Glottometrics 48, ……..

Lengths and L-motifs of Rhythmical Units in Formal British Speech

Andrew Wilson
Lancaster University

Abstract. The lengths of rhythmical units (as defined by Karl Marbe in 1904) were identified, and their frequencies counted, in twelve complete texts from the Aix-MARSEC database of formal spoken British English. The texts all belonged to the genre of current affairs commentary. L-motifs (i.e. maximal monotone non-decreasing sequences) of the rhythmical unit lengths were also identified, and the frequencies of the different L-motif lengths were counted. The frequencies of both rhythmical unit lengths and L-motif lengths were modelled using a continuous approach with the Zipf-Alekseev function. Good qualities of fit were obtained for both kinds of unit on all texts. The parameters *a* and *b* of the Zipf-Alekseev function for the rhythmical unit lengths (though not for the L-motif lengths) were also found to be related in the form of a further Zipf-Alekseev function. Further research should aim to extend the application of the motif approach to rhythmical units.

Keywords: *Rhythmical units; Motifs; L-motifs; Zipf-Alekseev function*.

Introduction

The study of rhythmical patterns of stresses, in both verse and prose, has a long history in quantitative linguistics (Kelih, 2008). In more recent years, Karl-Heinz Best (e.g. 2002; 2005) has renewed the quest with investigations into the rhythmical units proposed by Karl Marbe (1904). A rhythmical unit, according to Marbe's definition, is the span between one stressed syllable and the next, incorporating the first stressed syllable but not the second; in other words, it is a sequence of a stressed syllable followed by any number of unstressed syllables. Thus, if we represent a stressed syllable by 'S' and an unstressed one by 'U', SU is a rhythmical unit of length 2, SUU is a rhythmical unit of length 3, SUUU is a rhythmical unit of length 4, and so on. When a stressed syllable is followed immediately by another stressed syllable (i.e. SS), then we have a rhythmical unit of length 1, consisting only of the first stressed syllable; the second stressed syllable forms the beginning of the next rhythmical unit.

Marbe's rhythmical units take no account of punctuation or of any other textual or phonetic cues. In this respect, they differ from a very similar unit used in the so-called 'British tradition' of prosodic analysis. In this tradition (e.g. Crystal, 1969: ch.2), the study of stress patterns and the study of pitch movements are conflated. As units of structure, the British tradition recognizes syllables, feet, and intonation units (also known as 'tone groups'). It is the foot that bears the closest similarity to Marbe's rhythmical unit, but, in the British tradition, its identification is affected by the boundaries of another structural unit, namely the intonation unit. Unlike a traditional metrical foot, a foot in the British tradition is defined as the span between either (1) a stressed syllable and the next stressed syllable, or (2) an intonation unit boundary and the next stressed syllable, or (3) a stressed syllable and an intonation unit boundary (Halliday, 1963). Definition (1) gives a unit which is identical to Marbe's rhythmical unit, but definition (2) defines a unit containing no stressed syllable at all, whilst definition (3) curtails the unit short of the second stressed syllable, at the intervening boundary of an intonation unit.

Using a sample of texts from the Aix-MARSEC database of formal British speech, Wilson (2017) made a quantitative assessment of some of the assumptions of the British tradition and found it wanting: the intonation-unit lengths, when measured in feet, did not abide neatly by any discrete probability distribution, and there was also no systematic relationship between foot length and intonation-unit length using the Menzerath-Altmann equation. The latter result is particularly important because it indicates that, even if the criterion for fitting the probability distributions erred on the conservative side (a criterion of $C < 0.02$ was used), the foot is clearly not the correct immediate constituent of an intonation unit. This is contrary to the proposal of Halliday (1967), who claimed that the foot *is* the immediate constituent of an intonation unit. The precise causes of these modelling deficiencies in relation to the Aix-MARSEC data could not be identified with certainty. However, in contrast to these results, studies of Marbe's rhythmical units by Best and others (e.g. Knaus, 2008) have proven to be more fruitful, showing these units to behave in a law-like fashion in complete texts from a number of languages, though not generally in samples from texts. Nevertheless, to date, the amount of research on this topic remains relatively small: the most up-to-date version of Best's online bibliography of quantitative linguistics (July 2019) lists only nine studies on rhythmical units, mostly using German texts, but with individual studies also analysing data from ancient Greek, English, and Russian. This study therefore re-analyses some of the texts used in Wilson (2017) within that framework.

