

Different Parameters of the Negative Hypergeometric Distribution as a Discriminating Feature for Musical or Composer's style

ZUZANA MARTINÁKOVÁ-RENDEKOVÁ

Faculty of Music

Academy of Arts, Banská Bystrica

Jána Kollára 22, SK-974 01 Banská Bystrica

SLOVAKIA

Abstract: – In previous studies some hypotheses and testing methods have been introduced. The testing method used in this study is rank-frequency distribution of the pitch values in musical works of various musical and composer's styles. Using a special computer programme QUAMS (Köhler, 1995/1996), the iterative fitting of probability distributions (Altmann-Fitter 2.0 Programme) and Ord's criterion some interesting results were obtained. Testing compositions by Palestrina, Bach, Beethoven and Liszt the position of each composer was individual and characteristic. We suppose that discovering of discriminating criterion for musical or composer's style can bring in future some interesting results.

Key-Words: – MIDI-Compositions, Testing Methods, Rank-Frequency Distribution, Iterative Fitting of Probability Distributions, Ord's Criterion, Musical Style

1 Introduction¹

In previous investigations in the field of systems theory, computing analysis and applications of synergetics principles in musicology yielded some interesting results. They lead to the setting up of hypotheses which could be captured by mathematical formulations and tested by quantitative methods. We have observed that using Reinhard Köhler's QUAMS and the Altmann-Fitter 2.0 software for the analysis of the tone pitch the most empirical rank-frequency distributions abide by the *negative hypergeometric distribution*, however, its moments and parameters differ according to composition, composer and style. This fact thus may become a discriminating criterion for musical types and styles.

2 Hypotheses and their testing

Also in previous studies we pointed out that models of music must provide mathematically formulated hypotheses.

The data used in our analysis have been collected from compositions by selected composers, which are accessible on internet in MIDI device². In a previous study (Martináková-Rendeková, WSEAS, 2004) we have shown that the frequency distribution of pitch in a great number of analysed musical texts of MIDI device follows to the *negative hypergeometric distribution*:

$$P_x = \frac{\binom{M+x-1}{x} \binom{K-M+n-x-1}{n-x}}{\binom{K+n-1}{n}}, \quad x=0,1,2,\dots,n$$

where, K , M , n are parameters and x is the rank of the tone. By convention we began the ranking with 0.

In this study we would like to show that the parameters of this negative hypergeometric distribution (further NHG) are not always identical. Our theory is based on the assumption that the parameters depend on the style and composer.

To the main criteria of selection concerning the affiliation to the type and style of music: we selected compositions of the classical music of different styles and composers. The analysis was performed by means of the software QUAMS which can order all used tone pitches in the musical text according to type and frequency of tone pitches, i.e. the program is able to establish rank-frequency distributions of pitch values.

¹ Results published in this study precede a long-time collaboration with the professor of synergetics and quantitative linguistics at the University in Trier Reinhard Köhler and the professor of mathematical linguistics Gabriel Altmann who was many years active at the University in Bochum. I am deeply grateful to the both scientists for relevant advices and remarks to the subject of my study. Prof. Altmann has helped me by elaboration of gained data using statistical methods and Ord's criterion.

² <http://www.notation.com/MIDI-Composer.html>,
<http://www.classicalarchives.com>; <http://www.midifilearchive.com>;
<http://www.stormloader.com>, etc.

The data generated from MIDI musical texts were further analysed by the software Altmann-Fitter 2.0, automatically generated the most adequate distribution which was always the NHG distribution, however having in each case different parameters.

3 Probabilistic criterion of selection

A theory based on the number of tone types and their values in a composition which are discriminating features for musical or composer's style is not without its problems. It is not very probable that depending on style and composer the musical text must have a different number of tone pitch types.

