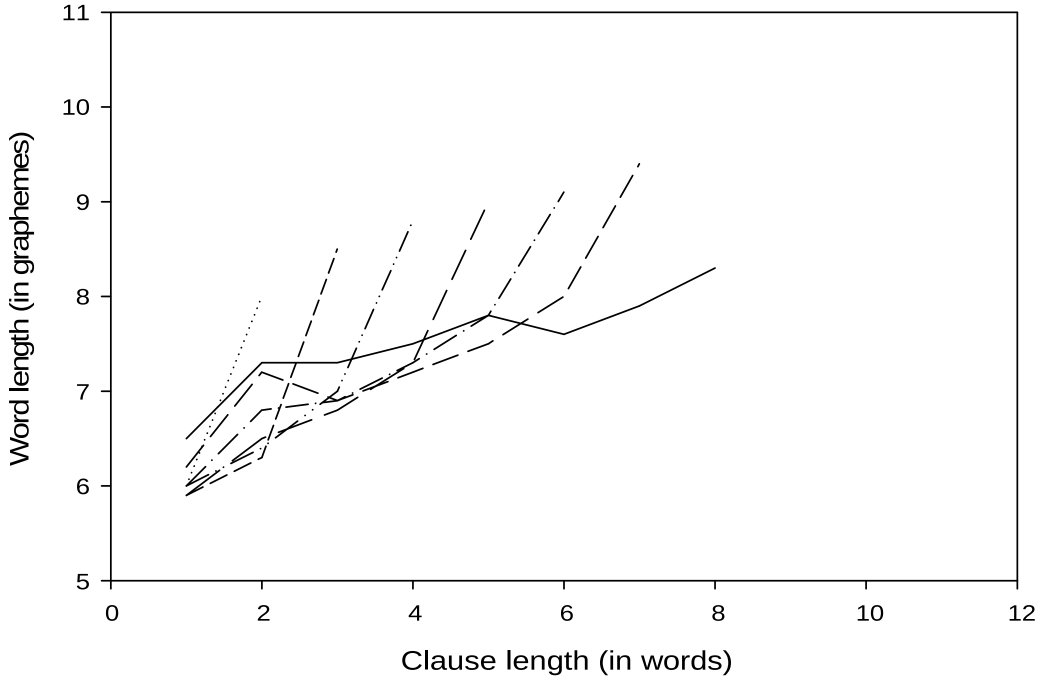

Figure 2. Average word length in different positions in clauses of different length (Niemikorpi 1991: 287)

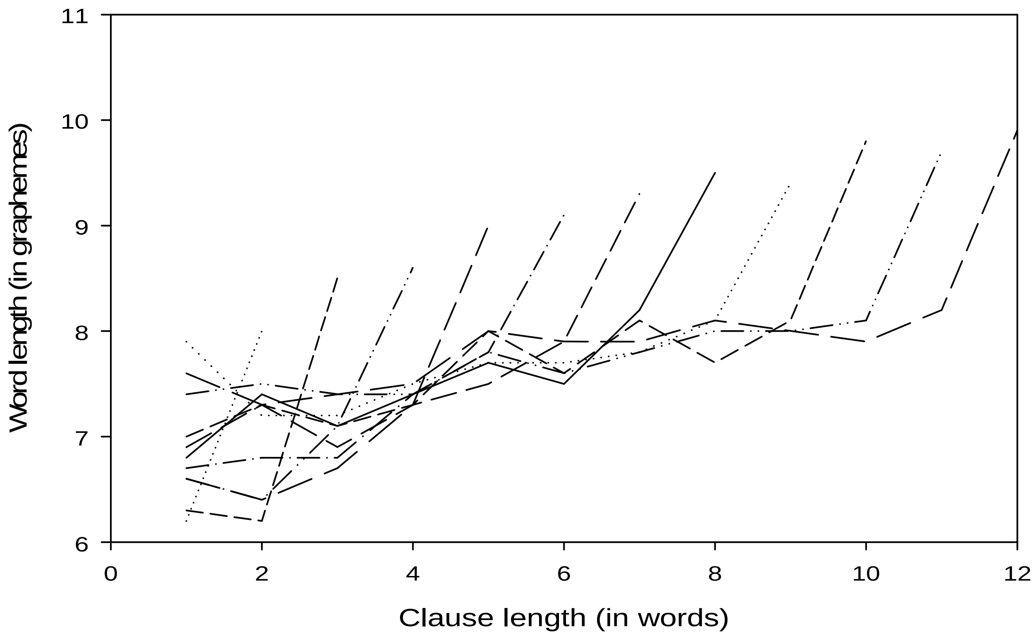

Figure 3. Average word length in different positions in main clauses of different length (Niemikorpi 1991: 292)

[2] By way of a general explanation, Niemikorpi referred to the fact that in case of Finnish, we are concerned with an SVO language, in which the subject precedes the predicate and the verb qualifiers are placed at the end. Also, the theme-rheme structure of clauses is likely to play a crucial role here.

There is, however, one major drawback with Niemikorpi's study: measuring word length in the number of graphemes per word may heavily skew the results obtained, or even lead to erroneous results. The reason for this is that graphemes (even less than phones or phonemes) are no direct constituents of the word; rather, units like syllables or morphemes should be taken as measuring units. Last not least, the necessity to proceed this way is related to the Menzerathian fact that long words tend to be composed of short syllables, and short words of long syllables. A "long" word measured in terms of graphemes, however, may consist either of many short syllables, or of a few long syllables – as a result, on the basis of Niemikorpi's study we cannot be sure about word length neither in general, nor with regard to sentence position, specifically.

In a subsequent study, Uhlířová (1997a,b) has analyzed the relation between word length and position in sentence with Czech material, in order to study a language typologically different from Finnish. For this purpose, she analysed ten short stories for schoolchildren by Czech writer Miloš Macourek (1926-2002), summing up to a corpus of ca. 10,000 running words. In contrast to Niemikorpi, Uhlířová did not analyze word length in clauses, but in sentences only[3], confining the presentation of results to the interval from 1 to 12 words per sentence, since longer sentences were considered to occur too rarely in the corpus to provide reliable results. Moreover, in her study word length was measured in the number of syllables per word.

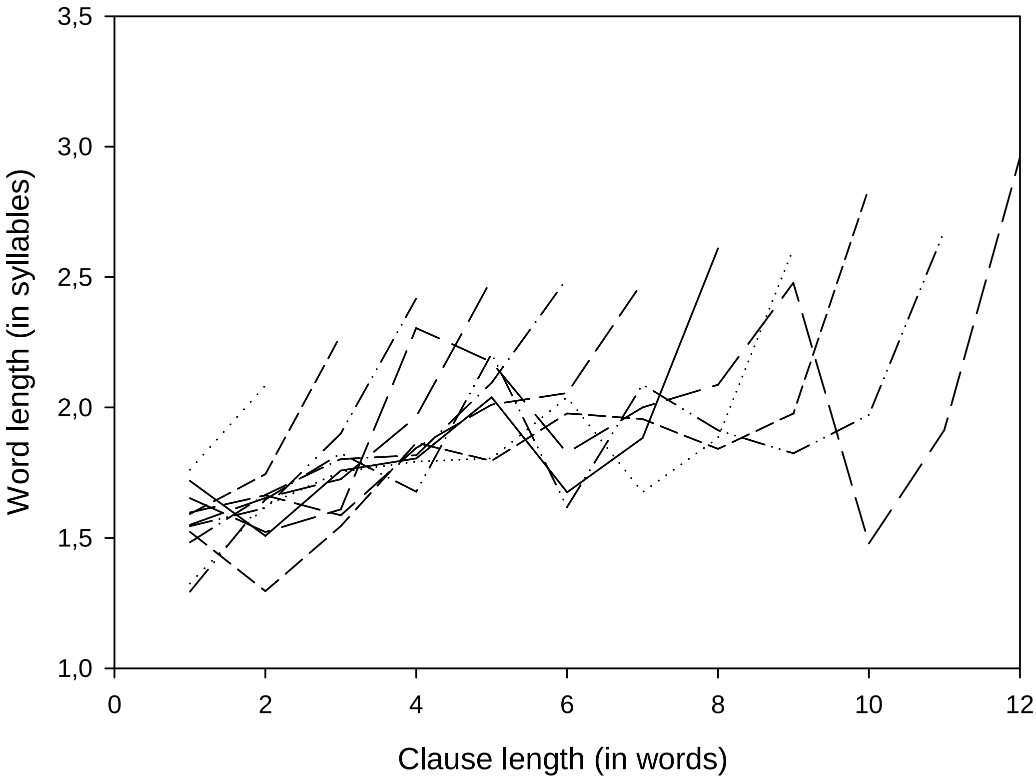

Figure 4. Average word length in sentences of different length in a corpus of Czech short stories (Uhlířová 1997a: 182, 1997b: 269)

[3] Unfortunately, the original term 'věta' (= sentence) used in Uhlířová's (1997a) original Czech text, has been translated as 'clause' in Uhlířová's (1997b) English version, thus given rise to serious misinterpretations.

According to Uhlířová, the two most salient points are those at the beginning of a sentence (i.e. the first position), and those at the end (i.e. the last position): the first position is the position of short words, whereas the end is the position of long words. With the exception of sentence length 1, the average word length is lowest in the first position with sentences of any length, word length then gradually increasing from position to position, at least up to sentence length 8, when occurrences become too rare to represent reliable results. Uhlířová (1997b: 269) suspected that the quantitative relationship between length and order of words in sentences need not be language specific, rather it may be of a more general nature, lending itself to be interpreted as a "quantitative manifestation of the structural principle of end weight".

In fact, a closer inspection of Figures 1 to 4 gives rise to the impression that not only the last, but also the first word in a sentence seem to play a specific role, as far as their length is concerned (at least on the average). We have no idea thus far, to what kind of linguistic factors this tendency might be related, in how far language specifics, discourse or sentence typological aspects, or other factors come into play, etc. Without a doubt, a more detailed study of this question is an obligatory task for future research, which cannot be tackled in the framework of this article.

Summarizing, one can say that, both for Finnish and Czech, similar tendencies seem to have been observed, in so far as words at the end of clauses and/or sentences tend to be longer. Such a tendency would seem to be reasonable, if one takes into account the fact that words occurring towards the end of an utterance carry more information than words at its beginning; hence these words do not occur frequently and, as a consequence of the well-known Zipfian rule, tend to be longer. As a result, we would be concerned with some specific kind of regulation between the position, information, length and frequency of words.

Given the situation outlined above, a number of open issues to be solved remain open. The first question is, if the above tendency holds true more generally for other languages, as well. The second question concerns the point, in how far the observed tendency characterizes clauses and sentences in the same way. On the one hand, it may be, theoretically speaking, possible that the above observations with regard to sentences are but a special case (i.e. a consequence) of tendencies ultimately characterizing clauses, since any sentence is, of course, composed of one or more clauses. On the other hand, this would mean that the course of a sentence composed of more than one clause would display some specific kind of oscillation, or steps, around clause boundaries, a phenomenon which has never before been reported about.

It goes without saying that, in the present study, we will not be able to give answers to all of these questions. We will therefore concentrate on sentences only, extending the empirical data base to more languages. In addition to this, in case the results will corroborate the findings reported above, we will attempt to model the overall trend common to all languages studied. Of course, whatever turns out to be the case, is likely to hold on the average only; further generalizations will automatically give rise to further hypotheses.

Linguistically, this hypothesis is reasonable: the "topic" introduced at the beginning of the sentence obtains "comments" or "predicates" whose length increases with increasing sentence. As has been said above, this phenomenon has been studied for Finnish (Niemikorpi 1991) and for Czech (Uhlířová 1997a,b) with regard to whole sentences. It should be mentioned that these two languages are both strongly synthetic. But what is the situation in languages tending towards analytism like English? How is the situation in languages like Japanese, having many long words, or in some American Indian languages, in which whole sentences can be formed of one word only? Do they conform to one of the alternatives of our hypothesis? And if not, how do we have to modify the hypothesis? Almost as a matter of fact, a second question emerges: If there is a trend, how can it be characterized: is it linear or not?

And, if it is linear, does the slope of the straight line increase with increasing sentence length? And if it is not linear, what form does it then have?

Taking into account that there are a number of uncertainty factors arising from previous research, the only way to get some security is to proceed inductively, by way of analyzing several languages and then elaborating on the results obtained. The languages chosen to test our hypothesis quite logically included languages with different typological profiles: Russian, Slovak, Latin are strongly inflecting, Hungarian is strongly agglutinative, English tends to analytism, Indonesian is moderately agglutinative.

For each of these languages, we shall use texts of different origin and different length, and we will study the progression of averages within sentences of a given length. In detail, the following hypotheses can be tested in isolation:

Hypothesis I:
Mean word length increases from the beginning to the end of a sentence

against the null hypothesis that no change of word length occurs. Given hypothesis (I) can be corroborated, we will then test the subsequent hypothesis (II):

Hypothesis II:
Word length increase from the beginning to the end of a sentence follows a common principle.

2. Testing

The hypothesis is very simple; it may, in fact, be scrutinized with longer individual texts, or a group of texts of the same sort (thus guaranteeing data homogeneity), maybe with non-homogeneous samples of texts from one and the same language. For each analysis performed in this study (see Appendix for the sources), groups of sentences with identical sentence length were formed, and for each individual length group, mean word length in the individual positions was calculated. Possible specifics of the length of first and last words in sentences (see above) are going to be ignored, a sentence rather being taken "as a whole", including first and last word.

In our first analyses it turned out quite soon that for short individual texts there seems to be no clearly overt tendency; this is most probably caused by the small number of sentences of equal length, displaying a vast variance and therefore rendering the count of averages non-reliable. In these cases – cf. the analyses of Russian and Slovak journalistic texts, and English below –, all individual texts were merged into a corpus of texts, and sentences of equal length from the whole corpus were pooled, in order to increase reliability; mean lengths in the given word position was then computed group-wise.

The results for all languages are presented in Tables 1 to 6: Column 1 contains the sentence length in terms of the number of words per sentence; Column 2 contains the mean word-lengths in a given position within a sentence; Column 3 shows the linear regression equation for word length depending on position; Column 4 (R^2) shows the determination coefficient of the regression; Column 5 contains the mean word length in the whole sentence; Column 6 (k) presents the number of sentences of length n in the given sample.

Still, even when merging individual texts to a larger corpus, the problem of data reliability remains, a problem which has not been solved even in statistics. What is a reliable sample size? In the present investigation we adhere to the following principles:

a. The number of sentences of a given length must be at least $k \geq 10$.

b. Sentences with less than 3 words are omitted from the analyses, and only sentences of length $n \geq 3$ are scrutinized, because lengths 1 and 2 do not make sense for regression analyses.
c. If sentences with length $n = 3,4,\ldots$ are represented by less than $k \geq 10$ sentences they, too, are excluded from analysis which begins with the “well represented” lengths.
d. At the upper end of length as well, sentences are scrutinized only up to the length n having minimal frequency of $k \geq 10$; more exactly, all those occurrences (possibly including such with less than 10) are taken account of up to that point, where there is no more occurrence of $k \geq 10$ in the whole sample.

2.1. Russian

As was mentioned above, no reliable results were obtained for relatively short individual journalistic texts; therefore data were pooled, and a corpus of 44 Russian journalistic texts, consisting of 860 sentences with 15001 words, was obtained and submitted to analysis. The results are presented in Table 1. The length n = 3, 4 are not well represented, with only 3 or 5 occurrences only, and are omitted from analysis. At length n = 31, there seems to be a break after which oscillation begins. As to an explanation of this break, it is possible that, in addition to statistical reasons – the data points above sentence length 30 not being represented by sufficient observations – linguistic factors come into play (cf. Grzybek et al. 2008); this problem will have to be studied in future with more voluminous data.

With regard to the regression analyses carried out in the present study, it should be mentioned that we are not so much interested in the goodness of the fitting results for the individual regression of each sentence length; rather, we are concerned with the tendency along sentences of different length. For the sake of simplicity we use a linear regression, even if in many cases different functions would be more appropriate to express the positional changes. Again, a thorough study as to which function(s) might be more appropriate, must be left to future research.

Table 1
Mean lengths (L) of words in individual positions (x) in sentences of length n in 44 Russian journalistic texts from the journal “Vremja” (M = mean of all positions, k = number of sentences)

n	**Mean syllabic lengths of words *L***	**Linear relation *L = a + bx***	R^2	***M***	***k***
3	3.33, 1.87, 3.67	-	-	-	3
4	3.00, 2.80, 3.20, 2.40	-	-	-	5
5	2.75, 2.75, 2.92, 2.92, 3.00	2.6670 + 0.0670x	0.89	2.87	12
6	2.07, 2.72, 2.72, 2.69, 2.48, 3.07	2.2000 + 0.1214x	0.47	2.63	29
7	2.33, 2.40, 3.13, 3.13, 2.33, 3.10, 3.20	2.3443 + 0.1146x	0.34	2.80	30
8	2.26, 2.67, 2.67, 2.81, 2.56, 2.67, 2.63, 3.22	2.3504 + 0.0746x	0.47	2.69	27
9	2.65, 2.76, 2.81, 2.86, 3.14, 2.62, 2.62, 2.54, 3.11	2.7433 + 0.0093x	0.01	2.79	37
10	2.43, 2.64, 2.75, 2.77, 2.48, 2.77, 2.89, 2.77, 2.75, 3.36	2.4313 + 0.0599x	0.51	2.76	44
11	2.66, 2.90, 2.59, 3.17, 2.85, 2.71, 2.54,	2.7311 + 0.0116x	0.03	2.80	41