Another recent focus in quantitative linguistics has been on the unit known as a motif. This has its origins in Moisei Boroda's attempt, in the 1970s, to systematize the analysis of music by identifying more objectively defined structural units (see the various studies in Orlov, Boroda & Nadarejšvili, 1982). Thus, in contrast to the well-known subjective concept of a musical motif, Boroda's motifs have a strict mathematical definition. Reinhard Köhler and colleagues later adapted this idea to verbal texts (Köhler & Naumann, 2008), and motifs are now a fruitful area of quantitative research on texts (see, e.g., Altmann, 2016). Motifs deal with linear progressions in texts and aim to capture rhythmical patterns at different linguistic levels. They can be either qualitative or quantitative: a qualitative motif is an uninterrupted linear sequence of unrepeated items (e.g. a string of entirely different parts of speech), whereas a quantitative motif is a maximal monotone non-decreasing sequence of numerical values (e.g. of the lengths of successive words or sentences in a text). Motifs are abstract units, with no grounding in traditional theories of linguistic analysis, but they have so far shown regular and lawful behaviour in texts. However, whilst there have been a number of studies looking at motifs of elements such as word lengths or parts of speech, there have not yet - to the present writer's knowledge - been any studies on motifs of rhythmical units.

The present study thus aims to move the study of rhythmical units forwards along two fronts. First, it adds more English data to the other data sets assembled by Best and his colleagues. Second, it considers motifs of rhythmical unit lengths, as well as the rhythmical unit lengths themselves.

Data and Method

The data for this study were drawn from the Aix-MARSEC database of formal British speech. This is a richly annotated corpus of around 52,000 running words in length, dating from the middle of the 1980s (Taylor & Knowles, 1988; Roach et al., 1993; Knowles, 1994; Auran, Bouzon & Hirst, 2004). Most of the texts contained within Aix-MARSEC are scripted radio broadcasts, though a few are live speech events - also, however, mostly scripted. The majority of the texts are also complete texts, and only complete texts were used in this study. The texts are distributed across several genres, including current affairs commentary, news, weather forecasts, poetry readings, charity appeals, and presentation speeches for honorary degrees. Speakers were selected for inclusion by virtue of their being speakers of British Received Pronunciation (RP). The texts were prosodically annotated, following a version of the British tradition, by two experienced phoneticians: Gerry Knowles (GOK) and Briony Williams (BJW). Based on their perceptions of the sound recordings, they marked up elements such as tonetic stress markers and the boundaries of intonation units. Some texts were annotated by just one phonetician, whilst other, mostly longer, texts were split between the two. A few sub-passages were annotated by both phoneticians, allowing a comparison of their annotation practices (e.g. Knowles, 1991). Other information was added later to the database, including a phonemic transcription of each word (from a dictionary source), the stressed and unstressed syllables in each word, and the time-lengths of various units measured in milliseconds.

A mixing of either speakers or annotators within the same text can lead to heterogeneity in the data, which affects the rhythms and frequency structures (Altmann, 1992). For this study, therefore, only the single genre of current affairs commentary was considered. This genre category is made up of twelve texts in Aix-MARSEC. Each of these texts was spoken by just one speaker, and each entire text annotated by just one phonetician. The identities of the speakers and phoneticians are shown in Table 1. As can be seen, each speaker is represented only once.

Table 1
Text references, speakers, and annotators of the twelve texts from Aix-MARSEC (Taylor & Knowles, 1988).

Text ref.	Speaker	Annotator
A01	Rosemary Hartill	BJW
A02	Gerald Butt	GOK
A03	Jon Silverman	BJW
A04	John Carlin	GOK
A05	James Morgan	GOK
A06	David Smeeton	BJW
A07	Laurie Margolis	GOK
A08	Keith Graves	BJW
A09	Graham Leach	GOK
A10	Alan MacDonald	BJW
A11	Peter Ruff	GOK
A12	Jim Biddulph	BJW

For each text, the linear sequence of stressed and unstressed syllables was extracted from the database and converted into rhythmical units using Marbe's definition. Nine of the twelve texts began with one or more unstressed syllables; these do not strictly form a rhythmical unit according to Marbe's definition, which requires a stressed syllable in first position, but they were included in the counts, as this seems also to have been the practice of Best (2002, p. 138). The lengths of the rhythmical units, in syllables, were then counted using a bespoke program written in Python.