However we know that the frequency distribution of pitch values is always the same in musical texts of different genres, styles and composers. If we accept that the octaves do not mean the same pitch type, then we can suppose that in the works by Webern who used the change of octave positions very frequently, there would

be more pitch types than in the works by Palestrina, for example. The main criterion for our decision was the affiliation to the style period and the establishment of homogeneous group (e.g. Masses by Palestrina, Preludes and Fugues by Bach, Piano Sonatas by Beethoven etc.).

3 Results of Fitting

Let us first demonstrate the results of fitting the negative hypergeometric distribution to all data. The results are shown in Tables 1a to 1g. The columns K , M and n are reserved for the parameters of the NHG distribution, columns Chi , DF and P show the result of fitting the NHG distribution to the rank-frequency data; the last four columns contain the mean ($M1$) and the first two central moments ($M2$, $M3$) which are necessary for scrutinizing Ord's criterion (see below).

Table 1a: G.P. da Palestrina: Missa Ascendo ad Patrem (ca.1550)

	K	M	n	Chi	DF	P	M1	M2	M3
Motet	4,5108	1,0639	26	28,33	21	0,13	6,13	25,52	108,57
Kyrie	3,6868	0,9536	23	13,19	19	0,83	5,95	25,02	107,14
Gloria	4,0821	1,0003	25	17,35	21	0,69	6,13	25,67	108,88
Credo	5,6046	1,1121	31	44,23	23	0,01	6,18	26,45	137,55
Sanctus	4,1017	0,9712	24	4,54	19	0,99	5,62	23,69	104,55
Benedictus	6,0538	1,1509	29	4,48	19	0,99	5,57	21,90	105,34
Agnus Dei I	3,2998	0,8618	23	6,98	19	0,99	6,01	26,94	117,82
Agnus Dei II	3,5667	0,9721	22	4,78	18	0,99	5,99	23,96	85,04

Table 1b: G.P. da Palestrina: Missa Ave Regina Coelorum (ca.1550)

	K	M	n	Chi	DF	P	M1	M2	M3
Kyrie	3,7066	1,0111	22	3,47	18	0,99	6,00	23,42	84,70
Gloria	6,2865	1,2121	31	12,95	22	0,93	5,98	23,62	105,95
Credo	4,6929	1,0857	26	33,92	21	0,04	6,02	24,19	97,59
Sanctus	3,6059	1,0147	21	1,93	17	≈ 1	5,93	22,64	79,00
Benedictus	4,3426	1,0712	23	7,46	18	0,98	5,77	21,69	85,71
Agnus Dei I	3,1982	1,0628	20	1,15	17	≈ 1	6,65	24,39	67,87
Agnus Dei II	3,5908	0,9872	21	4,54	17	0,99	5,77	21,92	79,52

Table 1c: G.P. da Palestrina: Missa Papae Marchelli (ca.1550)

	K	M	n	Chi	DF	P	M1	M2	M3
Kyrie	3,3304	0,8952	20	10,50	17	0,88	5,38	20,66	71,02
Gloria in excelsis Deo	4,4502	0,9850	26	24,09	21	0,29	5,76	24,06	101,32
Credo in unum Deum	4,2843	0,9843	25	23,67	21	0,31	5,74	23,89	103,46
Sanctus	6,7499	1,0989	35	13,57	23	0,94	5,63	24,25	122,13
Benedictus	4,3547	1,0202	24	8,86	19	0,98	5,39	21,89	104,96
Agnus Dei I	3,6098	0,8978	23	3,93	19	0,99	5,72	24,40	101,43
Agnus Dei II	4,4018	1,1259	24	26,71	19	0,11	6,14	23,37	83,84

Table 1d: G.P. da Palestrina: Missa Veni Sponsa Christi (ca.1550)

	K	M	n	Chi	DF	P	M1	M2	M3
Kyrie	4,7671	1,0470	28	9,28	21	0,98	5,98	25,50	121,07
Gloria	4,4642	1,0707	25	12,99	20	0,88	5,99	23,97	101,06
Credo	3,7476	1,0085	23	10,78	20	0,95	6,19	25,08	97,07
Sanctus	4,3477	1,0601	24	13,54	19	0,81	5,85	22,72	87,85