	2.66, 2.88, 2.61, 3.24				
12	2.40, 2.58, 2.37, 2.35, 2.40, 3.12, 2.86, 2.81, 2.56, 2.65, 2.98, 2.93	2.3627 + 0.0469x	0.40	2.67	43
13	2.22, 2.54, 3.02, 2.38, 2.82, 2.68, 2.74, 2.66, 2.66, 2.58, 2.76, 2.86, 3.24	2.4377 + 0.0381x	0.32	2.70	50
14	2.34, 2.83, 2.71, 3.12, 2.78, 3.15, 3.05, 2.61, 2.83, 2.71, 2.83, 2.63, 2.54, 3.02	2.7775 + 0.0025x	0.00	2.80	41
15	2.37, 2.61, 2.50, 2.76, 2.97, 2.63, 3.03, 3.11, 2.82, 2.45, 2.76, 3.08, 2.47, 2.89, 2.97	2.5953 + 0.0208x	0.14	2.76	38
16	2.31, 2.28, 2.90, 2.69, 2.90, 2.48, 3.00, 2.76, 3.03, 2.93, 3.00, 2.55, 2.90, 3.00, 3.38, 3.21	2.4450 + 0.0456x	0.51	2.83	29
17	2.57, 2.84, 2.68, 2.50, 2.98, 2.59, 2.86, 2.77, 2.57, 2.68, 2.36, 2.91, 2.75, 2.48, 2.80, 2.91, 3.18	2.6222 + 0.0121x	0.09	2.73	44
18	2.28, 2.69, 2.67, 2.82, 2.67, 3.05, 2.59, 2.77, 2.74, 3.08, 2.54, 2.44, 2.62, 2.69, 2.74, 2.79, 2.26, 2.95	2.6718 + 0.0017x	0.00	2.69	39
19	2.54, 2.80, 2.37, 2.93, 2.32, 2.37, 3.05, 2.71, 2.63, 2.63, 2.37, 2.93, 2.78, 3.10, 2.93, 2.71, 2.39, 2.78, 3.17	2.5414 + 0.0170x	0.13	2.71	41
20	2.47, 2.74, 2.91, 2.65, 3.32, 3.12, 3.12, 2.85, 2.62, 2.35, 2.56, 2.59, 2.85, 2.82, 2.79, 2.71, 2.82, 2.85, 2.65, 3.26	2.7778 + 0.0024x	0.00	2.80	34
21	2.30, 2.81, 2.85, 3.11, 2.48, 3.00, 2.74, 2.96, 2.63, 2.48, 3.11, 3.04, 2.33, 2.59, 2.59, 2.33, 2.63, 2.63, 2.15, 2.56, 3.07	2.7954-0.0100x	0.05	2.69	27
22	2.20, 2.80, 2.30, 3.17, 2.97, 2.83, 2.77, 2.87, 2.43, 2.77, 2.63, 2.23, 3.40, 2.80, 2.77, 2.70, 3.03, 2.40, 2.73, 2.50, 3.57, 2.90	2.6134 + 0.0129x	0.06	2.76	30
23	2.03, 2.47, 2.69, 2.56, 2.56, 2.38, 2.81, 2.88, 2.91, 2.88, 2.88, 2.84, 3.13, 2.47, 2.56, 2.84, 2.59, 2.94, 2.72, 2.59, 2.69, 2.91, 3.34,	2.4799 + 0.0204x	0.26	2.72	32
24	2.32, 2.64, 2.64, 2.77, 3.00, 2.86, 3.41, 2.36, 2.91, 2.77, 3.36, 3.05, 3.09, 2.50, 2.77, 2.77, 2.36, 2.95, 2.82, 2.50, 2.50, 2.95, 2.77, 3.41	2.7437 + 0.0054x	0.02	2.81	22
25	2.24, 2.52, 2.76, 2.36, 2.64, 2.36, 2.28, 2.88, 2.36, 2.80, 2.64, 2.68, 2.84, 2.40, 2.32, 2.60, 2.76, 2.32, 2.80, 2.68, 2.56, 2.64, 2.36, 3.04, 2.96	2.4432 + 0.0114x	0.13	2.59	25
26	2.67, 2.87, 2.53, 3.07, 2.73, 2.67, 2.47, 2.87, 3.27, 3.20, 2.20, 2.87, 3.07, 2.47, 3.13, 3.20, 3.33, 2.53, 3.27, 2.73, 3.07, 2.60, 2.93, 2.73, 2.80, 3.13	2.7584 + 0.0077x	0.04	2.86	15
27	2.94, 2.22, 2.72, 2.94, 2.94, 3.33, 3.44,	2.8913 - 0.0123x	0.09	2.72	18

	2.94, 2.72, 2.33, 2.22, 2.78, 3.00, 2.72, 2.72, 2.33, 2.89, 3.06, 2.72, 2.72, 2.33, 2.72, 2.17, 2.61, 2.56, 2.44, 2.89				
28	2.06, 2.65, 2.41, 2.59, 2.94, 2.65, 2.76, 2.82, 2.94, 2.71, 2.47, 2.29, 2.41, 2.76, 2.35, 2.53, 2.82, 3.12, 3.24, 2.12, 1.88, 2.18, 2.94, 2.47, 3.00, 2.88, 3.24, 3.18	2.4893 + 0.0116x	0.07	2.66	17
29	2.83, 2.11, 3.00, 2.83, 3.11, 3.00, 2.56, 3.28, 2.78, 2.89, 2.50, 2.50, 2.72, 2.44, 1.94, 2.67, 3.00, 2.89, 2.11, 2.17, 2.50, 3.11, 2.39, 3.39, 2.00, 3.17, 2.50, 2.33, 3.22	2.7538 - 0.0044x	0.01	2.69	18
30	2.22, 2.33, 3.11, 2.56, 3.11, 3.33, 2.44, 3.44, 3.56, 2.44, 2.44, 2.89, 2.67, 2.67, 2.78, 3.22, 3.56, 1.78, 2.89, 2.33, 3.00, 2.67, 2.89, 3.33, 3.11, 3.33, 3.56, 2.22, 1.67, 2.56	2.8300 - 0.0017x	0.00	2.80	9
31	1.80, 2.50, 2.70, 3.20, 2.50, 3.20, 3.00, 2.40, 2.80, 2.90, 1.80, 2.90, 3.10, 2.40, 2.90, 2.30, 2.80, 2.40, 2.30, 3.40, 2.30, 3.00, 2.00, 2.50, 2.90, 2.20, 1.60, 2.20, 2.00, 2.70, 2.50	2.7535 - 0.0124x	0.06	2.55	10

Considering the results presented in Table 1a, we may draw the following conclusions: Parameter *b* (cf. Column 3), indicating the positional change (increase or decrease) of word length, takes positive values with shorter sentences; this means that there is an increase of word length from the beginning of a sentence towards its end. To show the strength of the relation, we use the determination coefficient (R^2).[4] However, the increase gets slower as sentence length increases; this can be seen in Figure 5: at length $n = 22$, the values for *b* may become negative, and from this point on, *b* oscillates between positive and negative values, i.e. the observed tendency ceases to exist. This means that when a sentence attains a certain number of words, word length may even decrease towards the end of the sentence.

The trend is illustrated in Figure 5; with regard to the analyses of other texts and languages to follow below, the *x*- and *y*-axes are chosen in intervals identical for all Figures. In the given case, the trend can be captured by a straight line $b = 0.0859 - 0.0034n$ (in the interval $5 \geq n \leq 31$); a slightly better fitting can be attained by the function $b = 0.2349 \exp(-0.1616n)$, which will be used in the subsequent analysis, here yielding $R^2 = 0.69$.

Thus, as a result, hypothesis (I) is only partly corroborated: it holds for short sentences, but for long sentences (with the addition of subordinate clauses), the tendency disappears: towards the end of long sentences, average word length may cease to increase (i.e. increase may approximate zero).

It is possible of course that this conclusion will have to be modified by future research: after all, a long sentence consists of a number of clauses, and the addition of clause-specific rules may completely change the image.

[4] The values of the *t*- and *F*-tests were always satisfactory but mostly equal because of the software limitations

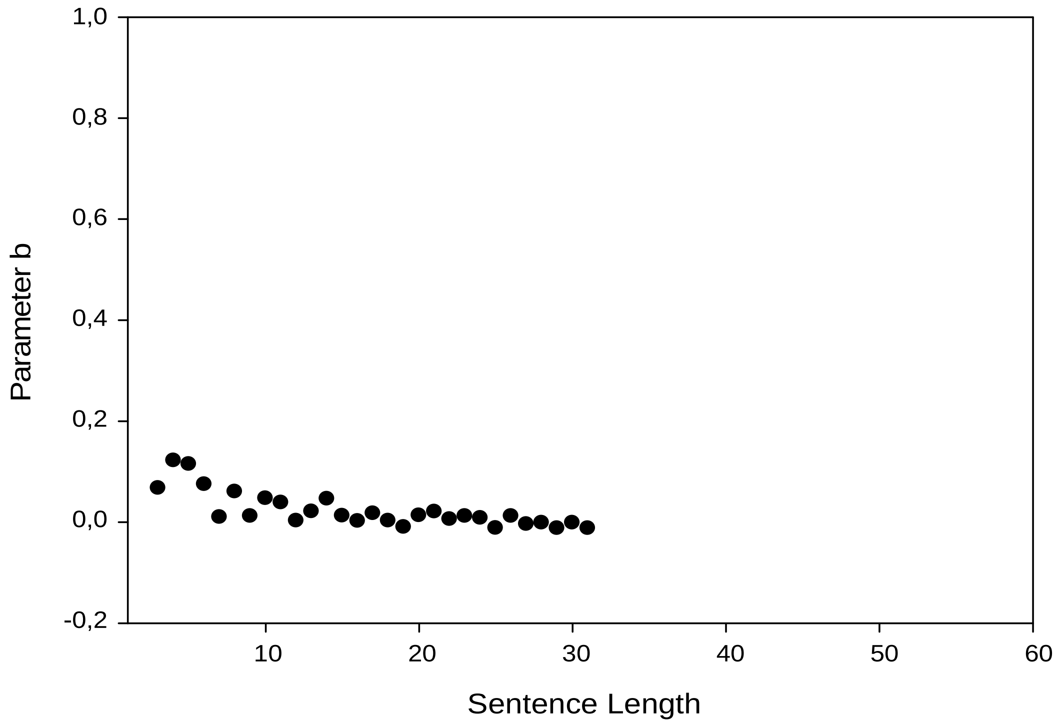

Figure 5. Decrease of coefficient *b* in a corpus of 44 Russian journalistic texts

In order to avoid the bias that we are concerned with a corpus-biased effect (due to data heterogeneity), it seems reasonable to test the hypothesis (and our above conclusion) on broader data material.

For this purpose, we now analyze a larger Russian text, namely L.N. Tolstoj's *Anna Karenina* in its complete length (with 19297 sentences and 258384 running words), in order to obtain a reliable number of sentences of all lengths from a single text (which still may be, of course, be composed of heterogeneous elements). This novel consists of 19297 sentences, so the data basis should be sufficient in case our hypothesis holds. The results are presented in Table 2. Since in *Anna Karenina*, more sentence lengths are (reliably) represented, we can take the interval $3 \leq n \leq 58$.

Table 2
Mean lengths (L) of words in individual positions (x) in sentences of length n in L.N. Tolstoj's *Anna Karenina* (M = mean of all positions, k = number of sentences)

n	**Mean syllabic lengths of words *L***	**Linear relation *L = a + bx***	***R*²**	***M***	***k***
3	1.64, 2.02, 2.48	1.2067 + 0.4200x	1.00	2.05	1093
4	1.63, 1.98, 2.06, 2.47	1.3850 + 0.2600x	0.95	2.04	1296
5	1.65, 2.06, 2.12, 2.17, 2.52	1.5490 + 0.1850x	0.89	2.10	1294
6	1.58, 2.01, 2.06, 2.14, 2.21, 2.60	1.5220 + 0.1651x	0.88	2.10	1145
7	1.66, 2.07, 2.05, 2.13, 2.18, 2.27, 2.59	1.6614 + 0.1186x	0.85	2.14	1115
8	1.68, 2.10, 2.10, 2.11, 2.18, 2.27, 2.22, 2.59	1.7518 + 0.0899x	0.77	2.16	992

9	1.73, 2.15, 2.10, 2.13, 2.20, 2.23, 2.24, 2.24, 2.65	1.8247 + 0.0722x	0.71	2.19	975
10	1.68, 2.10, 2.11, 2.17, 2.11, 2.15, 2.15, 2.22, 2.31, 2.67	1.8033 + 0.0661x	0.69	2.17	874
11	1.69, 2.08, 2.15, 2.21, 2.16, 2.19, 2.23, 2.13, 2.23 , 2.24, 2.75	1.8551 + 0.0554x	0.57	2.19	840
12	1.68, 2.14, 2.09, 2.22, 2.24, 2.31, 2.27, 2.15, 2.32, 2.22, 2.31, 2.67	1.9111 + 0.0473x	0.58	2.22	736
13	1.75, 2.15, 2.10, 2.14, 2.21, 2.25, 2.25, 2.29, 2.20, 2.25, 2.26, 2.37, 2.66	1.9312 + 0.0415x	0.66	2.22	648
14	1.76, 2.15, 2.17, 2.21, 2.21, 2.17, 2.26, 2.27, 2.25, 2.17, 2.29, 2.25, 2.34, 2.66	1.9765 + 0.0332x	0.57	2.23	643
15	1.76, 2.19, 2.20, 2.16, 2.18, 2.23, 2.24, 2.30, 2.35, 2.21, 2.23, 2.21, 2.28, 2.28, 2.66	2.0131 + 0.0274x	0.48	2.23	598
16	1.87, 2.24, 2.14, 2.27, 2.18, 2.18, 2.23, 2.22, 2.21, 2.17, 2.28, 2.26, 2.21, 2.25, 2.29, 2.71	2.0470 + 0.0218x	0.42	2.23	534
17	1.78, 2.10, 2.14, 2.18, 2.10, 2.25, 2.25, 2.37, 2.23, 2.29, 2.19, 2.26, 2.30, 2.13, 2.29, 2.31, 2.70	2.0046 + 0.0248x	0.48	2.23	489
18	1.77, 2.13, 2.16, 2.26, 2.25, 2.24, 2.22, 2.21, 2.28, 2.22, 2.23, 2.26, 2.23, 2.18, 2.26, 2.22, 2.27, 2.74	2.0444 + 0.0195x	0.37	2.23	489
19	1.86, 2.21, 2.32, 2.21, 2.09, 2.27, 2.31, 2.34, 2.25, 2.06, 2.28, 2.17, 2.28, 2.19, 2.24, 2.28, 2.37, 2.44, 2.75	2.0721 + 0.0187x	0.36	2.26	373
20	1.87, 2.23, 2.10, 2.15, 2.25, 2.17, 2.18, 2.36, 2.34, 2.20, 2.18, 2.18, 2.26, 2.18, 2.24, 2.27, 2.20, 2.17, 2.25, 2.81	2.0736 + 0.0148x	0.27	2.23	362
21	1.89, 2.27, 2.12, 2.27, 2.34, 2.23, 2.26, 2.18, 2.18, 2.20, 2.38, 2.33, 2.17, 2.31, 2.35, 2.27, 2.18, 2.28, 2.34, 2.22, 2.64	2.1293 + 0.0117x	0.28	2.26	320
22	1.82, 2.19, 2.19, 2.22, 2.20, 2.18, 2.19, 2.26, 2.33, 2.40, 2.25, 2.32, 2.18, 2.25, 2.36, 2.18, 2.27, 2.22, 2.31, 2.25, 2.52, 2.68	2.0832 + 0.0156x	0.41	2.26	330
23	1.75, 2.14, 2.18, 2.21, 2.27, 2.28, 2.35, 2.21, 2.32, 2.39, 2.30, 2.24, 2.35, 2.42, 2.37, 2.19, 2.21, 2.32, 2.30, 2.19, 2.24, 2.42, 2.75,	2.1109 + 0.0139x	0.31	2.28	271
24	1.90, 2.27, 2.20, 2.21, 2.32, 2.32, 2.22, 2.35, 2.33, 2.32, 2.39, 2.28, 2.32, 2.21, 2.09, 2.20, 2.11, 2.30,	2.1860 + 0.0060x	0.09	2.26	244

	2.20, 2.22, 2.32, 2.16, 2.34, 2.69				
25	1.94, 2.24, 2.17, 2.19, 2.26, 2.32, 2.37, 2.20, 2.30, 2.36, 2.27, 2.20, 2.21, 2.30, 2.24, 2.14, 2.04, 2.24, 2.20, 2.26, 2.23, 2.32, 2.36, 2.45, 2.72	2.1452 + 0.0089x	0.21	2.26	230
26	1.86, 2.06, 2.25, 2.23, 2.32, 2.16, 2.21, 2.29, 2.20, 2.31, 2.36, 2.36, 2.21, 2.16, 2.29, 2.23, 2.46, 2.17, 2.40, 2.24, 2.24, 2.30, 2.16, 2.25, 2.30, 2.71	2.1307 + 0.0095x	0.24	2.26	199
27	1.92, 2.22, 2.08, 2.19, 2.29, 2.14, 2.18, 2.12, 2.20, 2.29, 2.24, 2.17, 2.28, 2.26, 2.34, 2.32, 2.11, 2.25, 2.39, 2.19, 2.26, 2.31, 2.24, 2.48, 2.33, 2.27, 2.71	2.0873 + 0.0117x	0.43	2.25	194
28	1.88, 2.19, 2.06, 2.30, 2.13, 2.28, 2.22, 2.08, 2.12, 2.29, 2.26, 2.30, 2.06, 2.07, 2.30, 2.36, 2.18, 2.17, 2.24, 2.12, 2.13, 2.13, 2.19, 2.29, 2.08, 2.43, 2.26, 2.84	2.0866 + 0.0087x	0.18	2.21	189
29	1.94, 2.13, 2.43, 2.35, 2.21, 2.09, 2.26, 2.09, 2.26, 2.18, 2.14, 2.42, 2.58, 2.18, 2.22, 2.42, 2.35, 2.31, 2.21, 2.40, 2.27, 2.30, 2.25, 2.28, 2.12, 2.19, 2.27, 2.36, 2.67	2.1785 + 0.0062x	0.12	2.27	159
30	1.83, 2.29, 2.22, 2.27, 2.21, 2.23, 2.08, 2.17, 2.31, 2.29, 2.20, 2.12, 2.31, 2.37, 2.25, 2.23, 2.30, 2.37, 2.37, 2.32, 2.32, 2.23, 2.17, 2.12, 2.21, 2.38, 2.31, 2.20, 2.57, 2.61	2.1356 + 0.0082x	0.26	2.26	145
31	1.70, 2.12, 2.17, 2.16, 2.26, 2.21, 2.33, 2.23, 2.12, 2.21, 2.31, 2.40, 2.19, 2.28, 2.25, 2.48, 2.55, 2.23, 2.28, 2.20, 2.23, 2.18, 2.31, 2.19, 2.34, 2.26, 2.38, 2.32, 2.35, 2.44, 2.86	2.0939 + 0.0114x	0.33	2.28	130
32	2.03, 2.19, 2.12, 2.37, 2.09, 2.24, 2.12, 2.19, 2.24, 2.18, 2.08, 2.31, 2.43, 2.30, 2.30, 2.21, 2.45, 2.35, 2.26, 2.39, 2.15, 2.39, 2.27, 2.30, 2.04, 2.27, 2.07, 2.07, 2.33, 2.24, 2.19, 2.83	2.1676 + 0.0050x	0.09	2.25	115
33	1.88, 2.09, 2.07, 2.24, 2.41, 2.02, 2.25, 2.25, 2.18, 2.09, 2.21, 2.35, 2.04, 2.32, 2.12, 2.10, 2.25, 2.27, 2.36, 2.34, 2.51, 2.05, 2.39, 2.50, 2.31, 2.32, 2.12, 2.16, 2.35, 2.57, 2.42, 2.33, 2.62	2.0819 + 0.0103x	0.34	2.26	106