Having counted the lengths of the rhythmical units in each text, the linear sequence of lengths was then further analysed, using another Python program, to identify the quantitative length motifs (hereafter: L-motifs). As stated earlier, these are defined as maximal monotone non-decreasing sequences of the rhythmical-unit length values. Their lengths, in terms of rhythmical units, were also counted.

It was not necessary to undertake a further analysis of qualitative motifs, since all rhythmical units have the same basic structure, apart from the single unit of unstressed syllables at the starts of some texts. Length is therefore the only property that distinguishes, say, SUU from SUUU, or SU, or merely S; we do not have to consider a qualitative difference between rhythmical units of the same length - e.g., SUU, SUS, UUS, etc. - because it will never occur.

The frequencies of length classes can be modelled using either a discrete or a continuous approach (Mačutek & Altmann, 2007). The present study adopted the continuous approach, following broadly the reasoning of Hammerl (1989). To model the frequencies of both the rhythmical unit and L-motif lengths, a start was therefore made from the following differential equation, where *x* is the length of the unit or motif and *y* is the frequency of the length class:

$$\frac{dy}{y} \approx \frac{g(x)}{h(x)}\,dx$$

The function *g(x)* represents the contribution of the speaker, modelled as:

$$g(x) = k + r \ln x$$

and the function *h(x)* represents the contribution of the hearer, modelled as:

$$h(x) = mx$$

Inserting these into the original equation gives:

$$\frac{dy}{y} = \frac{k + r \ln x}{mx}\,dx$$

which, after integration and re-parameterization, results in the well-known Zipf-Alekseev function:

$$y = 1 + cx^{a + b \ln x}$$

The constant of 1 is added to the function, as we can never have a rhythmical unit or L-motif smaller than 1.

The Zipf-Alekseev function was fitted to the data using the *minpack.lm* package in R for Windows version 3.3.3 (Ihaka & Gentleman, 1996; Elzhov et al., 2016). Quality of fit was assessed using the determination coefficient given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(f - f_{pred})^2}{\sum_{i=1}^{N}(f - \bar{f})^2}$$

where f is the observed frequency and f_{pred} is the predicted frequency from the model. A fit was considered good if $R^2 > 0.9$ and still acceptable if $R^2 > 0.7$.

Results

Tables 2 to 13 show the results of fitting the Zipf-Alekseev function to the lengths of the rhythmical units in the twelve texts. Tables 14 to 25 show the results of fitting the Zipf-Alekseev function to the lengths of the L-motifs of rhythmical units in the twelve texts. Table 26 summarizes the parameter estimates in a convenient form.

Table 2
Text A01: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	121	118.0573
2	183	191.3157
3	134	119.6613
4	61	61.4821
5	22	30.5467
6	3	15.5133
c = 117.0573 a = 1.8787 b = -1.6988 R^2 = 0.9791		

Table 3
Text A02: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	98	97.3463
2	156	157.7661
3	90	87.0563
4	41	38.8456
5	11	16.9583
6	4	7.8048
c = 96.3463 a = 2.0787 b = -1.9857 R^2 = 0.9961		

Table 4
Text A03: Lengths of rhythmical units.

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	117	114.3709
2	153	161.9158
3	112	93.0483
4	37	45.0028
5	14	21.4179
6	2	10.6056
7	1	5.6455
c = 113.3709 a = 1.6932 b = -1.7138 R^2 = 0.9726		

Table 5
Text A04: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	176	176.1407
2	263	262.1543
3	130	134.0625
4	64	55.9231
5	22	22.9159
6	1	9.8984
c = 175.1407 a = 1.9893 b = -2.0384 R^2 = 0.9967		

Table 6
Text A05: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	102	100.6929
2	220	222.8524
3	133	126.8200
4	51	53.9363
5	18	21.7528
6	4	9.1320
$c = 99.6929$ $a = 2.7647$ $b = -2.3237$ $R^2 = 0.9971$		

Table 7
Text A06: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	100	99.1688
2	213	214.7343
3	124	120.8524
4	52	51.0557
5	15	20.5192
6	5	8.6172
7	4	4.0479
$c = 98.1688$ $a = 2.7308$ $b = -2.3204$ $R^2 = 0.9984$		

Table 8
Text A07: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	79	74.3252
2	164	173.2291
3	129	113.6517
4	56	56.1165
5	12	25.9909
6	1	12.2337
$c = 73.3252$ $a = 2.6698$ $b = -2.0744$ $R^2 = 0.9678$		