Benedictus	3,4505	0,8706	21	5,42	17	0,99	5,29	21,85	88,61
Agnus Dei I	3,4363	0,9384	21	6,37	17	0,99	5,73	22,25	75,57
Agnus Dei II	3,3401	0,9684	22	6,96	18	0,99	6,38	25,46	78,82

Table 1e: J.S. Bach: WTP I: Preludes and Fugues (1722)

	K	M	n	Chi	DF	P	M1	M2	M3
P+F1	6,1263	1,0829	53	13,08	37	≈ 1	9,31	61,72	557,75
P+F2	4,7639	0,9713	50	7,58	40	≈ 1	10,13	76,06	692,52
P+F3	5,2035	1,0736	52	40,69	41	0,48	10,56	74,53	682,81
P+F4	4,5908	1,0089	52	21,56	43	≈ 1	11,43	87,72	755,19
P+F5	4,1073	0,9385	47	10,15	40	≈ 1	10,56	81,56	753,18
P+F6	4,1752	0,9330	50	13,53	42	≈ 1	11,00	91,20	938,26
P+F7	5,3703	1,1350	52	45,07	41	0,31	10,87	75,11	688,89
P+F8	4,0663	0,9328	51	9,98	44	≈ 1	11,59	97,34	942,09
P+F9	5,7472	1,1192	52	21,91	38	0,98	9,89	67,26	641,40
P+F10	3,8272	0,8982	48	10,62	42	≈ 1	11,20	92,06	836,22
P+F11	3,7004	0,9416	46	11,03	40	≈ 1	11,34	90,53	835,83
P+F12	3,8831	0,9376	49	18,80	43	≈ 1	11,70	98,38	930,11
P+F13	4,6564	0,9763	49	14,08	39	≈ 1	10,18	74,44	643,95
P+F14	4,2174	0,9998	47	18,69	40	≈ 1	11,07	83,67	695,54
P+F15	5,3133	1,1243	55	26,42	44	0,98	11,80	85,42	694,70
P+F16	4,1119	0,9482	49	10,67	41	≈ 1	11,22	89,40	835,57
P+F17	5,0150	1,1007	45	17,13	36	≈ 1	9,83	62,65	478,97
P+F18	3,8696	0,9136	50	9,18	43	≈ 1	11,67	100,38	1039,33
P+F19	4,8226	1,0880	48	31,32	39	0,80	10,86	73,80	574,03
P+F20	4,9818	1,0430	54	17,54	44	≈ 1	11,23	87,12	835,90
P+F21	5,1065	1,0115	53	14,50	42	≈ 1	10,38	79,70	829,32
P+F22	3,6836	0,8576	47	17,69	41	≈ 1	10,54	88,02	919,70
P+F23	5,1855	1,0743	49	21,43	38	0,99	10,10	68,71	586,79
P+F24	3,1656	0,8376	47	19,69	43	≈ 1	12,17	109,72	1051,50