34	2.11, 2.13, 2.29, 2.25, 2.26, 2.38, 2.41, 2.30, 2.37, 2.35, 2.24, 2.42, 2.31, 2.31, 2.31, 2.41, 2.48, 2.31, 2.42, 2.29, 2.14, 2.13, 2.32, 2.47, 2.48, 2.45, 2.35, 2.26, 2.34, 2.12, 2.24, 2.23, 2.46, 2.81	2.2561 + 0.0041x	0.09	2.33	91
35	1.95, 2.15, 2.16, 2.28, 2.17, 2.48, 2.44, 2.40, 2.37, 2.33, 2.31, 2.28, 2.33, 2.32, 2.48, 2.32, 2.10, 2.43, 2.12, 2.17, 2.38, 2.15, 2.25, 2.25, 2.12, 2.07, 2.11, 2.16, 2.07, 2.41, 2.44, 2.10, 2.26, 2.58, 2.79	2.2339 + 0.0025x	0.02	2.28	81
36	1.99, 2.23, 2.26, 2.27, 2.21, 2.45, 2.47, 2.35, 2.13, 2.34, 2.30, 2.23, 1.99, 2.16, 2.15, 2.20, 2.37, 2.10, 2.41, 2.15, 2.26, 2.38, 2.12, 2.08, 2.12, 2.30, 2.38, 2.23, 2.37, 2.19, 2.29, 2.35, 2.17, 2.08, 2.29, 2.53	2.2246 + 0.0012x	0.01	2.25	86
37	1.92, 2.28, 2.10, 2.21, 2.12, 2.32, 2.06, 2.37, 2.22, 2.21, 2.32, 2.06, 2.13, 2.17, 2.19, 2.14, 2.51, 2.17, 2.29, 2.18, 2.29, 2.42, 2.15, 1.90, 2.37, 2.01, 2.35, 2.14, 2.03, 2.22, 2.31, 2.37, 2.37, 2.32, 2.21, 2.55, 2.77	2.1171 + 0.0063x	0.16	2.24	78
38	1.78, 2.35, 2.40, 2.10, 2.47, 2.19, 2.36, 2.29, 2.18, 2.47, 2.14, 2.25, 2.19, 2.13, 2.29, 2.29, 2.11, 2.01, 2.32, 2.38, 2.25, 1.92, 2.17, 2.14, 2.24, 2.04, 2.32, 2.29, 2.22, 2.00, 2.38, 2.58, 2.11, 2.28, 2.40, 2.24, 2.13, 2.63	2.1960 + 0.0021x	0.02	2.24	72
39	1.89, 2.07, 1.85, 1.93, 2.25, 2.02, 2.11, 2.42, 2.22, 1.96, 2.15, 2.33, 2.27, 2.07, 2.36, 2.36, 2.40, 2.38, 2.31, 2.24, 1.98, 2.02, 2.40, 1.98, 2.07, 2.29, 2.33, 2.04, 2.16, 2.00, 2.11, 2.09, 2.31, 2.05, 2.45, 2.24, 2.47, 2.29, 2.67	2.0677 + 0.0063x	0.15	2.19	55
40	1.72, 2.17, 2.13, 2.15, 2.06, 2.34, 1.96, 2.34, 2.15, 2.06, 2.38, 2.08, 2.19, 2.40, 2.15, 2.30, 2.42, 2.58, 2.38, 2.38, 2.23, 2.17, 2.38, 2.02, 2.15, 2.38, 2.36, 2.42, 2.21, 2.53, 2.62, 1.92, 2.21, 2.28, 2.19, 2.19, 1.91, 2.42, 2.47, 2.85	2.1148 + 0.0069x	0.14	2.26	53
41	2.06, 2.37, 2.00, 2.31, 2.10, 2.33, 1.96, 2.24, 1.96, 2.45, 2.02, 2.16, 2.41, 2.53, 2.18, 2.06, 2.47, 2.29,	2.1536 + 0.0038x	0.06	2.23	49

	2.04, 2.08, 2.37, 2.20, 2.04, 2.39, 2.45, 1.98, 2.31, 2.16, 2.08, 2.39, 2.33, 2.10, 2.24, 2.00, 2.20, 2.43, 2.39, 2.18, 2.31, 2.27, 2.76				
42	1.80, 2.30, 1.94, 2.18, 2.78, 2.40, 2.64, 2.38, 2.50, 2.06, 2.14, 2.44, 2.20, 2.36, 2.44, 1.88, 2.32, 2.26, 2.10, 2.08, 2.28, 2.52, 2.36, 2.34, 2.34, 2.34, 2.10, 2.12, 2.32, 2.38, 2.74, 2.36, 2.48, 2.60, 2.08, 2.48, 2.14, 2.18, 2.16, 2.64, 2.52, 2.96	2.2081 + 0.0054x	0.08	2.32	50
43	1.85, 2.46, 2.03, 2.49, 2.15, 1.95, 2.03, 2.28, 2.41, 2.38, 2.44, 2.05, 2.00, 2.10, 2.56, 2.05, 2.33, 2.03, 2.05, 2.28, 2.23, 2.36, 2.54, 2.36, 2.08, 1.95, 2.21, 2.10, 2.54, 2.33, 2.28, 1.79, 2.00, 1.92, 2.36, 2.18, 2.56, 2.23, 2.08, 2.31, 2.28, 2.64, 2.64	2.1566 + 0.0033x	0.04	2.23	39
44	2.10, 2.17, 2.34, 2.07, 2.49, 2.46, 2.00, 1.85, 2.68, 2.29, 2.68, 2.41, 2.78, 1.80, 2.22, 2.34, 2.34, 2.61, 2.37, 2.32, 2.49, 2.44, 1.98, 2.24, 2.20, 2.12, 2.63, 1.90, 2.05, 2.10, 2.41, 2.24, 2.27, 2.24, 2.22, 2.10, 2.17, 2.29, 2.24, 2.46, 2.34, 2.27, 2.27, 2.98	2.2710 + 0.0011x	0.00	2.29	41
45	2.07, 2.30, 2.13, 2.40, 2.00, 2.63, 2.10, 2.00, 2.87, 2.63, 2.37, 2.40, 2.27, 2.37, 2.17, 2.40, 2.33, 2.60, 2.33, 2.27, 2.10, 2.43, 2.40, 2.03, 2.30, 2.50, 2.30, 2.10, 2.33, 2.27, 2.23, 2.07, 2.33, 2.47, 2.17, 2.07, 2.67, 2.03, 2.17, 2.23, 2.57, 2.33, 2.20, 2.60, 3.03	2.2742 + 0.0022x	0.02	2.32	30
46	2.00, 2.18, 2.23, 1.92, 2.46, 2.05, 2.44, 2.21, 2.38, 2.23, 2.41, 2.21, 2.21, 2.28, 2.38, 2.46, 2.23, 2.26, 2.18, 2.46, 2.05, 2.13, 2.31, 2.13, 2.10, 2.56, 2.46, 2.13, 2.26, 2.21, 2.26, 2.10, 2.56, 2.18, 2.08, 2.18, 2.62, 2.44, 2.64, 2.21, 1.92, 2.44, 2.28, 2.36, 2.05, 2.77	2.2003 + 0.0031x	0.05	2.27	39
47	1.96, 2.11, 2.59, 1.85, 2.26, 2.30, 2.37, 2.11, 2.30, 2.00, 2.26, 2.48, 2.44, 2.26, 2.48, 1.93, 2.04, 2.04, 2.26, 1.96, 2.04, 2.07, 2.07, 2.22, 1.85, 2.44, 1.96, 2.15, 2.30, 2.19, 2.41, 1.96, 2.33, 1.78, 2.59, 2.11,	2.1734 + 0.0006x	0.00	2.19	27

	2.00, 2.22, 2.00, 2.00, 2.59, 2.33, 1.74, 2.26, 1.96, 2.48, 2.74				
48	1.69, 1.73, 2.42, 2.15, 2.19, 1.88, 2.19, 1.88, 2.31, 2.31, 2.12, 1.73, 2.04, 2.00, 2.58, 2.42, 1.96, 2.58, 2.23, 2.15, 2.50, 2.42, 2.23, 2.23, 2.08, 2.23, 1.92, 2.35, 2.31, 2.19, 2.27, 2.54, 2.50, 1.81, 2.12, 2.50, 2.19, 1.85, 2.35, 2.38, 1.96, 2.19, 2.19, 2.12, 2.35, 2.42, 2.73, 2.73	2.0496 + 0.0067x	0.13	2.21	26
49	1.88, 2.67, 2.13, 2.21, 1.92, 1.96, 1.92, 2.54, 2.04, 3.04, 2.17, 2.17, 2.71, 2.25, 2.13, 2.21, 2.25, 2.42, 2.42, 2.33, 2.25, 2.13, 2.17, 2.04, 2.21, 2.13, 2.25, 2.38, 2.33, 2.38, 2.13, 2.00, 2.04, 1.71, 2.50, 1.71, 2.54, 2.13, 2.50, 2.29, 2.08, 2.54, 2.13, 2.38, 2.33, 2.25, 2.79, 2.50, 2.83	2.1786 + 0.0035x	0.03	2.27	24
50	1.62, 2.50, 1.85, 1.88, 2.54, 2.08, 2.08, 1.73, 2.27, 2.27, 2.08, 2.46, 2.38, 2.31, 2.27, 1.85, 2.08, 1.96, 1.85, 2.12, 2.23, 2.27, 2.00, 2.23, 2.15, 2.62, 2.00, 1.88, 2.38, 2.19, 2.23, 2.27, 2.19, 2.35, 2.12, 2.19, 2.58, 2.31, 1.96, 2.58, 1.92, 1.73, 2.00, 2.00, 2.08, 2.27, 2.12, 2.12, 2.31, 2.96	2.0837 + 0.0033x	0.04	2.17	26
51	2.48, 2.05, 1.86, 2.14, 2.52, 2.43, 2.86, 2.29, 2.71, 2.33, 2.48, 2.19, 2.29, 2.29, 2.67, 2.33, 2.10, 2.48, 2.33, 2.62, 2.05, 2.43, 2.33, 2.05, 2.48, 2.57, 2.62, 2.38, 1.90, 1.95, 2.29, 1.81, 2.81, 1.81, 2.57, 2.48, 2.38, 2.62, 2.19, 1.57, 2.14, 1.90, 2.38, 1.95, 2.00, 2.52, 2.14, 2.43, 2.38, 2.52, 2.62	2.3663 - 0.0022x	0.01	2.31	21
52	1.53, 2.12, 2.18, 2.12, 2.29, 1.82, 2.35, 2.41, 2.53, 1.59, 2.59, 2.59, 3.00, 2.00, 3.00, 2.18, 2.12, 2.18, 2.59, 2.88, 2.29, 2.35, 2.12, 2.65, 1.88, 2.24, 2.76, 2.35, 2.06, 2.24, 2.12, 2.29, 2.06, 2.12, 2.82, 2.12, 1.94, 2.65, 1.88, 2.82, 2.29, 2.00, 2.00, 2.29, 2.18, 2.53, 2.29, 2.94, 2.47, 2.47, 1.88, 2.59	2.2290 + 0.0028x	0.02	2.30	17
53	1.76, 2.44, 2.52, 2.28, 2.36, 2.48, 1.60, 2.28, 1.92, 2.44, 2.08, 2.84,	2.2562 + 0.0014x	0.01	2.29	25

	2.48, 1.92, 2.20, 2.60, 2.68, 2.12, 2.56, 2.28, 2.16, 2.48, 2.24, 2.48, 2.00, 1.84, 2.36, 2.28, 2.40, 2.68, 2.20, 2.36, 2.20, 2.24, 2.32, 2.40, 2.16, 1.92, 2.40, 2.16, 2.68, 2.32, 2.48, 2.32, 2.16, 2.16, 2.48, 2.12, 1.96, 2.40, 2.28, 2.28, 2.80				
54	2.33, 1.87, 1.60, 2.27, 2.40, 2.27, 2.07, 2.27, 2.53, 1.80, 2.07, 1.93, 1.87, 2.00, 2.20, 1.93, 2.13, 2.07, 1.60, 2.00, 2.00, 2.00, 1.60, 2.27, 1.93, 2.13, 1.87, 2.73, 1.93, 2.67, 2.07, 2.33, 2.27, 1.87, 2.67, 1.87, 1.87, 2.20, 2.07, 2.27, 2.20, 1.93, 2.33, 2.47, 2.33, 2.07, 2.67, 2.00, 1.67, 2.07, 2.33, 2.33, 2.93, 2.60	1.9959 + 0.0054x	0.08	2.14	15
55	1.73, 2.00, 2.91, 3.09, 2.18, 1.82, 1.73, 1.82, 2.73, 2.27, 2.91, 2.73, 2.82, 3.18, 2.45, 2.18, 1.82, 2.55, 1.55, 2.09, 2.27, 2.45, 2.00, 1.55, 3.64, 1.91, 2.73, 1.91, 2.00, 2.45, 2.36, 2.27, 2.18, 2.00, 2.91, 2.45, 2.55, 2.09, 2.91, 2.73, 2.09, 2.45, 2.09, 2.00, 1.91, 2.00, 1.64, 3.00, 2.09, 2.27, 3.18, 2.45, 2.09, 2.00, 3.00	2.3097 + 0.0007x	0.00	2.33	11
56	1.58, 2.58, 2.25, 2.25, 2.50, 2.75, 1.92, 2.17, 2.58, 2.08, 1.67, 2.17, 2.08, 1.92, 1.83, 1.92, 1.75, 2.25, 2.58, 2.08, 1.92, 2.25, 2.50, 2.00, 2.17, 2.17, 2.17, 2.33, 2.08, 1.92, 2.42, 3.17, 2.17, 2.92, 1.75, 2.00, 2.25, 2.00, 2.42, 2.42, 1.58, 1.83, 2.75, 2.42, 2.17, 1.92, 2.42, 2.33, 2.67, 2.50, 2.25, 2.42, 1.83, 2.42, 2.33, 2.50	2.1207 + 0.0035x	0.03	2.22	12
57	1.64, 1.82, 2.09, 1.64, 2.55, 2.64, 2.64, 2.91, 1.27, 3.00, 1.73, 1.64, 3.09, 2.18, 2.27, 3.09, 1.64, 3.27, 1.91, 2.09, 1.64, 2.73, 1.91, 2.27, 2.18, 2.55, 1.64, 1.82, 1.82, 2.09, 2.18, 1.82, 2.36, 3.18, 1.91, 2.27, 1.55, 1.91, 2.27, 2.09, 2.09, 2.09, 2.18, 2.09, 2.09, 2.00, 1.91, 2.18, 2.27, 1.73, 1.91, 2.27, 2.27, 2.45, 2.18, 2.18, 2.64	2.2076 - 0.0012x	0.00	2.17	11
58	1.67, 2.08, 2.50, 2.08, 2.17, 1.92, 2.42, 2.25, 2.42, 2.42, 2.92, 2.42, 2.67, 2.00, 2.67, 2.67, 2.33, 1.83, 2.00, 2.50, 2.33, 2.33, 2.00, 2.42,	2.2970 - 0.0009x	0.00	2.27	12

	1.92, 1.92, 2.42, 2.17, 2.17, 2.75, 2.25, 2.08, 3.08, 2.33, 2.17, 2.58, 2.17, 2.92, 2.42, 2.33, 2.33, 1.67, 2.00, 2.50, 2.33, 1.92, 1.75, 2.17, 2.33, 1.83, 2.08, 2.33, 2.08, 2.33, 2.58, 1.92, 2.08, 2.83				

In Figure 6 one can see that the course of the parameter *b* is very smooth. This is most likely caused by the great number of sentences of each length. As can be seen from Figure 6, the decrease of parameter *b* is very regular. It can be expressed by $b = 0.8737 \exp(-0.2763n)$, with $R^2 = 0.97$. The results show that long texts display the expected tendency much more clearly. Interestingly enough, we again see that regression coefficient *b* may become negative for longer sentences, although in this case, for sentences with $n > 51$ only. The question remains open if this effect has to be considered as some usual dispersion, including the possibility that ultimately, there is no decrease of word length towards the end of a sentence, but rather some kind of "swinging in" around zero.

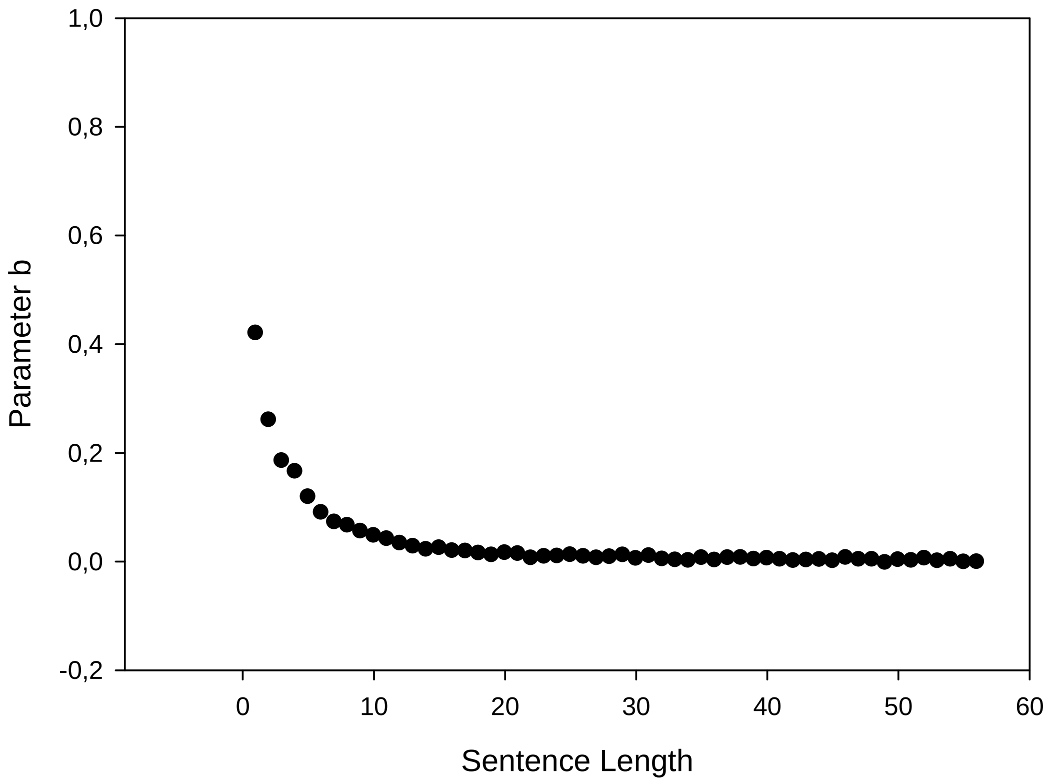

Figure 6. The decrease of parameter *b* in Tolstoj´s *Anna Karenina*

In any case, it seems obvious that the *regularity* of the decrease of parameter *b* with increasing sentence length very much depends on sample size.

Given these findings, the next step must include the analysis of data from further languages – in our case in Slovak, Hungarian, English, Latin and Indonesian texts – in order to arrive at more general conclusions. We therefore continue our analyses with Slovak texts because this language, despite its overall differences, still displays great kinship and similarity to Russian.