Table 9
Text A08: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	91	87.8551
2	145	154.0312
3	117	98.2172
4	41	50.5822
5	15	25.0585
$c = 86.8551$ $a = 2.0387$ $b = -1.7623$ $R^2 = 0.9443$		

Table 10
Text A09: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	79	74.7098
2	187	194.6206
3	145	132.1287
4	62	65.8654
5	27	30.4314
6	3	14.1701
7	2	6.9710
$c = 73.7098$ $a = 2.8788$ $b = -2.1431$ $R^2 = 0.9862$		

Table 11
Text A10: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	93	86.8228
2	197	209.6366
3	160	134.4915
4	50	64.0979
5	18	28.5556
6	3	12.9286
$c = 85.8228$ $a = 2.785$ $b = -2.169$ $R^2 = 0.959$		

Table 12
Text A11: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	85	78.2799
2	176	189.2477
3	151	127.5925
4	58	64.3508
5	15	30.2800
6	4	14.3836
$c = 77.2799$ $a = 2.7123$ $b = -2.0599$ $R^2 = 0.9535$		

Table 13
Text A12: Lengths of rhythmical units

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	87	86.2866
2	117	119.4741
3	69	63.3344
4	26	28.2033
5	9	12.5496
6	2	5.9939
$c = 85.2866$ $a = 1.7726$ $b = -1.8733$ $R^2 = 0.9933$		

Table 14
Text A01: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	29	31.0469
2	69	63.9156
3	32	42.7544
4	30	22.4100
5	13	11.2802
6	9	5.9087
7	2	3.3820
8	1	2.1841
c = 30.0469 a = 2.3769 b = -1.8909 R^2 = 0.9374		

Table 15
Text A02: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	19	19.0864
2	51	50.8128
3	35	35.4816
4	19	18.2787
5	8	8.9070
6	6	4.5596
7	2	2.6209
8	1	1.7548
c = 18.0864 a = 2.9562 b = -2.1562 R^2 = 0.9979		

Table 16
Text A03: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	26	21.9801
2	44	52.1596
3	53	38.6468
4	15	21.9037
5	6	11.7153
6	8	6.4100
7	2	3.7566
8	0	2.4312
9	0	1.7598
10	0	1.4128
11	1	1.2294
c = 20.9801 a = 2.5746 b = -1.8591 R^2 = 0.8919		

Table 17
Text A04: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	37	37.8308
2	84	82.1141
3	52	55.6249
4	31	29.1133
5	15	14.4861
6	11	7.4186
7	2	4.1008
8	2	2.5334
$c = 36.8308$ $a = 2.4729$ $b = -1.9244$ $R^2 = 0.9931$		

Table 18
Text A05: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	32	29.0329
2	50	56.1445
3	49	42.6779
4	31	26.2497
5	10	15.3991
6	6	9.1345
7	3	5.6396
8	1	3.6911
9	0	2.5914
10	1	1.9598
$c = 28.0329$ $a = 2.0276$ $b = -1.517$ $R^2 = 0.9539$		

Table 19
Text A06: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	29	26.1794
2	60	64.9122
3	54	47.4373
4	28	26.1430
5	7	13.5225
6	5	7.1384
7	0	4.0378
8	2	2.5330
9	1	1.7918
$c = 25.1794$ $a = 2.6887$ $b = -1.9403$ $R^2 = 0.9673$		

Table 20
Text A07: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	27	27.1553
2	60	59.4981
3	37	38.5435
4	21	19.2018
5	10	9.2152
6	4	4.6849
7	1	2.6819
8	0	1.7879
9	2	1.3799
$c = 26.1553$ $a = 2.5841$ $b = -2.0527$ $R^2 = 0.996$		

Table 21
Text A08: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	19	17.4854
2	46	48.5682
3	41	36.6230
4	18	20.3150
5	10	10.5208
6	4	5.5934
7	3	3.2312
8	1	2.1037
$c = 16.4854$ $a = 2.9434$ $b = -2.0408$ $R^2 = 0.982$		

Table 22
Text A09: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	35	35.7620
2	78	76.2291
3	38	42.1886
4	20	17.7593
5	12	7.3711
6	3	3.4269
7	2	1.9492
8	1	1.3843
9	0	1.1614
10	2	1.0703
$c = 34.762$ $a = 2.7538$ $b = -2.3661$ $R^2 = 0.9911$		