Table 1f: Beethoven: Piano Sonatas

	K	M	n	Chi	DF	P	M1	M2	M3
1 (537)	3,3794	0,8182	58	48,06	54	0,70	14,04	151,45	1843,27
2 (626)	3,310	0,7980	61	59,90	57	0,37	14,71	169,37	2073,24
3 (625)	3,8306	0,9141	66	75,36	59	0,07	15,55	173,87	2237,62
4 (868)	3,2369	0,8212	63	153,73	59	0,00	15,88	186,92	2455,01
5 (473)	3,5571	0,8149	64	76,04	59	0,07	14,38	170,18	2436,69
6 (404)	3,1631	0,8715	59	41,69	55	0,91	16,10	176,29	1934,68
7 (563)	3,3070	0,8552	65	56,38	61	0,64	16,81	193,52	2092,58
8 (455)	3,2955	0,8600	63	61,04	59	0,40	16,40	191,33	2361,90
9 (431)	3,2556	0,7598	62	32,11	58	0,998	14,47	169,85	2104,38
10 (392)	4,1738	0,9144	63	51,25	55	0,62	13,62	140,38	1804,95
11 (646)	3,8415	0,8847	65	156,51	58	0,00	14,66	163,71	2214,29
17 (550)	3,0394	0,7784	63	56,87	60	0,59	16,13	186,92	2464,33
18 (734)	3,4194	0,8127	64	259,47	59	0,00	14,74	180,65	2750,82
19 (410)	3,1534	0,6992	52	62,53	48	0,08	11,31	116,24	1467,24
21 (757)	2,9381	0,8136	66	115,52	62	0,00	17,99	232,21	3071,50
22 (292)	3,2101	0,8663	63	32,05	59	0,99	17,00	198,31	2384,92
23 (911)	3,0898	0,8164	68	46,74	65	0,96	17,97	229,60	2956,72
24 (274)	4,3974	0,9138	70	55,82	60	0,63	14,53	161,35	2252,52
25 (415)	3,2712	0,7770	59	52,46	55	0,57	13,78	157,79	2105,53
26 (430)	4,7031	0,9270	80	54,29	68	0,89	15,76	186,92	2739,10
27 (530)	3,2475	0,7603	67	19,72	63	≈1,00	15,69	196,68	2713,27
28 (434)	3,2957	0,8389	72	26,47	68	≈1,00	18,33	240,38	3272,27
29 (689)	3,5930	1,0411	77	150,04	74	0,00	22,31	269,77	2802,06
30 (501)	4,3523	0,8944	80	77,76	70	0,25	16,54	206,15	3166,79
31 (420)	3,6356	0,8820	73	34,92	68	0,99	17,71	223,84	3175,40
32 (538)	3,1670	0,8799	72	88,73	69	0,06	20,00	264,92	3282,48

Table 1g: Liszt

Franz Liszt	K	M	n	Chi ²	DF	P	M1	M2	M3	
Piano concerto Es Major (1855)	1	3.0094	0,8189	80	112,36	76	0,004	21,95	329,64	4231,48
	2	2,8314	0,8784	78	64,93	75	0,79	24,20	347,30	3794,00
	3	3,0794	0,8482	78	29,88	74	≈1	21,21	310,92	4469,67

Transcendental etude Heroica		2.8693	0,8738	77	9,61	73	≈ 1	23,45	337,50	4197,63
Trans. Etude Feux Follets		3.4422	0,8370	73	19,27	67	≈ 1	17,53	226,23	3325,78
Trans. Etude Mazeppa		3.22964	0,9118	80	31.4169	76	≈ 1	22.586	313.50	3929.384
Venezia et Napoli	Gondoliera	5.9192	0,9772	78	72,99	55	0,05	11,49	107,06	1575,26
	Canzone	4.028	0,7361	68	41,87	58	0,95	12,17	145,13	2214,19
	Tarantella	3.2588	0,8725	78	22,23	74	≈ 1	20,88	291,59	4043,20
Sonata B Minor		3.2149	0,9651	82	85,74	79	0,28	24,62	343,65	3884,81
Love dreams 3		3.03984	0,7257	73	17.4778	67	≈ 1	16.949	247.992	4247.316
Petrarca Sonnet 104		3.77872	0,9364	80	19.9301	71	≈ 1	19.082	248.269	3813.131
L'isle joyeuse		3.99045	0,7714	80	47.6254	71	0.9851	15.079	203.721	3518.497
Hungarian Rhapsody	1	2.7236	0,7916	75	37.3843	72	0.9998	21.799	315.205	3861.366
	3	2.6174	0,6323	69	10.0784	64	≈ 1	16.428	247.642	4021.724
Hungarian Dance No 1		3.8356	0,8646	71	41.0237	62	0.99	15.21	181.59	2769.09
Hungarian Dance No 5		3.5825	0,8372	57	13.6469	50	≈ 1	12.93	132.19	1577.97
Hungarian Dance No 6		3.6569	0,8352	61	30.01	54	0.99	13.64	143.34	1652.85

The fitting of NHG reveals different phenomena.