2.2. Slovak

Like in case of our analyses of Russian texts, no reliable data were obtained for short individual texts; we therefore merged 30 Slovak journalistic texts to one overall corpus of 433 sentences (with 6072 words). The results for this corpus are presented in Table 3. As can be seen, reliable data can be obtained for $4 \leq n \leq 23$ (with $k \geq 10$).

Table 3
Mean lengths (L) of words in individual positions (x) in sentences of length n in Slovak press texts (M = mean of all positions, k = number of sentences)

n	**Mean syllabic lengths of words *L***	**Linear relation *L = a + bx***	R^2	***M***	***k***
3	1.67, 2.67, 3.33	-	-	-	3
4	2.18, 1.82, 2.82, 3.27	1.4550 + 0.4270x	0.72	2.52	11
5	2.47, 2.13, 2.53, 2.33, 2.73	2.2220 + 0.0720x	0.26	2.44	15
6	2.52, 2.10, 2.43, 2.33, 2.38, 3.05	2.1293 + 0.0969x	0.33	2.47	21
7	2.69, 2.15, 2.38, 1.96, 2.73, 2.77, 3.00	2.1657 + 0.0900x	0.27	2.53	26
8	2.48, 2.62, 1.90, 2.71, 2.38, 2.38, 2.71, 3.05	2.2314 + 0.0661x	0.23	2.53	21
9	2.80, 1.95, 2.50, 2.20, 2.50, 2.30, 1.60, 3.05, 3.10	2.2111 + 0.0467x	0.07	2.44	20
10	2.35, 2.29, 2.77, 2.84, 2.19, 2.77, 2.55, 2.61, 2.71, 2.68	2.4153 + 0.0292x	0.15	2.58	31
11	2.59, 2.32, 2.45, 2.45, 2.55, 2.05, 2.32, 2.23, 2.68, 2.18, 2.68	2.4140 – 0.0008x	0.00	2.41	22
12	2.54, 2.54, 2.21, 2.25, 2.88, 2.04, 2.17, 2.75, 2.13, 2.71, 2.54, 2.75	2.3467 + 0.0173x	0.05	2.46	24
13	2.81, 2.48, 2.38, 2.29, 2.52, 2.19, 2.43, 2.29, 2.19, 2.52, 2.76, 2.48, 2.62	2.4388 + 0.0028x	0.00	2.46	21
14	2.91, 1.91, 2.14, 2.82, 2.91, 2.95, 2.95, 2.68, 2.18, 2.41, 2.23, 2.64, 2.50, 2.91	2.5521 + 0.0039x	0.00	2.58	22
15	2.52, 2.81, 2.33, 2.81, 2.48, 2.71, 2.33, 2.86, 2.48, 2.43, 2.86, 2.62, 2.52, 2.43, 2.81	2.5809 + 0.0024x	0.00	2.60	21
16	2.76, 2.33, 1.94, 2.85, 2.55, 2.27, 1.91, 2.73, 2.39, 2.97, 2.58, 2.39, 2.33, 2.61, 2.94, 2.88	2.3308 + 0.0231x	0.11	2.53	33
17	2.57, 2.67, 2.27, 2.80, 2.40, 2.27, 2.60, 2.23, 2.30, 2.53, 2.20, 2.40, 2.13, 2.03, 2.30, 2.27, 2.53	2.5584 – 0.0196x	0.23	2.38	30
18	3.43, 1.86, 2.00, 2.57, 2.57, 2.43, 2.57, 2.57, 2.43, 2.71, 3.00, 2.57, 3.00, 1.71, 1.71, 2.14, 2.71, 2.43	2.6052 – 0.0145x	0.03	2.47	7
19	2.06, 2.28, 2.67, 2.28, 3.06, 2.50, 2.00,	2.3428 + 0.0197x	0.09	2.54	18

	2.56, 2.78, 2.56, 3.06, 2.56, 2.00, 2.44, 2.67, 2.22, 2.39, 3.22, 2.94				
20	2.72, 2.83, 2.33, 2.56, 3.00, 3.06, 2.39, 1.83, 2.17, 2.72, 2.78, 2.61, 2.67, 2.61, 2.44, 3.00, 2.50, 2.56, 2.72, 2.67	2.5975 + 0.0010x	0.00	2.61	18
21	2.67, 2.47, 2.40, 1.93, 3.73, 3.13, 3.07, 2.33, 2.60, 2.67, 2.47, 1.80, 3.40, 1.67, 2.20, 2.27, 2.13, 2.27, 2.27, 2.47, 2.67	2.7527 – 0.0225x	0.08	2.50	15
22	3.50, 2.10, 2.90, 2.60, 2.50, 2.20, 2.70, 3.40, 2.00, 3.50, 2.10, 2.50, 2.50, 3.00, 2.20, 2.80, 2.20, 1.90, 2.20, 2.80, 3.00, 3.00	2.7091 – 0.0079x	0.01	2.62	10
23	2.45, 3.09, 2.55, 2.00, 3.27, 2.18, 2.64, 2.18, 1.73, 3.18, 2.91, 2.27, 2.00, 2.64, 2.73, 2.09, 2.00, 2.73, 2.00, 2.00, 2.36, 1.91, 3.00	2.6158 – 0.0154x	0.05	2.43	11

Although Slovak and Russian are typologically quite similar to each other, the Slovak results strongly differ from those for our Russian data: In the Slovak texts, the regression coefficient quite soon takes values of $b < 0$, as can be seen in Figure 7.

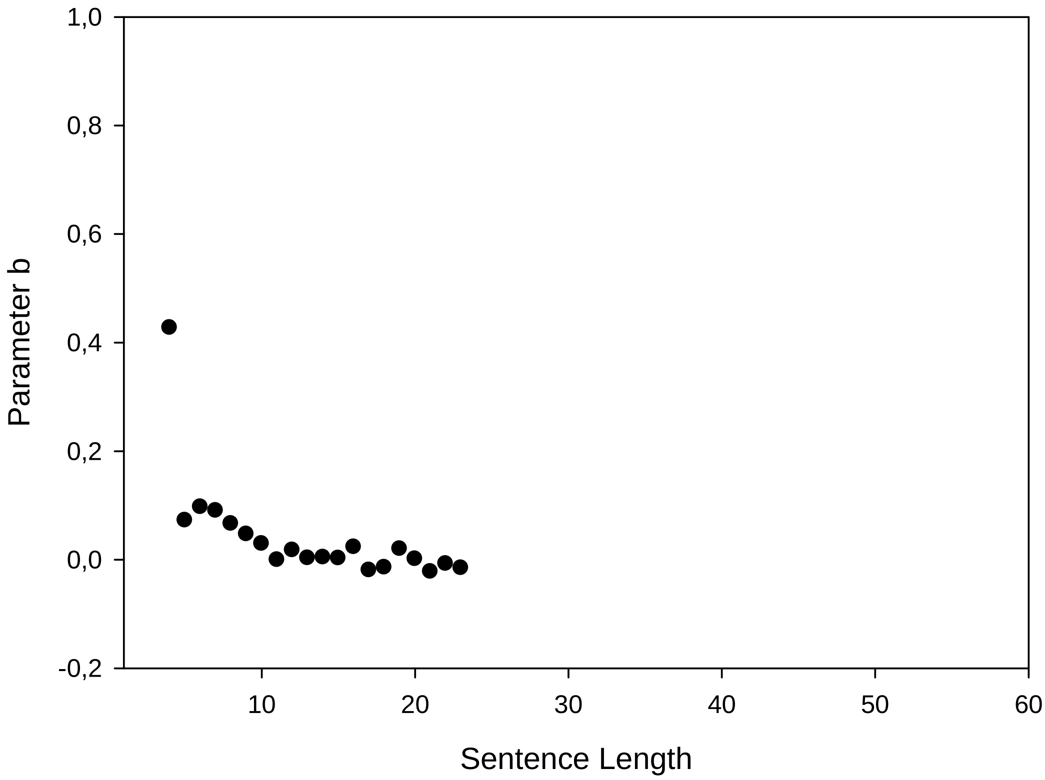

Figure 7. Dependence of *b* on sentence length in Slovak

Here we obtain b = 25.2594 exp(-1.0298n) with R^2 = 0.88. Although the tendency is, in principle, the same as observed before, the rapid decrease of parameter *b* is conspicuous. Since there is no easy explanation for this phenomenon, which may be caused by the text material, a mixture of press texts, let us additionally analyze a longer Slovak text, which is more likely to be homogeneous, the novel *Veterná ružica* (1995), written by Vincent Šikula. For this text, which comprises 13961 words (in 881 sentences), reliable results are obtained for $3 \leq n \leq 25$, which are presented in Table 4.

Table 4
Mean lengths (L) of words in individual positions (x) in sentences of length n in Slovak text *Veterná ružica* by Šikula
(M = mean of all positions, k = number of sentences)

n	**Mean syllabic lengths of words *L***	**Linear relation *L = a + bx***	R^2	***M***	***k***
3	1.95, 1.59, 2.40	1.5300 + 0.2250x	0.31	1.98	58
4	1.91, 1.40, 1.74, 2.50	1.3600 + 0.2110x	0.35	1.89	78
5	1.83, 1.43, 1.59, 1.91, 2.62	1.2580 + 0.2060x	0.51	1.88	90
6	1.78, 1.51, 1.71, 1.75, 1.79, 2.63	1.3487 + 0.1466x	0.49	1.86	68
7	1.71, 1.40, 1.42, 1.54, 1.83, 1.77, 2.69	1.1814 + 0.1461x	0.51	1.77	48
8	1.71, 1.42, 1.97, 2.06, 1.81, 1.81, 1.90, 3.00	1.3868 + 0.1274x	0.46	1.96	31
9	1.87, 1.70, 1.78, 1.78, 1.83, 1.70, 1.26, 1.96, 2.70	1.5939 + 0.0497x	0.13	1.84	23
10	2.08, 1.33, 1.61, 1.67, 1.75, 2.00, 1.75, 1.83, 2.22, 2.44	1.4993 + 0.0670x	0.40	1.87	36
11	1.44, 1.56, 1.44, 2.00, 1.93, 1.85, 1.81, 1.67, 1.67, 1.59, 2.63	1.4547 + 0.0544x	0.29	1.78	27
12	2.07, 1.75, 1.82, 1.79, 1.89, 1.96, 1.79, 1.82, 2.00, 1.79, 2.04, 2.82	1.7044 + 0.0396x	0.24	1.96	28
13	2.03, 1.41, 1.93, 1.83, 2.00, 1.86, 2.03, 1.76, 1.90, 1.62, 1.97, 1.83, 2.59	1.7242 + 0.0258x	0.14	1.90	29
14	1.90, 1.70, 1.60, 1.75, 2.15, 2.00, 2.20, 1.90, 1.90, 2.40, 1.50, 1.75, 2.10, 2.75	1.7126 + 0.0345x	0.19	1.97	20
15	1.62, 1.71, 1.67, 2.05, 2.14, 1.76, 2.05, 2.05, 1.48, 1.76, 1.62, 1.86, 2.14, 1.52, 2.95	1.6740 + 0.0273x	0.11	1.89	21
16	2.00, 1.14, 1.90, 2.19, 2.19, 2.00, 2.14, 2.14, 2.19, 1.43, 1.90, 1.95, 1.71, 1.76, 1.86, 2.90	1.8033 + 0.0187x	0.05	1.96	21
17	2.06, 1.71, 1.35, 2.12, 1.76, 1.94, 1.76, 2.71, 1.53, 1.71, 1.88, 2.12, 1.82, 1.59, 1.94, 1.76, 2.88	1.7475 + 0.0192x	0.06	1.92	17
18	1.75, 1.83, 1.75, 2.42, 1.75, 1.83, 2.25, 1.83, 1.75, 2.33, 1.58, 2.25, 2.00, 1.42, 2.08, 2.00, 1.83, 2.75	1.8239 + 0.0150x	0.06	1.97	12
19	2.07, 1.93, 1.47, 2.47, 2.40, 1.73, 1.93, 1.93, 2.07, 1.93, 1.87, 1.87, 1.27, 2.20, 2.20, 2.00, 1.80, 1.73, 2.60	1.9504 + 0.0022x	0.00	1.97	15
20	2.00, 1.55, 1.95, 1.95, 2.20, 2.00, 1.85, 1.95, 1.75, 1.90, 1.75, 2.10, 1.65, 1.95, 1.70, 1.35, 1.65, 1.65, 2.20, 2.20	1.9084 - 0.0041x	0.01	1.87	20

21	1.86, 1.43, 2.14, 2.07, 1.86, 1.86, 2.43, 1.79, 2.50, 2.36, 2.57, 2.07, 1.86, 2.00, 1.64, 1.64, 1.79, 1.57, 1.36, 1.86, 3.07	1.9784 + 0.0008x	0.00	1.99	14
22	2.38, 1.75, 1.50, 1.88, 2.13, 2.38, 2.00, 1.88, 1.88, 1.38, 2.25, 2.75, 1.63, 2.00, 1.38, 1.75, 2.25, 1.50, 2.50, 2.25, 2.75, 2.88	1.8008 + 0.0215x	0.10	2.05	8
23	2.44, 1.33, 1.33, 2.11, 2.78, 2.11, 2.00, 1.22, 2.56, 2.11, 2.33, 1.56, 1.78, 2.00, 1.56, 1.44, 1.89, 2.11, 2.44, 2.22, 1.89, 2.00, 3.11	1.8459 + 0.0140x	0.04	2.01	9
24	1.83, 1.58, 1.67, 1.58, 2.00, 1.83, 1.67, 2.17, 2.33, 1.58, 2.42, 2.25, 1.67, 2.25, 2.00, 1.92, 1.75, 2.25, 2.33, 1.50, 1.75, 1.75, 2.50, 2.75	1.7295 + 0.0194x	0.16	1.97	12
25	2.00, 1.75, 1.67, 1.92, 1.92, 2.00, 2.25, 1.83, 1.75, 1.83, 2.00, 1.83, 1.83, 2.08, 1.67, 1.83, 2.42, 1.75, 1.58, 2.00, 1.92, 1.75, 2.00, 2.42, 2.50	1.8101 + 0.0100x	0.09	1.94	12

The course of parameter *b* can be modelled by the regression equation $b = 0.4325 \exp(-0.1812n)$ with $R^2 = 0.95$. Here the decrease is not as rapid as in the press texts, and in the analyzed domain ($3 \leq n \leq 25$), $b < 0$ in only one case (at sentence length 20). Figure 8 illustrates the results of the regression analyses.

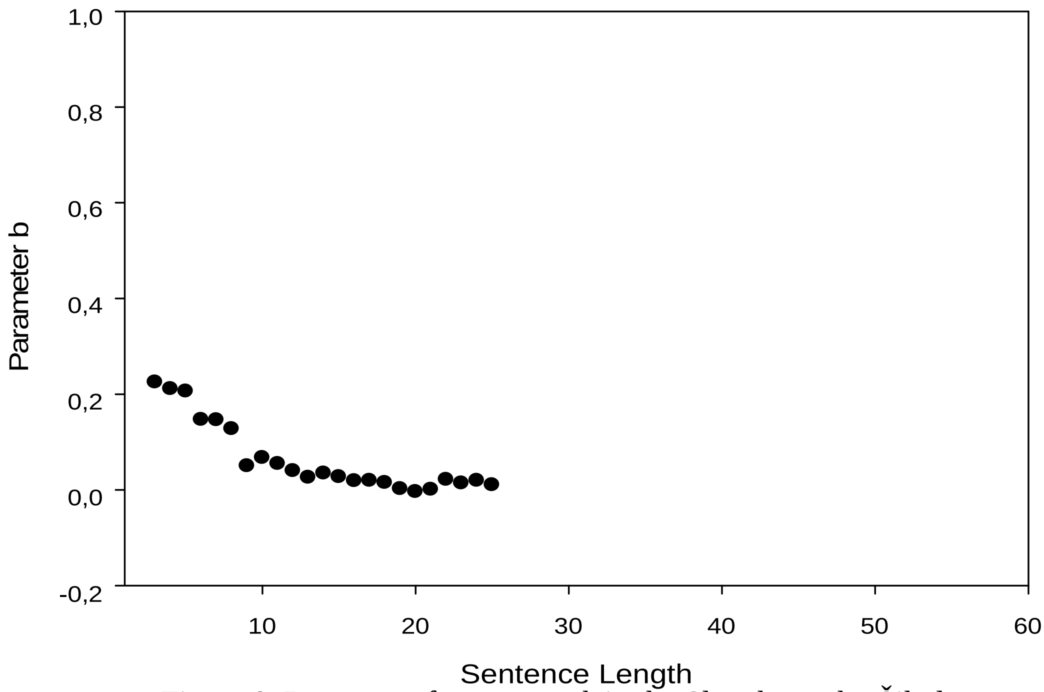

Figure 8. Decrease of parameter *b* in the Slovak text by Šikula

We may therefore conclude that the general tendency predicted by hypothesis (I) – i.e., decrease of *b* with increasing text length – holds also for Slovak.

2.3. Hungarian

For Hungarian, we analyzed two longer texts, both of approximately equal length, to compare the results with other languages, on the one hand, and to get additional information as to the degree of the variation within a given language.

The first data were taken from the novel *Az ezüst kecske* (1897) by Sándor Bródy (1863-1924), with 3356 sentences (and 46049 words), the second Hungarian text was a novel by Frigyes Karinthy (1887-1938), *Utazás a koponyám körül* (1937), with 44125 words (from 3440 sentences).

The results for the first text are presented in Table 5; all length classes up to the point when k was 10 for the last time are taken into consideration (in this case n = 40).