Table 23
Text A10: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	28	27.6531
2	63	63.6578
3	46	44.8456
4	23	24.1899
5	14	12.3628
6	6	6.5028
7	2	3.6977
8	1	2.3512
9	1	1.6936
c = 26.6531 a = 2.5668 b = -1.924 R^2 = 0.9971		

Table 24
Text A11: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	32	32.9326
2	76	73.9997
3	37	41.5179
4	20	17.5105
5	10	7.2541
6	7	3.3681
7	2	1.9196
8	1	1.3694
c = 31.9326 a = 2.8615 b = -2.4074 R^2 = 0.9882		

Table 25
Text A12: Lengths of L-motifs

Length	Observed frequency	Predicted (Zipf-Alekseev)
1	12	12.6608
2	42	40.6756
3	25	28.9064
4	21	14.4775
5	3	6.8307
6	3	3.4627
7	2	2.0494
8	1	1.4571
c = 11.6608 a = 3.4287 b = -2.398 R^2 = 0.9503		

Table 26
Summary of parameters a, b, and c.

Rhythmical Units	c	a	b
A01	117.0573	1.8787	-1.6988
A02	96.3463	2.0787	-1.9857
A03	113.3709	1.6932	-1.7138
A04	175.1407	1.9893	-2.0384
A05	99.6929	2.7647	-2.3237
A06	98.1688	2.7308	-2.3204
A07	73.3252	2.6698	-2.0744
A08	86.8551	2.0387	-1.7623
A09	73.7098	2.8788	-2.1431
A10	85.8228	2.785	-2.169
A11	77.2799	2.7123	-2.0599
A12	85.2866	1.7726	-1.8733
L-motifs			
A01	30.0469	2.3769	-1.8909
A02	18.0864	2.9562	-2.1562
A03	20.9801	2.5746	-1.8591
A04	36.8308	2.4729	-1.9244
A05	28.0329	2.0276	-1.517
A06	25.1794	2.6887	-1.9403
A07	26.1553	2.5841	-2.0527
A08	16.4854	2.9434	-2.0408
A09	34.762	2.7538	-2.3661
A10	26.6531	2.5668	-1.924
A11	31.9326	2.8615	-2.4074
A12	11.6608	3.4287	-2.398

Discussion

The fitting of the Zipf-Alekseev function was successful for both the rhythmical unit lengths and the lengths of the L-motifs of rhythmical units. Good fits were obtained in all cases, apart from the case of the L-motifs in text A03; but, even here, the quality of fit was very satisfactory and fell only marginally below the threshold set for a good fit ($R^2 = 0.8919$).

These findings provide further support for the validity of Marbe's (1904) rhythmical units as a model for the rhythmical structure of texts. Perhaps more interestingly, they also suggest that texts may possess higher levels of rhythmical organization than Marbe's rhythmical units alone. This is not to suggest that Marbe's units are constituents of any hypothetical higher-level units: unlike the place of feet in Halliday's (1967) version of the British tradition of prosodic analysis, there has never been any suggestion that Marbe's units form part of a larger constituency hierarchy. Nevertheless, it does seem that writers or speakers are controlling the rhythms of their texts by varying how often they deploy a shorter rhythmical unit after a sequence of equally or increasingly long ones.

As regards the parameter estimates from the present set of formal British English texts, parameter *c* equates roughly, in all cases, to the size of the first length class. (This feature of the

Zipf-Alekseev function is already well known from other research – see, e.g., Koch 2014 for a review.) The other two parameters (*a* and *b*) are then quite similar for both rhythmical unit lengths and L-motif lengths. For rhythmical unit lengths, parameter *a* has a mean of 2.3327 (SD = 0.4576) and parameter *b* has a mean of -2.0136 (SD = 0.2154). In the case of L-motif lengths, parameter *a* has a mean of 2.6863 (SD = 0.3488) and parameter *b* has a mean of -2.0397 (SD = 0.2608). Furthermore, if we consider *b* as a function of *a*, it can be shown that this is, too, a Zipf-Alekseev function given for rhythmical units as:

$$b = -0.2989 * a\text{^}(2.2408 - 0.2854 * \ln a)$$

with $R^2 = 0.9586$. However, for motifs we do not obtain a smooth relation.

Further research should attempt to replicate these results using other texts, especially texts from other languages and other genres, and compare the parameter estimates of the Zipf-Alekseev function. Other properties of the L-motif lengths for rhythmical units might also be investigated, along the lines suggested in Altmann's (2016) programmatic paper.