As can be seen, in works by Palestrina the parameters attain values in the following intervals:

$$n \in <20, 35>$$

$$K \in (3.1, 6.8) \quad \bar{K} = 4.2436 \quad \sigma(K) = 0.9055 \quad \sigma(\bar{K}) = 0.1681 \\ M \in (0.85, 1.21) \quad \bar{M} = 1.0173 \quad \sigma(M) = 0.0821 \quad \sigma(\bar{M}) = 0.0153$$

For the two works by Bach we can ascertain the following intervals of parameters:

$$n \in <45, 58>$$

$$K \in (3.00, 6.2) \quad \bar{K} = 4.5546 \quad \sigma(K) = 0.7083 \quad \sigma(\bar{K}) = 0.1022 \\ M \in (0.81, 1.135) \quad \bar{M} = 0.9955 \quad \sigma(M) = 0.7989 \quad \sigma(\bar{M}) = 0.0115$$

For Beethoven we have

$$n \in <58, 80> \\ K \in (2.9, 4.4) \quad \bar{K} = 3.4948 \quad \sigma(K) = 0.4488 \quad \sigma(\bar{K}) = 0.0881 \\ M \in (0.69, 1.05). \quad \bar{M} = 0.8467 \quad \sigma(M) = 0.0678 \quad \sigma(\bar{M}) = 0.0133$$

For Liszt we have

$$n \in <57, 82> \\ K \in (2.6, 5.92) \quad \bar{K} = 3.4504 \quad \sigma(K) = 0.7311 \quad \sigma(\bar{K}) = 0.1723 \\ M \in (0.63, 0.98) \quad \bar{M} = 0.8396 \quad \sigma(M) = 0.0843 \quad \sigma(\bar{M}) = 0.0199$$

(a) The parameter n is small with Palestrina and grows until Beethoven. Liszt does not differ from Beethoven, the utmost range is attained.

(b) The parameter M seems to decrease with time.

5 Ord's criterion

In order to visualize the differences between empirical distributions one sets up Ord's criterion using the first three empirical moments and defines the quantities

$$I = \frac{m_2}{m'_1} \quad S = \frac{m_3}{m_2}$$

which can be found for each composer in the above tables as $M1, M2, M3$. The values $<I, S>$ which are shown in the Table 2 are inserted in a Cartesian coordinate system as can be seen in Figure 1. The straight line $S = 2I - 1$ represents the upper boundary of the negative hypergeometric distribution.