Table 5
Mean lengths (L) of words in individual positions (x) in sentences of length n in the Hungarian text *Az ezüst kecske* by Sándor Bródy
(M = mean of all positions, k = number of sentences)

n	**Mean syllabic lengths of words**	**Linear relation** $L = a + bx$	R^2	M	k
3	1.71, 1.87, 2.41	1.2967 + 0.3500x	0.91	2.2115	118
4	1.72, 1.97, 1.97, 2.50	1.4550 + 0.2340x	0.85	2.086	159
5	1.75, 2.04, 1.99, 1.93, 2.63	1.5730 + 0.1650x	0.61	1.9972	191
6	1.82, 1.86, 2.03, 2.00, 1.99, 2.66	1.6040 + 0.1303x	0.64	2.044	213
7	1.69, 2.03, 2.12, 1.97, 2.00, 2.04, 2.56	1.7000 + 0.0896x	0.56	2.0681	188
8	1.65, 2.00, 1.92, 2.15, 1.92, 1.94, 2.03, 2.81	1.6186 + 0.0964x	0.49	2.0595	198
9	1.80, 2.18, 1.93, 2.12, 2.06, 2.08, 2.09, 2.29, 2.61	1.8081 + 0.0642x	0.59	2.0593	179
10	1.76, 2.09, 2.17, 2.02, 2.04, 2.06, 1.96, 2.17, 2.04, 2.78	1.8200 + 0.0525x	0.37	2.0524	187
11	1.74, 2.11, 2.07, 2.23, 2.03, 2.20, 2.10, 2.12, 1.92, 2.34, 2.72	1.8589 + 0.0475x	0.40	2.1291	179
12	1.71, 2.19, 2.11, 2.13, 2.14, 2.07, 1.91, 2.04, 2.24, 2.26, 2.07, 2.64	1.8920 + 0.0360x	0.35	2.108	140
13	1.76, 2.29, 2.18, 2.21, 2.07, 1.97, 2.13, 2.16, 2.01, 2.08, 2.13, 2.18, 2.71	1.9665 + 0.0254x	0.21	2.1422	143
14	1.94, 2.39, 1.92, 2.30, 2.15, 2.07, 2.03, 2.14, 2.10, 2.31, 2.42, 2.04, 2.17, 2.72	2.0175 + 0.0234x	0.20	2.1262	105
15	1.75, 2.20, 2.22, 2.13, 2.31, 2.22, 2.10, 2.13, 1.87, 2.20, 2.13, 2.02, 2.04, 2.15, 2.65	2.0313 + 0.0138x	0.09	2.1452	134
16	1.71, 2.12, 2.17, 2.12, 2.09, 2.04, 2.09, 2.12, 2.04, 2.07, 2.03, 2.10, 2.17, 2.05, 2.25, 2.78	1.9125 + 0.0246x	0.31	2.1939	139
17	1.80, 2.09, 2.35, 2.00, 1.97, 2.10, 1.99, 2.23, 2.16, 1.95, 2.03, 1.97,	1.9236 + 0.0223x	0.20	2.1408	115

	2.07, 2.16, 2.04, 2.23, 2.97				
18	1.75, 2.10, 2.07, 2.23, 2.20, 2.25, 2.14, 2.25, 2.06, 2.04, 2.26, 2.24, 1.96, 2.41, 2.20, 1.98, 2.07, 2.66	2.0233 + 0.0143x	0.16	2.1223	102
19	1.84, 2.09, 2.21, 2.16, 2.06, 2.08, 2.37, 2.17, 2.08, 2.15, 2.26, 2.10, 2.07, 2.21, 2.13, 2.18, 2.11, 2.11, 2.79	2.0240 + 0.0143x	0.19	2.1228	89
20	1.80, 2.00, 2.28, 2.05, 2.14, 2.01, 2.17, 2.19, 2.08, 2.05, 1.93, 2.23, 2.00, 2.06, 2.29, 1.89, 2.08, 2.13, 2.12, 2.71	1.9854 + 0.0119x	0.14	2.158	83
21	1.61, 2.18, 2.11, 2.05, 2.09, 2.02, 1.95, 2.26, 2.36, 2.32, 1.94, 2.21, 2.03, 2.02, 2.36, 2.14, 2.03, 2.15, 2.42, 2.26, 3.08	1.8997 + 0.0247x	0.31	2.1668	66
22	1.77, 1.97, 2.23, 2.12, 1.94, 2.17, 2.30, 2.52, 2.22, 1.93, 2.30, 2.25, 2.03, 2.06, 2.16, 2.09, 2.22, 2.33, 2.39, 2.12, 2.32, 2.90	1.9816 + 0.0188x	0.27	2.1108	69
23	1.91, 2.17, 1.92, 2.17, 2.28, 2.09, 2.09, 2.03, 2.16, 2.05, 2.41, 2.22, 2.13, 2.11, 2.03, 2.31, 1.89, 2.02, 2.19, 2.45, 2.25, 2.31, 2.97	1.9818 + 0.0166x	0.25	2.1703	64
24	1.54, 2.10, 2.08, 2.17, 1.81, 2.20, 1.93, 2.00, 2.07, 2.41, 2.12, 1.95, 1.95, 2.19, 2.07, 2.29, 2.02, 2.08, 2.31, 2.44, 2.12, 2.44, 2.29, 2.83	1.8572 + 0.0228x	0.41	2.197	59
25	1.83, 2.29, 2.02, 2.14, 1.74, 2.43, 2.17, 1.93, 2.17, 1.95, 2.17, 2.19, 1.76, 2.50, 2.14, 2.26, 2.05, 2.00, 2.10, 2.52, 2.21, 2.38, 2.19, 2.21, 3.14	1.9424 + 0.0182x	0.22	2.1807	42
26	1.70, 2.16, 2.14, 2.05, 2.14, 1.88, 2.44, 2.23, 1.93, 2.09, 2.09, 2.44, 2.26, 2.16, 2.14, 1.93, 2.21, 2.02, 2.12, 2.16, 2.02, 2.26, 2.05, 2.19, 2.07, 2.56	2.0418 + 0.0067x	0.08	2.1419	43
27	1.96, 2.33, 2.38, 2.09, 1.93, 1.78, 2.40, 2.20, 2.07, 2.11, 2.00, 2.09, 2.07, 2.31, 2.22, 2.00, 2.00, 2.18, 2.27, 1.89, 1.98, 2.49, 2.24, 2.09, 2.24, 1.93, 2.84	2.0658 + 0.0061x	0.05	2.18	45
28	1.74, 2.20, 2.26, 2.09, 2.40, 2.09, 2.29, 2.09, 1.97, 2.03, 2.34, 1.94, 2.20, 1.89, 2.31, 2.06, 2.31, 2.14, 2.11, 2.17, 2.37, 2.31, 2.00, 1.94, 1.71, 2.63, 1.77, 3.23	2.0566 + 0.0074x	0.04	2.1324	35
29	1.71, 1.92, 2.21, 2.17, 2.42, 2.08,	2.0890 + 0.0080x	0.06	2.1514	24

	2.42, 2.67, 2.08, 2.13, 2.17, 2.13, 2.38, 2.13, 1.83, 2.08, 2.25, 2.42, 2.25, 1.88, 2.75, 2.00, 1.96, 2.50, 1.92, 2.00, 2.17, 2.50, 2.92				
30	1.88, 2.19, 1.96, 2.19, 2.27, 2.23, 2.23, 2.00, 2.23, 2.42, 1.96, 2.00, 2.19, 1.81, 2.00, 2.15, 2.62, 2.12, 2.00, 1.92, 2.27, 2.15, 2.73, 2.23, 2.00, 2.46, 2.23, 1.92, 2.50, 2.38	2.0625 + 0.0072x	0.08	2.1643	26
31	2.05, 2.11, 2.16, 2.37, 2.16, 1.84, 1.89, 2.47, 1.89, 2.37, 2.32, 2.53, 2.05, 1.63, 2.47, 2.37, 1.63, 2.84, 2.26, 1.79, 2.42, 1.68, 2.26, 2.47, 2.32, 2.11, 1.74, 1.79, 1.95, 2.47, 2.89	2.1187 + 0.0033x	0.01	2.2069	19
32	1.77, 2.77, 2.00, 1.69, 2.08, 2.23, 2.38, 2.38, 2.15, 2.54, 2.46, 2.00, 2.23, 2.15, 2.54, 2.38, 2.00, 2.62, 2.38, 2.00, 2.31, 1.77, 2.08, 1.77, 2.62, 2.15, 2.46, 2.62, 2.08, 2.54, 2.00, 3.00	2.1337 + 0.0073x	0.05	2.1756	13
33	1.64, 2.09, 2.27, 2.00, 1.82, 2.18, 2.27, 2.55, 2.18, 2.36, 2.00, 1.82, 2.18, 2.09, 1.91, 2.09, 2.27, 2.00, 2.00, 1.73, 2.09, 2.45, 1.82, 2.45, 2.36, 2.09, 2.45, 1.73, 2.18, 2.09, 2.00, 2.00, 2.73	2.0423 + 0.0044x	0.03	2.1715	11
34	1.69, 1.85, 2.23, 1.85, 2.69, 2.23, 2.23, 2.23, 1.69, 2.08, 2.38, 2.00, 2.23, 1.85, 2.15, 2.00, 2.54, 1.77, 2.62, 2.69, 2.38, 3.08, 2.46, 2.31, 1.92, 1.85, 1.92, 1.85, 2.54, 2.23, 2.46, 1.92, 2.23, 3.15	2.0307 + 0.0105x	0.08	2.2548	13
35	1.73, 2.53, 1.93, 2.20, 1.93, 2.67, 2.33, 2.47, 2.13, 2.07, 2.47, 1.93, 2.47, 2.40, 2.20, 1.73, 2.47, 2.47, 1.53, 1.73, 2.33, 2.60, 1.87, 2.00, 1.93, 1.80, 2.47, 1.73, 2.47, 2.27, 2.20, 3.07, 1.87, 2.47, 3.1	2.1109 + 0.0058x	0.03	2.1185	15
36	1.78, 2.33, 1.67, 1.67, 2.44, 1.78, 2.00, 2.22, 2.33, 1.78, 2.00, 2.44, 2.22, 1.78, 2.78, 2.22, 1.67, 2.44, 1.78, 2.33, 2.67, 2.00, 2.56, 1.89, 2.00, 1.56, 2.56, 1.89, 1.67, 2.44, 2.56, 1.67, 1.44, 2.00, 2.10, 3.10	2.0238 + 0.0044x	0.01	2.2149	9
37	1.71, 1.71, 1.71, 1.71, 2.14, 2.00, 1.29, 2.14, 2.14, 2.29, 2.43, 1.86, 1.86, 2.29, 2.86, 2.57, 2.14, 2.29, 2.43, 2.71, 1.71, 2.00, 1.71, 2.86,	1.9008 + 0.0129x	0.12	2.2152	7

	1.86, 2.43, 1.86, 2.57, 2.29, 2.29, 2.14, 1.86, 2.14, 2.00, 1.90, 2.10, 3.40				
38	1.60, 1.40, 2.20, 2.20, 3.10, 1.90, 1.80, 3.00, 2.20, 2.60, 1.80, 1.80, 2.20, 2.00, 2.10, 1.60, 2.60, 2.40, 2.50, 2.00, 2.30, 2.20, 1.60, 2.40, 2.20, 1.10, 1.60, 2.50, 2.00, 1.80, 2.20, 2.50, 3.30, 2.20, 2.10, 1.70, 2.40, 2.80	2.0539 + 0.0052x	0.02	2.1049	10
39	2.67, 3.00, 3.00, 2.33, 2.33, 3.00, 2.67, 2.67, 2.67, 1.00, 2.33, 1.33, 2.33, 3.67, 1.67, 1.33, 2.33, 3.67, 1.33, 2.00, 2.00, 2.33, 3.00, 1.67, 2.67, 1.33, 1.33, 2.00, 1.00, 1.33, 1.67, 3.33, 2.00, 2.33, 2.00, 2.30, 2.70, 2.00, 3.00	2.4878 - 0.0124x	0.04	2.1467	3
40	1.50, 2.00, 2.10, 2.00, 2.20, 2.00, 3.00, 2.10, 2.50, 1.50, 2.70, 2.00, 1.80, 1.80, 2.30, 2.40, 2.10, 1.80, 1.70, 2.10, 1.50, 1.50, 1.90, 2.20, 2.10, 1.80, 1.60, 1.80, 1.50, 2.40, 1.70, 1.70, 2.00, 1.90, 2.20, 2.00, 2.50, 2.30, 2.20, 2.50	2.0262 - 0.0002x	0.00	2.1553	10

Again one sees that in the linear regression the slope (*b*) decreases from highly significant positive values to small negative ones as shown in Figure 9. The *b*-sequence is very regular; a straight line seems to be no good model, but again, with $b = 0.7432 \exp(-0.2756n)$, the determination coefficient is very satisfying ($R^2 = 0.97$).

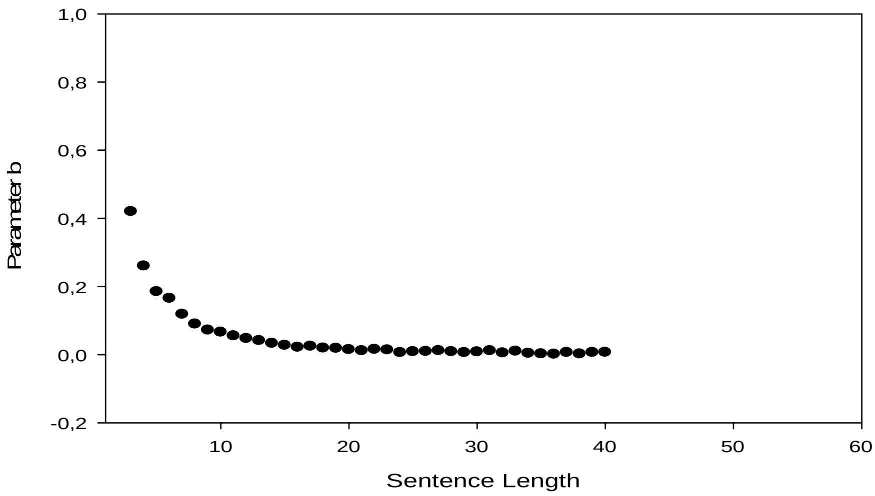

Figure 9. Decrease of *b* in Bródy’s Hungarian text

The results for second Hungarian text are represented in Table 6. The function expressing the decrease of *b* in Figure 10 is $b = 0.3061\exp(-0.1778n)$, with $R^2 = 0.92$.

Table 6
Mean lengths (*L*) of words in individual positions (*x*) in sentences of length *n* in the Hungarian text *Utazás a koponyám körül* by Frigyes Karinthy
(*M* = mean of all positions, *k* = number of sentences)

n	**Mean syllabic lengths of words**	**Linear relation** $L = a + bx$	R^2	***M***	***k***
3	2.11, 2.08, 2.50	1.8400 + 0.1950x	0.69	2.23	210
4	2.04, 1.89, 2.02, 2.58	1.6950 + 0.1750x	0.55	2.13	223
5	1.95, 2.05, 2.16, 1.94, 2.45	1.8430 + 0.0890x	0.45	2.11	195
6	1.98, 2.13, 2.09, 2.09, 2.10, 2.61	1.8607 + 0.0874x	0.54	2.17	215
7	2.13, 2.06, 2.10, 1.95, 2.20, 2.16, 2.67	1.9071 + 0.0686x	0.42	2.18	188
8	1.90, 2.03, 2.19, 2.01, 2.16, 2.09, 2.24, 2.80	1.7918 + 0.0857x	0.59	2.18	182
9	1.87, 2.25, 2.07, 2.17, 2.14, 2.24, 2.25, 2.22, 3.01	1.8383 + 0.0817x	0.52	2.25	153
10	2.18, 2.43, 2.05, 2.18, 2.08, 2.19, 2.07, 2.27, 2.13, 2.90	2.0727 + 0.0319x	0.14	2.25	153
11	2.09, 2.02, 2.18, 2.01, 2.26, 2.30, 2.20, 2.32, 1.98, 2.13, 2.84	1.9855 + 0.0377x	0.27	2.21	128
12	1.88, 2.06, 2.27, 2.07, 2.32, 2.13, 2.24, 2.44, 2.13, 2.19, 2.15, 2.72	1.9835 + 0.0359x	0.37	2.22	125
13	1.96, 2.33, 2.19, 1.98, 2.39, 2.12, 2.24, 2.28, 2.27, 2.34, 2.18, 2.13, 2.96	2.0300 + 0.0327x	0.27	2.26	113
14	1.91, 2.10, 2.25, 2.20, 1.98, 2.27, 2.20, 2.19, 2.11, 2.13, 2.10, 2.18, 2.17, 2.92	1.9821 + 0.0282x	0.26	2.19	119
15	1.76, 2.08, 2.16, 2.15, 2.27, 2.35, 2.17, 2.26, 2.17, 2.02, 2.28, 2.23, 2.03, 2.17, 2.78	2.0000 + 0.0240x	0.25	2.19	92
16	2.00, 2.31, 2.21, 2.23, 2.07, 2.19, 2.48, 2.22, 2.26, 2.35, 2.16, 2.30, 2.23, 2.19, 2.46, 2.91	2.0795 + 0.0243x	0.31	2.29	96
17	2.03, 2.22, 2.32, 2.03, 2.36, 2.06, 2.31, 2.41, 2.03, 2.24, 2.20, 2.57, 2.22, 2.24, 2.16, 2.28, 2.87	2.0954 + 0.0191x	0.21	2.27	87
18	1.89, 2.16, 2.04, 2.40, 2.41, 2.15, 2.15, 2.36, 2.19, 2.24, 2.35, 2.44, 2.03, 2.32, 2.23, 2.00, 2.31, 2.88	2.0909 + 0.0170x	0.17	2.25	75
19	2.09, 2.11, 2.32, 2.33, 1.97, 2.08, 2.05, 2.15, 2.24, 2.30, 2.12, 2.05, 2.03,2.29,2.27,2.12,2.18,2.36,3.02	2.0377 + 0.0177x	0.19	2.21	66
20	2.04, 1.96, 2.21, 2.02, 2.16, 2.00, 2.13, 2.32, 2.39, 2.52, 2.45, 2.27, 2.57, 2.13, 2.43, 2.30, 1.95, 2.34, 2.45, 3.18	1.9968 + 0.0280x	0.34	2.29	56
21	2.10, 1.84, 2.20, 2.31, 2.17, 2.17, 2.29, 2.14, 2.17, 2.01, 2.29, 2.31, 2.00, 2.23, 2.11, 2.34, 2.04, 2.04, 2.36, 1.97, 2.91	2.0728 + 0.0107x	0.10	2.19	70
22	1.95, 2.49, 2.11, 2.49, 2.10, 2.13, 2.48, 2.30, 2.15, 2.43, 2.34, 2.15, 2.26, 2.18, 2.25, 2.15, 2.15, 2.07, 2.30, 2.36, 2.84, 2.85	2.1481 + 0.0129x	0.13	2.30	61
23	2.12, 2.21, 2.40, 2.27, 2.50, 2.31, 2.40, 2.02, 2.23, 2.21, 2.06, 2.02, 2.04, 2.15, 1.92, 2.31,	2.1686 + 0.0070x	0.05	2.25	52