Acknowledgement

I am grateful to Adrian Tucker for writing the program to identify and count motifs.

Data access statement

The data used in this study forms part of the Aix-MARSEC database, version 2, from the Speech & Language Data Repository (SLDR); http://sldr.org, identifier: sldr000033.

References

Altmann, G. (1992). Das Problem der Datenhomogenität. In B. Rieger (Ed.), *Glottometrika 13* (pp. 287–298). Bochum: Brockmeyer.

Altmann, G. (2016). On Köhlerian motifs. In E. Kelih, R. Knight, J. Mačutek & A. Wilson (Ed.), *Issues in quantitative linguistics 4* (pp. 2-8). Lüdenscheid: RAM-Verlag.

Auran, C., Bouzon, C., & Hirst, D. J. (2004). The Aix-MARSEC project: An evolutive database of spoken English. In B. Bel & I. Marlien (Eds.), *Proceedings of the second international conference on speech prosody* (pp. 561–564). Japan: Nara.

Best, K.-H. (2002). The distribution of rhythmic units in German short prose. *Glottometrics, 3,* 136–142.

Best, K.-H. (2005). Längen rhythmischer Einheiten. In R. Köhler, G. Altmann, & R. Piotrowski (Eds.), *Quantitative Linguistik - Quantitative Linguistics. Ein internationales Handbuch* (pp. 208–214). Berlin: Mouton de Gruyter.

Best, K.-H. (2019). *Bibliographische Übersicht zum Göttinger Projekt zur Quantitativen Linguistik*. Retrieved August 7, 2019, from http://wwwuser.gwdg.de/~kbest/litlist.htm

Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.

Elzhov, T.V., Mullen, K.M., Spiess, A.-N., & Bolker, B. (2016). *minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds.* R package version 1.2-1. https://CRAN.R-project.org/package=minpack.lm

Halliday, M.A.K. (1963). The tones of English. *Archivum Linguisticum, 15,* 1-28.

Halliday, M. A. K. (1967). *Intonation and grammar in British English.* The Hague: Mouton.

Hammerl, R. (1989). Untersuchungen zur Verteilung der Wortarten im Text. In L. Hrebicek (Ed.), *Glottometrika 11* (pp. 142-156). Bochum: Brockmeyer.

Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational Graphics and Statistics, 5,* 299–314.

Kelih, E. (2008). *Geschichte der Anwendung quantitativer Verfahren in der russischen Sprach- und Literaturwissenschaft.* Hamburg: Kovač.

Knaus, M. (2008). Zur Verteilung rhythmischer Einheiten in russischer Prosa. *Glottometrics, 16,* 57-62.

Knowles, G. (1991). Prosodic labelling: The problem of tone group boundaries. In S. Johansson & A.-B. Stenström (Eds.), *English computer corpora: selected papers and research guide* (pp. 149–163). Berlin: Mouton de Gruyter.

Koch, V. (2014). *Quantitative film studies: regularities and interrelations exemplified by shot lengths in Soviet feature films*. 2 vols. Ph.D. dissertation, Karl-Franzens-Universität Graz.

Köhler, R. & Naumann, S. (2008). Quantitative text analysis using L-, F- and T-segments. In B. Preisach & D. Schmidt-Thieme (Ed.), *Data analysis, machine learning and applications* (pp. 637-646). Berlin: Springer.

Mačutek, J. & Altmann, G. (2007). Discrete and continuous modelling in quantitative linguistics. *Journal of Quantitative Linguistics, 14,* 81-94.

Marbe, K. (1904). *Über den Rhythmus der Prosa*. Giessen: J. Ricker'sche Verlags-buchhandlung.

Orlov, J.K., Boroda, M.G. & Nadarejšvili, I.S. (1982). *Sprache, Text, Kunst. Quantitative Analysen.* Bochum: Brockmeyer.

Roach, P., Knowles, G., Varadi, T., & Arnfield, S. (1993). MARSEC: a machine-readable spoken English corpus. *Journal of the International Phonetic Association, 23,* 47–54.

Taylor, L. J., Knowles, G. (1988). *Manual of information to accompany the SEC Corpus: The machine-readable corpus of Spoken English.* Retrieved July 2, 2013, from http://icame.uib.no/sec/

Wilson, A. (2017). Units and constituency in prosodic analysis: a quantitative assessment. *Journal of Quantitative Linguistics, 24,* 163-177.