Table 2: The $\langle I, S \rangle$ -values for the four composers

Palestrina		Bach		Beethoven		Liszt	
I	S	I	S	I	S	I	S
4,1631	4,2543	6,6294	9,0368	10,7870	12,1708	15,0178	12,8367
4,2050	4,2822	7,5084	9,1049	11,5139	12,2409	14,3512	10,9243
4,1876	4,2415	7,0578	9,1615	11,1814	12,8695	14,6591	14,3756
4,2799	5,2004	7,6745	8,6091	11,7708	13,1340	14,3923	12,4374
4,2153	4,4133	7,7235	9,2347	11,8345	14,3183	12,9053	14,7009
3,9318	4,8100	8,2909	10,2879	10,9497	10,9744	13,8803	12,5339
4,4825	4,3734	6,9098	9,1717	11,5122	10,8132	9,3177	14,7138
4,0000	3,5492	8,3986	9,6783	11,6665	12,3446	11,9252	15,2566
3,9033	3,6166	6,8008	9,5361	11,7381	12,3896	13,9650	13,8660
3,9498	4,4856	8,2196	9,0834	10,3069	12,8576	13,9582	11,3046
4,0183	4,0343	7,9832	9,2326	11,1671	13,5257	14,6317	17,1268
3,81793,7	3,4894	8,4085	9,4543	11,5883	13,1839	13,0106	15,3589
591	3,9516	7,3124	8,6506	12,2558	15,2273	13,5102	17,2712
3,6677	2,7827	7,5583	8,3129	10,2776	12,6225	14,4596	12,2503
3,7990	3,6277	7,2390	8,1328	12,9077	13,2273	15,0744	16,24014
3,8401	3,4376	7,9679	9,3464	11,6653	12,0262		
4,1771	4,2111	6,3733	7,6452	12,7769	12,8777		
4,1620	4,3307	8,6015	10,354	11,1046	13,9605		
4,3073	5,0363	6,7956	7,7782	11,4507	13,3439		
4,0612	4,7949	7,7578	9,5948	11,8604	14,6539		
4,2657	4,1570	7,6782	10,4055	12,5354	13,7954		
3,8062	3,5875	8,3510	10,4487	13,1140	13,6129		
4,2642	4,7478	6,8030	8,5401	12,0919	10,3868		
4,0017	4,2161	9,0156	9,5835	12,4637	15,3616		
4,0517	3,8704			12,6392	14,1860		
3,8838	3,8666			13,2460	12,3905		
4,1304	4,0554						
3,8831	3,3964						
3,9906	3,0958						

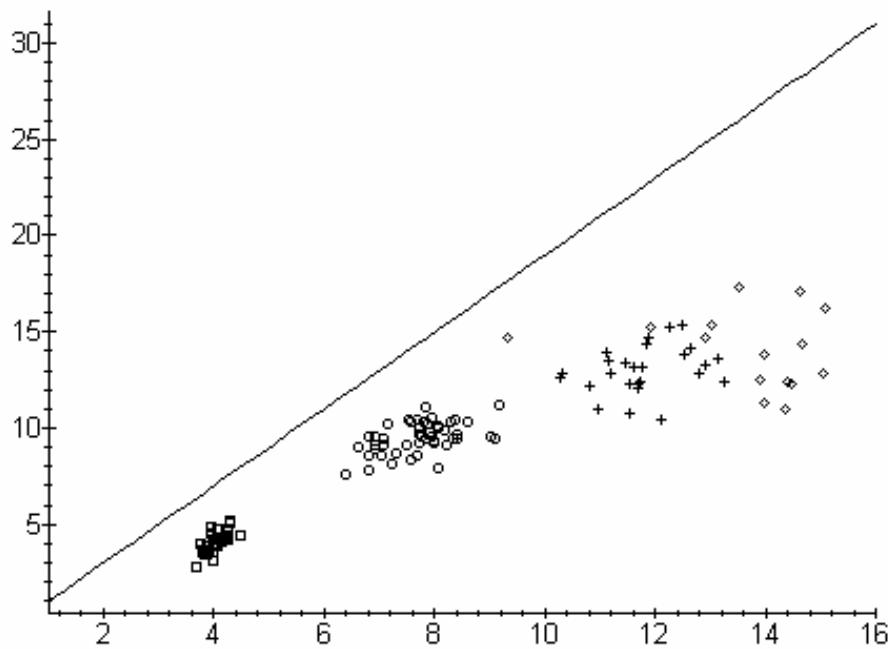


Fig. 1: Ord's criterion for four composers: Palestrina (box), Bach (circle), Beethoven (cross), Liszt (diamond)

As can be seen the position of each composer is very characteristic. They can be clearly distinguished without discrimination analysis. Besides, one sees both a

historical development from Palestrina to Liszt and the lost of compactness in the same direction.