	2.27, 2.13, 2.33, 2.23, 2.46, 2.23, 2.98				
24	1.89, 2.27, 1.95, 2.14, 1.93, 2.36, 2.09, 2.23, 1.98, 2.09, 2.52, 2.07, 2.07, 2.48, 2.57, 2.11, 2.07, 2.55, 2.48, 2.23, 2.39, 2.32, 2.18, 3.00	1.9890 + 0.0208x	0.32	2.25	44
25	1.66, 2.28, 2.14, 1.79, 2.10, 2.24, 2.14, 2.24, 2.55, 2.17, 2.38, 2.21, 2.21, 2.24, 2.07, 2.31, 2.21, 2.38, 2.14,2.03,1.90,2.24,2.14,2.10,2.90	2.0603 + 0.0100x	0.10	2.19	29
26	1.84, 2.30, 1.95, 2.16, 2.27, 1.97, 2.35, 2.38, 2.51, 2.05, 1.97, 2.59, 2.19, 2.22, 2.14, 2.30, 2.22, 2.08, 2.08, 2.08, 2.03, 2.22, 2.19, 2.14, 2.11, 3.11	2.1042 + 0.0078x	0.06	2.21	37
27	2.05, 2.11, 2.22, 2.00, 2.24, 2.43, 1.95, 2.38, 2.35, 2.51, 2.27, 2.00, 2.00, 2.22, 2.22, 2.62, 2.24, 2.46, 2.38, 2.68, 2.03, 2.41, 2.22, 1.97, 2.16, 2.49, 2.86	2.1274 + 0.0107x	0.13	2.28	37
28	2.03, 2.13, 2.17, 2.17, 2.27, 2.40, 2.13, 2.10, 2.33, 2.20, 2.10, 2.13, 2.10, 2.20, 1.93, 2.20, 2.50, 2.00, 2.00, 2.03, 2.33, 1.87, 2.27, 2.30, 2.13, 2.37, 2.13, 3.10	2.0968 + 0.0072x	0.07	2.20	30
29	2.24, 2.08, 2.00, 1.80, 2.28, 2.32, 1.96, 2.48, 1.92, 2.12, 2.20, 1.96, 1.76, 2.52, 2.16, 2.04, 2.12, 1.88, 2.40, 2.44, 2.32, 2.32, 2.28, 2.60, 2.40, 2.20, 2.36, 2.92, 2.92	1.9523 + 0.0193x	0.33	2.24	25
30	1.85, 2.47, 2.09, 2.06, 2.09, 2.47, 2.00, 2.18, 2.06, 2.21, 2.18, 2.12, 2.24, 2.24, 2.29, 2.47, 2.00, 1.94, 2.15, 2.24, 2.44, 2.12, 2.21, 2.03, 2.38, 2.15, 2.44, 2.56, 2.29, 3.00	2.0547 + 0.0115x	0.20	2.23	34
31	2.16, 2.24, 1.92, 2.12, 2.20, 2.84, 2.16, 2.36, 2.72, 2.20, 2.08, 2.24, 2.56, 2.08, 1.92, 2.48, 2.00, 2.28, 2.56, 2.16, 2.04, 2.00, 2.24, 2.08, 2.76, 2.24, 2.36, 2.20, 1.88,1.92,2.92	2.2446 + 0.0007x	0.00	2.26	25
32	1.78, 2.33, 2.00, 2.22, 2.15, 2.37, 2.04, 2.04, 2.63, 2.22, 2.26, 2.48, 2.19, 1.89, 1.96, 2.52, 2.19, 2.07, 2.07, 2.22, 2.70, 2.63, 2.22, 1.59, 2.11, 2.33, 2.59, 2.00, 2.37, 2.11, 2.56, 2.48	2.1244 + 0.0063x	0.05	2.23	27
33	2.22, 2.00, 2.39, 1.94, 2.06, 1.89, 2.06, 2.50, 2.28, 2.28, 2.67, 2.00, 2.22, 2.28, 2.33, 2.22, 1.67, 3.17, 2.39, 2.50, 2.33, 2.06, 2.56, 2.56, 2.00, 1.94, 2.28, 2.22, 2.33, 2.11, 2.17, 2.28, 2.83	2.1589 + 0.0062x	0.04	2.26	18
34	2.18, 2.23, 2.45, 2.05, 1.82, 2.45, 2.41, 1.91, 1.86, 2.18, 1.68, 2.23, 2.68, 2.14, 2.32, 2.23, 2.23, 1.91, 2.05, 2.68, 2.27, 2.23, 2.36, 2.09, 1.95, 2.50, 2.32, 2.50, 2.05, 2.27, 2.14, 2.05, 2.36, 2.82	2.1174 + 0.0061x	0.06	2.22	22
35	2.27, 1.91, 2.64, 1.55, 2.09, 1.91, 1.73, 1.64, 2.73, 1.73, 3.18, 1.82, 1.82, 1.91, 2.18, 2.18, 2.09, 1.45, 2.64, 2.55, 2.55, 2.18, 2.00, 2.18,	2.0832 + 0.0012x	0.00	2.11	11

	1.64, 2.55, 1.73, 2.27, 2.00, 2.55, 1.91, 1.73, 1.91, 1.82, 2.64				
36	1.64, 2.14, 2.50, 2.50, 1.93, 2.79, 2.64, 2.43, 1.79, 2.50, 2.07, 2.50, 2.36, 2.29, 2.14, 2.00, 2.00, 2.07, 2.36, 2.43, 1.93, 1.79, 1.86, 2.36, 2.50, 2.43, 2.07, 2.21, 1.93, 2.86, 2.14, 1.86, 2.36, 1.79, 2.50, 2.64	2.2351 - 0.0002x	0.00	2.23	14
37	3.00, 2.50, 2.50, 2.75, 1.63, 2.50, 2.50, 1.63, 2.88, 2.13, 2.38, 2.75, 3.00, 2.63, 2.88, 2.75, 2.63, 2.00, 3.13, 2.38, 1.75, 2.50, 2.88, 2.13, 2.50, 2.25, 1.75, 1.63, 1.25, 2.38, 2.13, 2.25, 1.75, 2.25, 1.75, 2.13, 2.63	2.6394 - 0.0159x	0.14	2.34	8
38	1.80, 2.20, 2.20, 1.80, 2.70, 2.20, 2.00, 1.70, 1.90, 2.00, 2.00, 2.30, 2.20, 2.00, 2.10, 1.70, 3.10, 2.70, 2.60, 2.30, 1.60, 3.00, 2.10, 2.80, 2.10, 2.20, 2.70, 2.40, 2.00, 2.40, 2.30, 2.20, 2.50, 2.40, 3.20, 3.10, 2.00, 3.00	1.9653 + 0.0173x	0.21	2.30	10
39	2.14, 2.86, 2.00, 2.29, 2.00, 2.29, 1.57, 1.86, 1.86, 2.71, 1.71, 1.71, 3.14, 2.14, 2.71, 2.29, 1.86, 1.71, 2.29, 2.00, 2.71, 2.00, 1.86, 2.43, 1.29, 1.71, 2.14, 2.00, 2.57, 1.43, 2.57, 2.00, 1.57, 2.43, 1.57, 1.14, 2.43, 1.86, 3.00	2.2024 - 0.0052x	0.02	2.10	7
40	1.50, 2.67, 2.83, 2.00, 1.83, 1.33, 2.50, 2.17, 2.83, 2.00, 2.00, 2.67, 3.33, 2.50, 3.00, 2.00, 2.83, 1.17, 2.00, 2.00, 2.50, 1.67, 2.50, 2.00, 2.00, 2.50, 1.50, 2.50, 3.17, 1.83, 3.50, 2.67, 2.83, 2.33, 2.33, 3.00, 2.67, 2.17, 1.67, 3.00	2.1448 + 0.0094x	0.04	2.34	6
41	2.38, 2.38, 2.25, 2.00, 1.50, 2.13, 2.25, 2.13, 2.63, 2.25, 2.88, 2.50, 1.75, 1.50, 2.13, 2.63, 2.38, 2.00, 2.00, 2.75, 2.50, 2.13, 2.25, 1.63, 2.25, 2.13, 2.25, 2.00, 2.63, 2.75, 1.88, 2.50, 2.50, 2.00, 2.38, 2.13, 2.13, 1.88, 2.13, 2.63, 2.25	2.1965 + 0.0015x	0.00	2.23	8
42	1.50, 1.80, 2.10, 2.00, 2.90, 2.50, 2.10, 1.80, 2.10, 2.90, 2.50, 1.90, 1.60, 2.30, 3.30, 2.20, 2.70, 2.30, 3.00, 2.00, 2.40, 2.00, 2.40, 2.00, 2.70, 1.90, 2.10, 2.30, 2.40, 3.00, 2.20, 2.20, 1.80, 2.50, 1.90, 2.40, 2.00, 2.20, 2.60, 1.60, 2.00, 3.10	2.2050 + 0.0029x	0.01	2.27	10
43	2.67, 1.67, 1.83, 2.67, 2.33, 1.75, 1.83, 3.00, 1.83, 2.50, 2.08, 2.67, 2.25, 3.08, 1.83, 2.42, 1.92, 2.08, 1.67, 2.08, 1.83, 2.00, 2.17, 2.25, 2.92, 2.58, 2.08, 1.83, 2.42, 2.42, 2.17, 2.17, 1.67, 2.67, 2.42, 1.92, 2.75, 2.25, 1.92, 2.42, 1.92, 2.17, 2.75	2.1940 + 0.0016x	0.00	2.23	12

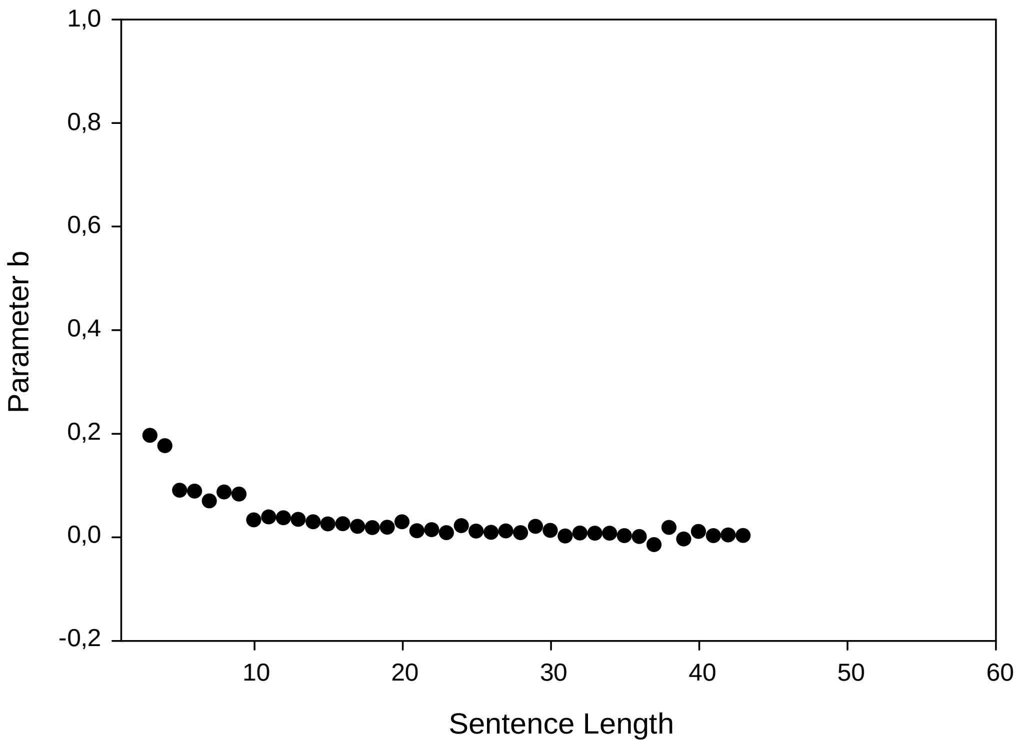

Figure 10. The decrease of *b* in Karinthy's Hungarian text

2.4. Latin

In order to check the hypothesis in a more extreme linguistic situation, we next analyze a Latin text by Seneca (ca. 1-65), *Epistularum moralium ad Lucilium liber primus*, summing up to 6226 words (in 365 sentences). Latin is a strongly inflecting language; therefore one might expect that the results strongly differ from the ones observed above for languages which are synthetic to a lesser degree. The results are presented in Table 7.

Table 7
Mean lengths (L) of words in individual positions (x) in sentences of length n in the Latin text by Seneca

n	**Mean syllabic lengths of words**	**Linear relation** $L = a + bx$	R^2	***M***	***k***
3	2.5, 2.5, 2.0	-	-	-	-
4	1.4, 2.3, 2.6, 2.9	1.1600 + 0.4580x	0.91	2.305	9
5	1.9, 1.9, 1.7, 2.3, 2.6	1.5340 + 0.1780x	0.60	2.068	9
6	2.2, 2.2, 2.5, 2.5, 2.9, 2.9	1.9693 + 0.1554x	0.96	2.513	13
7	2.2, 1.9, 2.5, 2.4, 2.5, 2.7, 2.8	1.9557 + 0.1157x	0.75	2.419	12
8	2.2, 2.1, 2.3, 2.7, 2.1, 1.9, 2.7, 2.5	2.1232 + 0.0418x	0.12	2.311	14
9	2.2, 2.3, 2.0, 2.1, 2.0, 2.5, 3.0, 2.6, 2.3	2.0036 + 0.0622x	0.27	2.314	12
10	1.8, 1.9, 1.9, 2.1, 2.1, 2.3, 2.1, 2.7, 2.6, 2.8	1.6340 + 0.1078x	0.89	2.227	16

11	2.3, 2.2, 2.1, 1.9, 2.3, 1.9, 2.5, 1.8, 2.3, 2.4, 2.8	2.0373 + 0.0300x	0.13	2.217	17
12	2.0, 1.9, 1.7, 2.1, 2.7, 2.0, 3.1, 2.9, 2.5, 2.3, 2.6, 2.7	1.8786 + 0.0752x	0.37	2.368	15
13	2.2, 2.3, 2.8, 2.4, 2.2, 2.1, 2.4, 2.3, 2.5, 2.1, 2.2, 3.1, 2.7	2.2100 + 0.0259x	0.12	2.392	12
14	2.3, 2.4, 2.6, 2.3, 2.1, 2.2, 2.5, 2.5, 2.3, 2.1, 2.6, 2.3, 2.4, 2.6	2.3186 + 0.0077x	0.04	2.376	18
15	2.2, 2.4, 2.2, 2.0, 2.2, 2.3, 2.8, 2.4, 2.5, 2.4, 2.2, 2.2, 2.6, 2.4, 2.85	2.1176 + 0.0282x	0.29	2.343	20
16	1.6, 2.0, 2.0, 2.2, 2.8, 2.4, 2.2, 1.9, 2.0, 2.4, 1.8, 2.3, 2.7, 1.7, 2.89, 2.6,	1.9493 + 0.0313x	0.14	2.216	9
17	1.9, 1.7, 2.6, 2.4, 2.6, 2.2, 2.2, 2.4, 2.5, 1.9, 2.4, 2.4, 2.5, 2.0, 2.72, 2.5, 2.7	2.0782 + 0.0284x	0.22	2.334	25
18	2.0, 1.8, 2.2, 2.4, 2.3, 2.3, 2.5, 2.7, 1.8, 2.0, 3.1, 2.4, 2.2, 2.1, 2.42, 1.8, 2.3, 2.3	2.1841 + 0.0069x	0.01	2.25	12
19	2.0, 2.3, 2.0, 2.5, 2.4, 2.4, 2.0, 2.1, 2.2, 2.4, 2.3, 2.8, 2.3, 2.3, 2.42, 2.4, 2.6, 2.5, 2.8	2.0960 + 0.0251x	0.35	2.347	12
20	1.5, 1.8, 2.1, 2.6, 2.5, 2.8, 2.2, 2.1, 2.9, 2.3, 1.8, 2.4, 2.9, 2.0, 2.58, 2.4, 2.4, 2.4, 2.4, 2.8	2.0602 + 0.0272x	0.17	2.346	12
21	2.5, 2.1, 1.8, 2.3, 2.2, 2.7, 1.9, 2.1, 2.1, 2.0, 2.9, 3.2, 2.3, 2.6, 2.08, 2.5, 1.7, 2.5, 2.3, 2.9, 2.9	2.1118 + 0.0220x	0.11	2.353	12
22	1.9, 2.0, 2.2, 2.6, 2.0, 2.3, 2.7, 2.5, 2.3, 2.1, 2.3, 2.3, 2.9, 2.1, 2.92, 2.0, 2.9, 2.3, 1.8, 2.0, 2.2, 2.3	2.2673 + 0.0025x	0.00	2.295	12
23	1.9, 2.6, 2.7, 1.9, 1.9, 2.5, 2.6, 3.1, 2.6, 2.1, 2.3, 2.6, 2.6, 2.2, 2.6, 1.9, 1.7, 2.3, 2.6, 2.4, 2.6, 2.7, 3.4	2.2516 + 0.0147x	0.06	2.428	11
24	2.8, 2.3, 2.8, 1.8, 2.7, 1.8, 1.8, 1.8, 2.8, 1.8, 2.7, 2.7, 2.0, 1.3, 2.5, 1.8, 3.3, 2.3, 2.2, 1.8, 1.8, 2.3, 2.5, 2.3,	2.3433 - 0.0065x	0.01	2.262	6
25	2.0, 2.4, 2.1, 2.1, 1.5, 1.8, 2.1, 1.9, 2.6, 2.2, 2.4, 1.9, 2.3, 1.9, 2.1, 1.8, 2.1, 2.4, 2.0, 2.2, 1.7, 2.1, 2.2, 2.1, 2.0	2.0740 + 0.0002x	0.00	2.076	10

The course of b is shown in Figure 11. As can be seen, b decreases again, b = 0.8698 exp(-0.1786n) with R^2 = 0.75. The reliable interval of Seneca's text (i.e., $k \geq 10$) is relatively short ($6 \leq n \leq 25$) for this text.

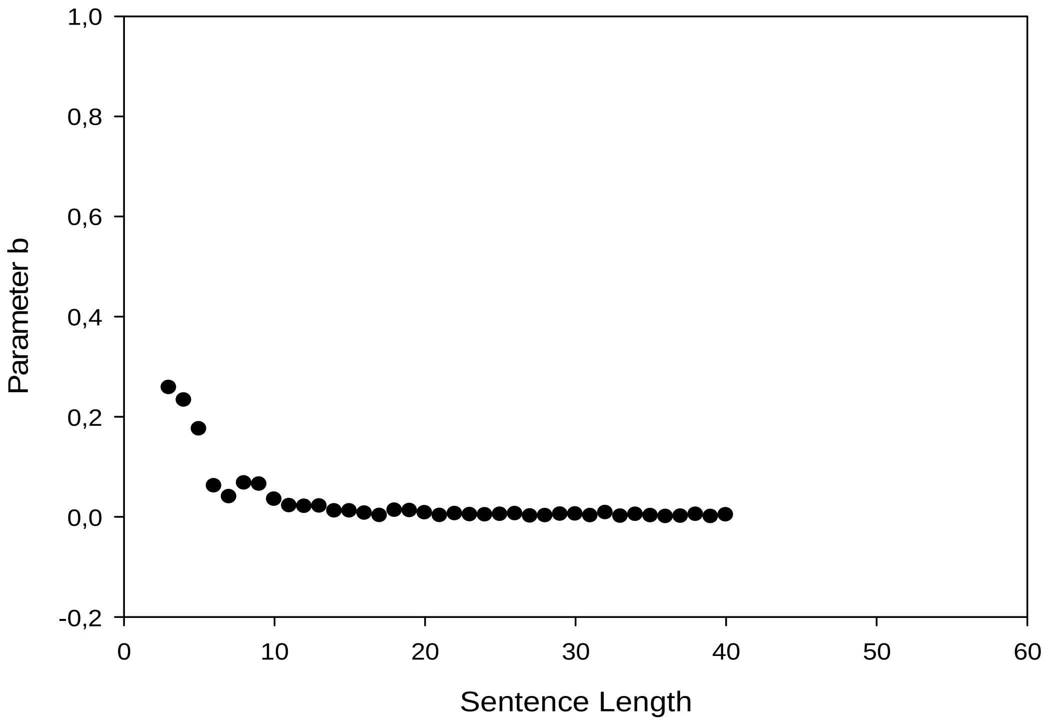

Figure 11. Decrease of *b* in Seneca's Latin text

2.5. English

Let us now consider English, a language which, from an evolutionary point of view, can be regarded as to be increasingly analytic. For our analysis, journalistic texts were combined to a corpus of 2268 sentences and 43320 words, and submitted to analysis. The results obtained are presented in Table 8. We considered the data for $3 \le n \le 40$.

Table 8
Mean lengths (*L*) of words in individual positions (*x*) in sentences of length *n* in English press texts (*M* = mean of all positions, *k* = number of sentences)

n	**Mean syllabic lengths of words**	**Linear relation** ***L = a + bx***	R^2	***M***	***k***
3	1.18, 1.48, 1.70	0.9393 + 0.2576x	0.99	1.45	33
4	1.17, 1.25, 1.28, 1.94	0.8282 + 0.2328x	0.72	1.41	64
5	1.19, 1.48, 1.50, 1.73, 1.94	1.0417 + 0.1750x	0.96	1.57	48
6	1.34, 1.46, 1.34, 1.50, 1.58, 1.66	1.2648 + 0.0614x	0.78	1.48	74
7	1.62, 1.39, 1.42, 1.62, 1.37, 1.50, 1.93	1.3928 + 0.0395x	0.19	1.55	76
8	1.23, 1.52, 1.44, 1.49, 1.59, 1.54, 1.46, 2.01	1.2334 + 0.0671x	0.55	1.54	71
9	1.19, 1.44, 1.50, 1.33, 1.66, 1.52, 1.61, 1.53, 1.98	1.2048 + 0.0646x	0.63	1.53	64
10	1.49, 1.48, 1.39, 1.47, 1.52, 1.60, 1.36, 1.57, 1.71, 1.87	1.3567 + 0.0346x	0.47	1.55	77
11	1.34, 1.36, 1.37, 1.45, 1.56, 1.46, 1.53,	1.3329 + 0.0218x	0.30	1.46	107

	1.50, 1.32, 1.43, 1.79				
12	1.40, 1.46, 1.51, 1.52, 1.56, 1.50, 1.39, 1.60, 1.51, 1.47, 1.57, 1.88	1.3980 + 0.0204x	0.35	1.53	104
13	1.44, 1.54, 1.63, 1.63, 1.57, 1.61, 1.61, 1.58, 1.69, 1.69, 1.63, 1.64, 1.93	1.4829 + 0.0211x	0.55	1.63	72
14	1.60, 1.76, 1.50, 1.58, 1.56, 1.39, 1.53, 1.69, 1.52, 1.56, 1.53, 1.60, 1.68, 1.97	1.5222 + 0.0110x	0.11	1.60	62
15	1.24, 1.50, 1.47, 1.58, 1.50, 1.59, 1.46, 1.46, 1.50, 1.59, 1.35, 1.38, 1.39, 1.49, 1.93	1.4072 + 0.0111x	0.11	1.50	74
16	1.37, 1.54, 1.73, 1.56, 1.64, 1.49, 1.47, 1.56, 1.54, 1.40, 1.29, 1.63, 1.39, 1.49, 1.70, 1.91	1.4857 + 0.0068x	0.04	1.54	70
17	1.45, 1.62, 1.64, 1.66, 1.52, 1.70, 1.51, 1.74, 1.55, 1.68, 1.51, 1.61, 1.47, 1.55, 1.47, 1.44, 1.98	1.5737 + 0.0021x	0.01	1.59	87
18	1.38, 1.64, 1.55, 1.58, 1.58, 1.55, 1.64, 1.55, 1.55, 1.58, 1.53, 1.54, 1.53, 1.49, 1.66, 1.57, 1.64, 2.13	1.4747 + 0.0126x	0.20	1.59	76
19	1.33, 1.50, 1.45, 1.63, 1.43, 1.41, 1.53, 1.46, 1.41, 1.58, 1.61, 1.68, 1.55, 1.66, 1.47, 1.67, 1.41, 1.46, 1.91	1.4150 + 0.0119x	0.24	1.53	76
20	1.39, 1.61, 1.62, 1.69, 1.71, 1.55, 1.50, 1.43, 1.68, 1.64, 1.50, 1.42, 1.57, 1.70, 1.40, 1.50, 1.48, 1.88, 1.62, 1.98	1.5146 + 0.0075x	0.08	1.59	84
21	1.33, 1.60, 1.65, 1.54, 1.63, 1.39, 1.55, 1.44, 1.49, 1.54, 1.43, 1.56, 1.49, 1.69, 1.63, 1.46, 1.56, 1.51, 1.29, 1.45, 1.83	1.5014 + 0.0021x	0.01	1.53	80
22	1.73, 1.29, 1.52, 1.68, 1.48, 1.64, 1.68, 1.41, 1.59, 1.38, 1.65, 1.77, 1.53, 1.45, 1.70, 1.45, 1.85, 1.55, 1.80, 1.42, 1.35, 2.00	1.5217 + 0.0057x	0.04	1.59	66
23	1.47, 1.41, 1.33, 1.47, 1.57, 1.47, 1.51, 1.35, 1.92, 1.61, 1.90, 1.35, 1.55, 1.71, 1.55, 1.61, 1.35, 1.35, 1.73, 1.37, 1.37, 1.59, 1.76	1.4875 + 0.0039x	0.02	1.53	49
24	1.36, 1.52, 1.43, 1.71, 1.57, 1.45, 1.62, 1.64, 1.69, 1.67, 1.53, 1.41, 1.71, 1.47, 1.59, 1.66, 1.52, 1.26, 1.57, 1.50, 1.55, 1.57, 1.50, 1.95	1.5162 + 0.0035x	0.03	1.56	58
25	1.47, 1.44, 1.65, 1.38, 1.62, 1.58, 1.58, 1.67, 1.55, 1.75, 1.25, 1.56, 1.67, 1.69, 1.71, 1.73, 1.78, 1.35, 1.47, 1.42, 1.64, 1.45, 1.65, 1.84, 1.65	1.5238 + 0.0045x	0.05	1.58	55
26	1.28, 1.48, 1.60, 1.60, 1.51, 1.52, 1.61, 1.52, 1.63, 1.54, 1.93, 1.52, 1.58, 1.58, 1.64, 1.51, 1.42, 1.82, 1.42, 1.67, 1.51, 1.70, 1.60, 1.48, 1.52, 1.94	1.5039 + 0.0058x	0.09	1.58	67
27	1.55, 1.61, 1.78, 1.61, 1.43, 1.57, 1.84,	1.5907 + 0.0011x	0.00	1.61	51

	1.71, 1.67, 1.39, 1.69, 1.47, 1.51, 1.82, 1.49, 1.59, 1.51, 1.53, 1.45, 1.61, 1.49, 1.67, 1.61, 1.69, 1.39, 1.69, 2.00				
28	1.51, 1.35, 1.60, 1.60, 1.42, 1.49, 1.60, 1.81, 1.84, 1.47, 1.67, 1.77, 1.60, 1.58, 1.53, 1.65, 1.23, 1.60, 1.47, 1.51, 1.56, 1.65, 1.51, 1.53, 1.53, 1.56, 1.60, 1.84	1.5519 + 0.0016x	0.01	1.58	43
29	1.32, 1.50, 1.80, 1.52, 1.64, 1.43, 1.61, 1.70, 1.39, 1.77, 1.64, 1.61, 1.50, 1.41, 1.68, 1.34, 1.64, 1.66, 1.59, 1.57, 1.66, 1.57, 1.66, 1.55, 1.48, 1.75, 1.48, 1.86, 1.77	1.5200 + 0.0046x	0.08	1.59	44
30	1.36, 1.57, 1.55, 1.50, 1.60, 1.50, 1.50, 1.50, 1.33, 1.52, 1.86, 1.48, 1.60, 1.52, 1.45, 1.62, 1.79, 1.74, 1.57, 1.43, 1.71, 1.45, 1.79, 1.74, 1.57, 1.36, 1.71, 1.45, 1.45, 1.95	1.4915 + 0.0052x	0.09	1.57	42
31	1.43, 1.60, 1.40, 1.45, 1.50, 1.85, 1.63, 1.55, 1.48, 1.75, 1.53, 1.40, 1.68, 1.65, 1.55, 1.58, 1.73, 1.50, 1.70, 1.63, 1.55, 1.53, 1.70, 1.63, 1.53, 1.48, 1.65, 1.43, 1.38, 1.63, 1.83	1.5466 + 0.0018x	0.02	1.58	40
32	1.17, 1.61, 1.33, 1.42, 1.42, 1.64, 1.47, 1.47, 2.00, 1.58, 1.42, 1.28, 1.47, 1.44, 1.44, 1.53, 1.56, 1.69, 1.47, 1.72, 1.56, 1.69, 1.61, 1.58, 1.86, 1.61, 1.53, 1.39, 1.83, 1.33, 1.50, 2.03	1.4236 + 0.0078x	0.14	1.55	36
33	1.19, 1.65, 1.73, 1.46, 1.62, 1.54, 1.73, 1.62, 1.62, 1.65, 1.77, 1.31, 1.81, 1.46, 1.50, 1.96, 1.62, 1.69, 1.42, 1.42, 2.04, 1.31, 1.54, 1.31, 1.50, 1.69, 1.42, 1.65, 1.54, 1.85, 1.15, 1.50, 2.08	1.5727 + 0.0008x	0.00	1.59	26
34	1.40, 1.40, 1.65, 1.60, 1.30, 1.40, 1.30, 1.50, 1.55, 1.50, 1.45, 1.25, 1.65, 1.55, 1.55, 1.35, 1.55, 1.90, 1.85, 1.50, 1.85, 1.55, 1.40, 1.50, 1.65, 1.35, 1.75, 1.55, 1.35, 1.85, 1.45, 1.40, 1.60, 1.65	1.4529 + 0.0045x	0.07	1.53	20
35	1.73, 1.57, 1.63, 1.47, 1.37, 1.70, 1.70, 1.50, 1.47, 1.47, 1.63, 1.63, 1.50, 1.47, 1.93, 1.77, 1.53, 1.43, 1.67, 1.70, 1.30, 1.87, 1.53, 1.50, 1.53, 1.50, 1.27, 1.80, 1.57, 1.50, 1.80, 1.33, 1.53, 1.67, 2.10	1.5599 + 0.0017x	0.01	1.59	30
36	1.54, 1.85, 1.54, 1.58, 1.46, 1.69, 1.73, 1.42, 1.69, 1.85, 1.46, 1.35, 1.96, 1.46, 1.46, 1.85, 1.38, 1.77, 1.46, 1.69, 1.54, 1.35, 1.38, 1.46, 1.65, 2.15, 1.62, 1.50, 1.46, 1.85, 1.23, 1.54, 1.54, 1.46, 1.73, 1.92	1.6017 - 0.0001x	0.00	1.60	26
37	1.53, 1.63, 1.32, 1.21, 1.68, 1.79, 1.63,	1.4969 + 0.0006x	0.00	1.51	19

	1.32, 1.42, 1.89, 1.37, 1.74, 1.68, 1.37, 1.53, 1.68, 1.26, 1.16, 1.32, 1.63, 1.47, 1.21, 1.37, 1.58, 1.37, 1.53, 2.00, 1.21, 1.26, 1.58, 1.26, 1.58, 1.95, 1.79, 1.26, 1.26, 1.95				
38	1.41, 1.06, 1.35, 1.41, 1.47, 1.59, 1.47, 1.76, 1.59, 1.35, 1.94, 1.24, 1.82, 1.65, 1.53, 1.59, 1.76, 1.76, 1.35, 1.35, 1.35, 1.53, 1.71, 1.76, 1.59, 1.47, 1.71, 1.47, 1.65, 1.41, 1.47, 1.24, 1.53, 1.71, 1.88, 1.06, 1.82, 1.94	1.4610 + 0.0044x	0.05	1.55	17
39	1.33, 2.17, 1.58, 1.58, 1.92, 1.08, 1.42, 2.17, 1.67, 1.58, 1.67, 1.42, 1.67, 2.08, 1.08, 1.75, 1.75, 1.50, 1.67, 1.75, 1.58, 1.25, 2.00, 1.58, 2.08, 1.67, 1.58, 1.08, 1.50, 1.83, 1.33, 1.75, 1.50, 1.42, 1.92, 1.58, 1.75, 1.75, 1.75	1.6343 + 0.0000x	0.00	1.63	12
40	1.53, 1.60, 1.33, 1.33, 1.80, 1.60, 1.87, 1.80, 1.20, 1.40, 1.67, 1.93, 1.60, 1.40, 1.80, 1.47, 1.73, 1.53, 1.20, 1.33, 1.40, 1.67, 1.47, 1.80, 1.67, 1.53, 1.53, 1.67, 1.33, 1.27, 2.07, 1.67, 2.00, 1.20, 1.47, 1.40, 1.60, 1.93, 1.40, 2.40	1.5187 + 0.0035x	0.02	1.59	15
41	1.78, 1.44, 1.11, 2.22, 1.89, 1.00, 1.56, 1.56, 1.67, 1.56, 1.33, 1.67, 1.78, 1.22, 1.89, 1.33, 1.78, 1.67, 1.33, 1.67, 1.56, 1.78, 1.33, 1.33, 1.11, 1.44, 1.22, 1.44, 1.22, 1.22, 1.78, 1.11, 2.22, 1.56, 1.11, 1.67, 1.44, 1.67, 1.22, 1.89, 1.67	1.5678 - 0.0021x	0.01	1.52	9
42	1.42, 1.75, 1.67, 1.33, 1.67, 1.33, 2.00, 1.83, 1.58, 1.83, 1.33, 1.67, 1.58, 1.50, 1.42, 1.50, 1.33, 1.42, 1.33, 1.58, 1.83, 1.42, 1.42, 1.67, 1.83, 1.67, 1.25, 1.50, 1.58, 1.33, 1.33, 1.33, 1.67, 1.83, 1.33, 1.58, 1.25, 1.67, 1.25, 1.75, 1.75, 2.00	1.5660 - 0.0005x	0.00	1.56	12
43	1.27, 1.73, 1.36, 1.18, 1.82, 1.64, 1.82, 1.64, 1.18, 1.73, 1.91, 2.09, 1.82, 1.64, 1.82, 1.55, 1.73, 1.64, 1.45, 1.64, 1.09, 1.82, 2.00, 1.09, 1.45, 1.27, 1.45, 1.27, 1.36, 1.91, 1.18, 1.82, 1.55, 1.64, 1.55, 2.09, 1.45, 2.09, 1.82, 1.45, 1.00, 1.55, 2.55	1.5575 + 0.0022x	0.01	1.61	11
44	1.70, 1.50, 1.30, 1.50, 1.60, 2.00, 1.70, 1.80, 1.60, 1.50, 1.30, 1.80, 2.00, 1.90, 1.30, 1.70, 1.50, 1.90, 1.40, 1.20, 1.20, 1.50, 1.50, 1.30, 1.90, 1.20, 1.60, 1.60, 1.20, 1.90, 1.90, 1.90, 1.40, 1.50, 1.80, 1.07, 1.30, 1.60, 1.40, 1.30, 1.20, 1.50, 1.40, 2.30	1.6094 - 0.0015x	0.01	1.58	10

The course of *b* is illustrated in Figure 12: In English, the decrease of parameter *b* can be expressed by the function $b = 0.6980 \exp(-0.3118n)$, with $R^2 = 0.95$.

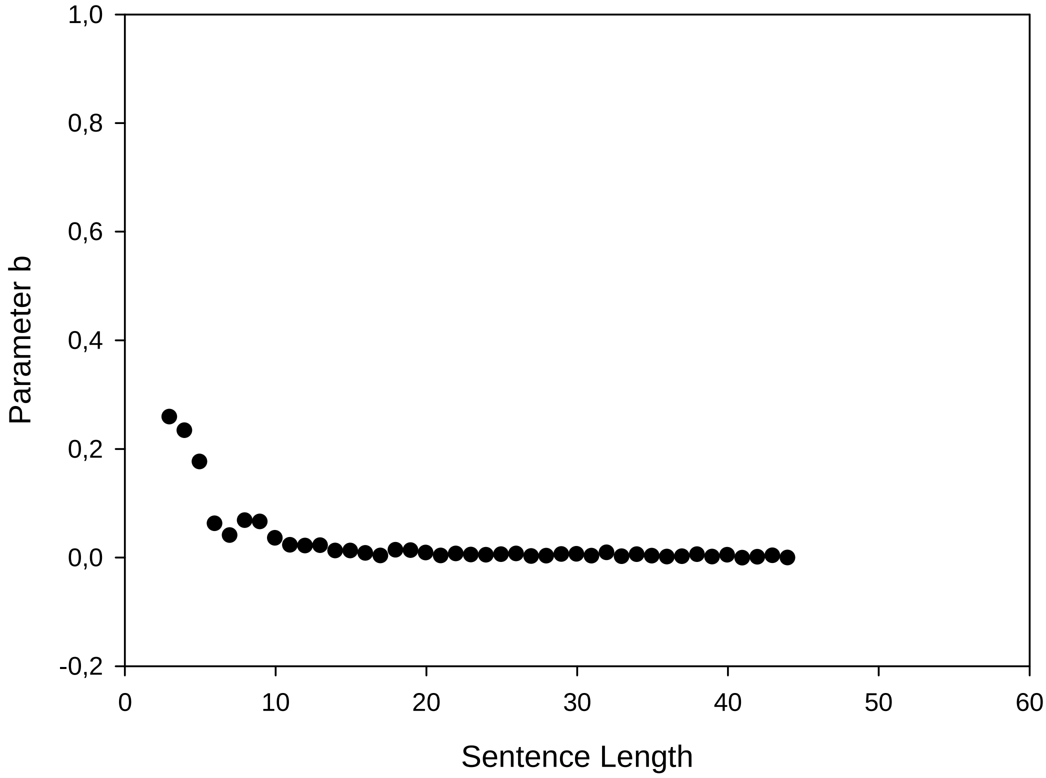

Figure 12. Decrease of *b* in a corpus of English journalistic texts

2.6. Indonesian

The last language analyzed here is a longer Indonesian text containing the well-known fairy tale *Burung api* [= *The fire bird*] by Pak Ojik (1971); this text consists of 409 sentences (4260 words). Indonesian is a language with a low degree of inflection and a higher degree of composition and derivation. Since the text is a fairy tale, the sentences are not too long; yet, we obtained reliable results for the sentence length interval $3 \leq n \leq 17$. The results are presented in Table 9 and the decrease of *b* in Figure 13.