4 Conclusion

In future we have to compare as great number of composers and compositions as possible. Surely more composers of different style periods used the same or similar tone space, however, the relations between the parameters could be different. Further, the hypothesis can be set up that the compactness of the $\langle I, S \rangle$ -sets

References:

- [1] Altmann, G.: *The Art of Quantitative Linguistics*. Journal of Quantitative Linguistics 4 (1-3) 1997, pp. 13-22.
- [2] Grzybek, P., Altmann, G.: *Oscillation in the frequency-length relationship*. In: Glottometrics 5, 2002, pp. 97-1007.
- [3] Grzybek, P.: *History of Quantitative Linguistics III. Nikolaj Gavrilovic Cernysevskij - A Forerunner of Quantitative Stylistics in Russia*. In: Glottometrics 7, 2004, pp. 91-93.
- [4] Grzybek, P.: *History of Quantitative Linguistics IV. Anton Semenovic Budilovic (1846-1908) – A Forerunner of Quantitative Linguistics in Russia?* In: Glottometrics 7, 2004, pp. 94-96.
- [5] Grzybek, P. (eds.): Contributions to the science of Text and Language. Word Length Studies and Related Issues. In: Text, Speech and Language Technology. Springer Netherlands 2006
- [6] Kelih, E.; Antić, G.; Grzybek, P.; Stadlober, E.: *Classification of Author and/or Genre? The Impact of Word Length*. In: http://www-gewi.kfunigraz.ac.at/quanta/publ/2004_ke-ga-pg-est_do_author-genre.pdf
- [7] Köhler, R.: *Synergetic Linguistics*. In: Contributions to Quantitative Linguistics. Dordrecht, Boston, London : Kluwer, 1993, pp. 41-52.
- [8] Köhler, R.; Martináková-Rendeková, Z.: *A systems theoretical approach to language and music*. In: Systems. New Paradigms for Human Sciences. Walter de Gruyter, Berlin – New York 1998, pp. 514-546.
- [9] Köhler, R., Altmann, G.: *Probability distributions of syntactic units and properties*. In: Journal of Quantitative Linguistics 7, 2000, pp. 189-200.
- [10] Martináková-Rendeková, Z.: *Systems Theoretical Modelling in Musicology*. In: Mathematics and Computers in Modern Science. Acoustics and Music, Biology and Chemistry, Business and Economics. WSES (World Scientific and Engineering Society) Press 2000, pp. 122-127.
- [11] Martináková-Rendeková, Z.: *Synergetische und systemtheoretische Aspekte der Musikanalyse*. In: Semiotische Berichte 1-4/O2, Wien 2002, pp. 191-216.
- [12] Martináková-Rendeková, Z.: *Systems Theoretical Modelling of the 20th Century Music (An Endeavour to Categorization)*. In: WSEAS Transactions on Acoustics and Music, Vol. 1, January 2004, pp. 1-6.
- [13] Martináková-Rendeková, Z.: *Rank-Frequency Distribution of Pitch in Musical Texts Using Altmann-Fitter 2.0 and Reinhard Köhler's QUAMS Computer Programmes*. WSEAS 2004. In: www.wseas.org
- [14] Popescu, Ioan-Iovitz: *On a Zipf's Law Extension to Impact Factors*. Submitted to Glottometrics June 23, 2003, <http://www.geocities.com/iipopescu>
- [15] Wimmer, G.; Köhler, R.; Grotjahn, R.; Altmann, G.: *Towards a Theory of Word Length Distribution*. Journal of Quantitative Linguistics 1, 1994, pp. 98-106.
- [16] Wimmer, G.; Witkovský, V.; Altmann, G.: *Modification of Probability Distributions Applied to Word Length Research*. Journal of Quantitative Linguistics 6, 1999, pp. 257-268.
- [17] Wimmer, G.; Šidlík, P.; Altmann, G.: *A New Model of Rank-Frequency Distribution*. Journal of Quantitative Linguistics Vol.6, No.2, 1999, pp. 188-193.
- [18] Wimmer, G.; Altmann, G.: Models of frank-frequency distributions in language and music. In: Quantitative Linguistics (Festschrift in honour of Luděk Hřebíček), Uhlirova, L., Wimmer, G., Altmann, G., Köhler, R. (eds.), 2001, pp. 283-294.