Table 9
Mean lengths (*L*) of words in individual positions (*x*) in sentences of length *n*
in the Indonesian text *Burung api* - (*M* = mean of all positions, *k* = number of sentences)

n	**Mean syllabic lengths of words *L***	**Linear relation *L = a + bx***	R^2	*M*	*k*
3	2.76, 3.12, 3.35	2.4867 + 0.2950x	0.98	3.08	17
4	2.56, 2.28, 2.67, 2.78	2.3100 + 0.1050x	0.40	2.57	18
5	2.57, 2.62, 2.71, 2.57, 3.00	2.4510 + 0.0810x	0.50	2.69	21
6	2.53, 2.53, 2.22, 2.22, 2.53, 2.72	2.3633 + 0.0271x	0.07	2.46	32
7	2.32, 2.68, 2.53, 2.32, 2.62, 2.38, 2.88	2.3657 + 0.0418x	0.19	2.53	34

8	2.44, 2.59, 2.30, 2.41, 2.22, 2.52, 2.52, 2.85	2.3211 + 0.0356x	0.21	2.48	27
9	2.50,2.33,2.30,2.43,2.57,2.47, 2.47, 2.37, 2.53	2.3894 + 0.0103x	0.10	2.44	30
10	2.67, 2.67, 2.27, 2.36, 2.58, 2.21, 2.18, 2.27, 2.52, 3.06	2.4273 + 0.0094x	0.01	2.48	33
11	2.45, 2.64, 2.21, 2.42, 2.64, 2.70, 2.42, 2.36, 2.42, 2.48, 2.48	2.4855 - 0.0018x	0.00	2.47	33
12	2.26, 2.57, 2.61, 2.26, 2.52, 2.39, 2.26, 2.65, 2.48, 2.52, 2.17, 2.96	2.3611 + 0.0169x	0.08	2.47	23
13	2.85, 2.50, 2.74, 2.47, 2.85, 2.65, 2.59, 2.50, 2.12, 2.62, 2.29, 2.71, 2.94	2.6550 - 0.0075x	0.02	2.60	34
14	2.43, 2.64, 2.29, 2.36, 2.29, 3.00, 2.50, 2.71, 2.29, 2.36, 2.07, 2.21, 2.29, 2.64	2.5240 - 0.0120x	0.04	2.43	14
15	2.08, 2.92, 3.25, 2.17, 2.58, 2.92, 2.42, 2.25, 2.58, 2.50, 2.58, 2.42, 2.25, 3.00, 2.83	2.5533 + 0.0038x	0.00	2.58	12
16	2.67, 2.56, 2.22, 2.11, 2.56, 2.78, 3.22, 3.22, 2.67, 2.11, 2.78, 2.33, 2.89, 1.89, 2.67, 3.00	2.5515 + 0.0063x	0.01	2.61	9
17	2.47, 2.60, 2.47, 2.40, 2.47, 3.00, 2.47, 2.33, 2.93,2.67,2.00,2.33,2.73,2.53, 2.07, 2.53, 2.47	2.5823 - 0.0093x	0.03	2.50	15

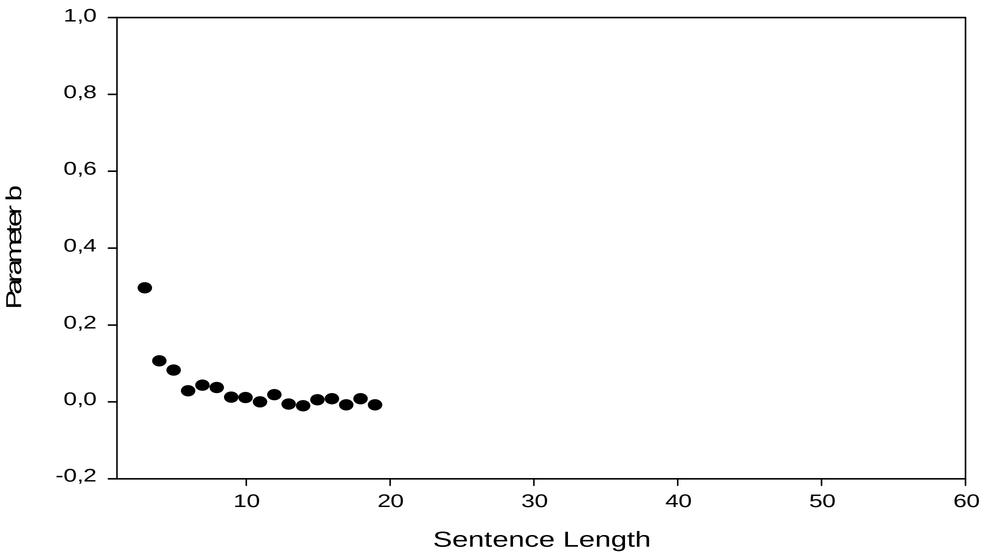

Figure 13. The decrease of parameter *b* in the Indonesian text *Burung api*

Though the number of well represented lengths is rather small, parameter *b* clearly decreases exponentially with increasing sentence length, just as in all other cases above. Here we obtain $b = 2.5394 \exp(-0.7295n)$ with $R^2 = 0.97$.

3. Conclusions

Summarizing the results from all texts and languages, we can state that a common tendency can be observed across (typologically different) languages. According to this tendency, word length is not stable in the course of a sentence, but changes; specifically, there seems to be an increase from the beginning of a sentence to its end, the degree of increase following a straight line (if one ignores possible specifics of first and last words, as was pointed out above). As a result, this increase is slowed down for longer sentences (signalized by decreasing parameter *b*), resulting in some kind of “zero-increase” towards the end of longer sentences (signalized by parameter $b \rightarrow 0$). The overall exponentially decreasing tendency of *b* across the languages studied in this text can be presented lucidly in Table 10. The data are ordered according to the increasing exponent, which is clearly associated with the decreasing multiplicative constant.

Table 10
The decrease of parameter *b* in individual languages

Language	***b***
Slovak (press texts)	25.2594 exp(-1.0298n)
Indonesian (Ojik)	2.5394 exp(-0.7295n)
Latin (Seneca)	2.5420 exp(-0.4556n)
English (press texts)	0.6980 exp(-0.3116n)
Russian (Tolstoj)	0.8737 exp(-0.2763n)
Hungarian (Bródy)	0.7432 exp(-0.2756n)
Slovak (Šikula)	0.4325 exp(-0.1812n)
Hungarian (Karinthy)	0.3061 exp(-0.1778n)
Russian (press texts)	0.2349 exp(-0.1616n)

As can be seen from Table 10, the parameter values of *b* do not seem to represent some constant for a given language; rather, *b* may significantly vary with a given language. For the time being, we do not have any reliable interpretation as to factors influencing *b*. It seems most reasonable to assume that the coefficients in the exponential function are related not only to characteristics of the given language, but also to certain text sort specifics, and by way of this to average word and sentence length, possibly influenced by the number (and type) of clauses constituting the sentences.

Though we do not consider the number of analyzed texts to be sufficient to arrive at general conclusions, hypothesis (I) has been corroborated in all cases: Word length increases towards the end of a sentence; this conclusion must be specified, however, in one important aspect: the longer the sentence, the more this increase is being braked (i.e., the slower is the increase); in very long sentences it can even turn negative, although it is as well possible that values of $b > 0$ do not significantly differ from zero. The slight oscillation in very long sentences may as well indicate the loss of control of word length. In any case, we can conjecture that the change of *b* according to the change of *n*, i.e. *db/dn* is slowed down to the extent almost achieved by *b*. Re-writing this assumption in a formal way, we obtain

(1) $$\frac{db}{dn} = -cb.$$

Solving this elementary differential equation we obtain

(2) $$b = ae^{-cn}$$

where *a* is the integration constant. In fact, function (2) has turned out to be adequate in the above analyses; the numerical values of fitting it to individual samples are presented in Table 10.

Since formula (2) can be also be written as $b = aq^n$ (writing $e^{-c} = q$), which represents a geometric sequence, we may also conclude that the increasing tendency of word length diminishes geometrically as sentence length increases. However, the fact that frequently the first word is longer than the second, and the last word longer than the last but one suggests that the mechanism controlling word length in sentence is not that simple as presented above. There are surely other factors influencing the smooth course of the regression. We consider them as boundary conditions which must be studied for each case separately.

Our data suggest that the phenomenon observed is neither language specific nor restricted to particular genres; most probably, we are concerned with a latent cross-linguistic mechanism. Its genesis can be directly derived from the unified theory of linguistic laws (cf. Wimmer, Altmann 2005) as shown above. Thus we can consider it as a candidate for a text law. However, more texts in more languages must be scrutinized in order to find all the boundary conditions determining the values of the parameters *a* and *c* in (2).

References

Ahrens, H. (1965). *Verborgene Ordnung: Die Beziehung zwischen Satzlänge und Wortlänge in deutscher Erzählprosa von Barock bis heute.* Düsseldorf: Schwann.

Altmann, G. (1983). H. Arens´ "Verborgene Ordnung" und das Menzerathsche Gesetz. In: Faust, M., et al. (eds.), *Allgemeine Sprachwissenschaft, Sprachtypologie und Textlinguistik: 31-39*. Tübingen: Narr.

Grzybek, P., Kelih, E., Stadlober, E. (2007). Long sentences, long words – short sentences, long words? *Presentation at the 31. Jahrestagung der Gesellschaft für Klassifikation: "Data Analysis, Machine Learning, and Application"*. (Freiburg, Germany 2007.)

Grzybek, P., Kelih, E., Stadlober, E. (2008). The relation between word length and sentence length: an intra-systemic perspective in the core data structure. *Glottometrics 16, 111-121.*

Grzybek, P., **Stadlober, E.** (2007). Do we have problems with Arens´ law? A new look at the sentence-word relation. In: Grzybek, P., Köhler, R. (eds.), *Exact Methods in the Study of Language and Text: 205-218*. Berlin: de Gruyter.

Grzybek, P., Stadlober, E., Kelih, E. (2007). The relationship of word length and sentence length: the inter-textual perspective. In: R. Decker and H.-J. Lenz (eds.), *Advances in Data Analysis: 611-618*. Berlin: Springer.

Grzybek, P., Stadlober, E., Kelih, E., Antić, G. (2005). Quantitative text typology: the impact of word length. In: C. Weihs, and W. Gaul (eds.), *Classification – The Ubiquitous Challenge: 53-64.* Berlin: Springer.

Kelih, E., Grzybek, P., Antić, G., Stadlober, E. (2006). Quantitative text typology: the impact of sentence length. In: M. Spiliopoulou et al. (eds.), *From Data and Information Analysis to Knowledge Engineering: 382-389.* Berlin: Springer.

Niemikorpi, A. (1991). Suomen kielen sanaston dynamiikkan. *Acta Wasaensia 26, Kielitiede 2.* Vaasa: Vaasan yliopisto.

Niemikorpi, A. (1997). Equilibrium of Words in the Finnish Frequency Dictionary. *JQL 4(1-3), 190-196.*

Saukkonen, P. (1994). Main trends and results of quantitative linguistics in Finland. *JQL 1(1), 2-15.*

Strauss, U., Fan, F., Altmann, G. (2008). *Problems in Quantitative Linguistics 1*. Lüdenscheid: RAM:

Uhlířová, L. (1997a). O vztahu mezi délkou slova a jeho polohou ve větě. *Slovo a slovesnost 58, 174-184.*

Uhlířová, L. (1997b). Length vs. Order: Word Length and Clause Length from the Perspective of Word Order. *Journal of Quantitative Linguistics 4(1-3), 266-275.*

Wimmer, G., Altmann, G. (2005). Unified derivation of some linguistic laws. In: Köhler, R., Altmann, G., Piotrowski, R.G. (eds.), *Quantitative Linguistics. An International Handbook: 791-807*. Berlin-New York: de Gruyter.

Texts and their sources

English

1. *A Few Good Shirts*, Jonathan Tepperman, Newsweek, March 27, 2009, www.newsweek.com/2009/03/27/a-**few-good-shirts**.html
2. *All Things Greek: To Hellenic and Back*, Jeremy McCarter, Newsweek, March 18, 2010, http://www.newsweek.com/2010/03/18/all-things-greek-to-hellenic-and- back. html
3. *Calming the Bear*, Barton Biggs, Newsweek, March 6, 2009, www.newsweek. com/2009/03/06/**calming-the-bear**.html
4. *Don't talk to aliens, warns Stephen Hawking*, Jonathan Leake, the Sunday Times, April 25, 2010, http://www.timesonline. co.uk/tol/news/science /space/article 7107207. ece
5. *Dr. No*, Owen Matthews, Newsweek, March 11, 2010, www.newsweek.com/ 2010/03/11/**dr-no**.html
6. *Geoducks: Happy as Clams*, Craig Welch, Smithsonian Magazine, March 2009, http://www.smithsonianmag.com/science-nature/Happy-As-Clams.html
7. *If the Stones Could Speak, Searching for the Meaning of Stonehenge*, Caroline Alexander, National Geographic, June 2008, http://ngm.nationalgeographic.com/ 2008/06/stonehenge/geiger-photography
8. *J. D. Salinger Outlived His Legend*, Malcolm Jones, Newsweek Web Exclusive, January, 27, 2010, http://www.newsweek.com/2010/01/27/j-d-salinger-outlived- his-legend.html
9. *Letter from China, Crazy English*, Evan Osnos, the New Yorker, April 28, 2008, http://www.newyorker.com/reporting/2008/04/28/080428fa_fact_osnos
10. *Mourning the Death of Handwriting*, Claire Suddath, August 03, 2009, Time Magazine, http://www. time.com/time/magazine/article/0,9171,1912419,00.html
11. *One Small Step*, George Alexander, July 17, 2009, Newsweek, www.newsweek. com/2009/07/16/**one-small-step**.html
12. *Palin With a Pedigree*, Michael Isikoff and Michael Hirsh, Newsweek, March 12, 2010, www.newsweek.com/2010/03/11/**palin-with-a-pedigree**.html
13. *Pregnant? Eat Chocolate!* Rob Stein, Washington Post, May 8, 2008, http://voices. washingtonpost.com/checkup/2008/05/bad_news_for_pregnant_women_ea.html
14. *Scanning a Stradivarius*, Erica R. Hendry, May 2010, Smithsonian Magazine, http://www.smithsonianmag.com/arts-culture/Scanning-a-Stradivarius.html
15. *The Case Against Goldman Sachs*, Stephen Gandel, Time Magazine, April 22, 2010, http://www.time.com/time/business/article/0,8599,1983747,00.html
16. *The Greatest Shakespeare Hoax*, Doug Stewart, Smithsonian Magazine, June 2010, http://www.smithsonianmag.com/history-archaeology/The-Greatest-Shakespeare-Hoax.html
17. *The iPad Launch: Can Steve Jobs Do It Again?* Stephen Fry, Time Magazine, April 01, 2010. http://www.time.com/time/business/article/0,8599, 1976935,00.html
18. *The Pill at 50: Sex, Freedom and Paradox*, Nancy Gibbs Thursday, Time Magazine, April 22, 2010, http://www.time.com/time/health/article/0,8599,1983712,00.html

19. *Trading Unit Propels Morgan Stanley to Profit*, Michael J. de la Merced, New York Times, April 21, 2010, http://www.nytimes.com/2010/04/22/business/22 morgan.html
20. *Unlocking Mysteries of the Parthenon*, Evan Hadingham, Smithsonian Magazine, February 2008, http://www.smithsonianmag.com/ history-archaeology/ Unlocking-Mysteries-of-the-Parthenon.html

Latin

Seneca, L.A., *Epistularum moralium ad Lucilium liber primus*
http://www.thelatinlibrary.com/sen/seneca.ep1.shtml (28.7.2010)

Hungarian

Bródy, S., *Az ezüst kecske*. 1898.
Karinthy, F., *Utazás a koponyám körül.* 1937.

Slovak

Journalistic texts from a Slovak corpus
Šikula, V. *Veterná ružica.* 1995.

Russian

44 journalistic texts from the journal Vremja (http://quanta-textdata.uni-graz.at/).
L.N. Tolstoj, *Anna Karenina*. 1877.

Indonesian

Pak Ojik, *Burung Api*. Djakarta: Pustaka Jaya. 1971

Glottometrics 20

2010
RAM-Verlag

Glottometrics

Glottometrics ist eine unregelmäßig erscheinende Zeitschrift für die quantitative Erforschung von Sprache und Text

Beiträge in Deutsch oder Englisch sollten an einen der Herausgeber in einem gängigen Textverarbeitungssystem (vorrangig WORD) geschickt werden

Glottometrics kann aus dem **Internet** heruntergeladen, auf **CD-ROM** (in PDF Format) oder in **Buchform** bestellt werden

Glottometrics is a scientific journal for the quantitative research on language and text published at irregular intervals

Contributions in English or German written with a common text processing system (preferably WORD) should be sent to one of the editors

Glottometrics can be downloaded from the **Internet**, obtained on **CD-ROM** (in PDF) or in form of **printed copies**

Herausgeber – Editors

Bestellungen der CD-ROM oder der gedruckten Form sind zu richten an
Orders for CD-ROM´s or printed copies to

RAM-Verlag RAM-Verlag@t-online.de

Herunterladen / Downloading: http://www.ram-verlag.de

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Glottometrics – 20 (2010) –. – Lüdenscheid: RAM-Verl., 2010
Erscheint unregelmäßig. – Auch im Internet als elektronische Ressource unter der Adresse http://www.ram-verlag.de verfügbar.-
Bibliographische Deskription nach 20 (2010)
ISSN 1617-8351

Glottometrics

Glottometrics ist eine unregelmäßig erscheinende Zeitschrift für die quantitative Erforschung von Sprache und Text

Beiträge in Deutsch oder Englisch sollten an einen der Herausgeber in einem gängigen Textverarbeitungssystem (vorrangig WORD) geschickt werden

Glottometrics kann aus dem **Internet** heruntergeladen, auf **CD-ROM** (in PDF Format) oder in **Buchform** bestellt werden

Glottometrics is a scientific journal for the quantitative research on language and text published at irregular intervals

Contributions in English or German written with a common text processing system (preferably WORD) should be sent to one of the editors

Glottometrics can be downloaded from the **Internet**, obtained on **CD-ROM** (in PDF) or in form of **printed copies**

Herausgeber – Editors

G. Altmann	ram-verlag@t-online.de
K.-H. Best	kbest@gwdg.de
F. Fan	fanfengxiang@yahoo.com
P. Grzybek	peter.grzybek@uni-graz.at
L. Hřebíček	ludek.hrebicek@seznam.cz
R. Köhler	koehler@uni-trier.de
J. Mačutek	jmacutek@yahoo.com
G. Wimmer	wimmer@mat.savba.sk
A. Ziegler	arne.ziegler@uni-graz.at

Bestellungen der CD-ROM oder der gedruckten Form sind zu richten an
Orders for CD-ROM´s or printed copies to

RAM-Verlag RAM-Verlag@t-online.de

Herunterladen / Downloading: http://www.ram-verlag.de

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Glottometrics. –20 (2010-10-136) –. – Lüdenscheid: RAM-Verl., 2010
Erscheint unregelmäßig. – Auch im Internet als elektronische Ressource unter der Adresse http://www.ram-verlag.de verfügbar.-
Bibliographische Deskription nach 20 (2010)
ISSN 1617-8351

Contents