Glottometrics 22, 2011, 5-16

Euphony in Slovak lyric poetry

Radek Čech
Ioan-Iovitz Popescu
Gabriel Altmann

1. Total euphony

Euphony is a well known phenomenon in poetry, especially since the introduction of rhyme which directly evokes it. In classical Javanese poetry either the vowels had their prescribed place in the line or they expressed some special mood. In poetic constructions of this kind the placing of some sounds is conscious and represents some kind of binding. However, there are also cases where the poet is not quite aware of the phonic component of his text; he cares for the content, but subconsciously he creates a construction containing elements of euphony. If we want to demonstrate it, we must set up a definition of euphony and present a method of its measurement.

Euphony in a line will be defined here as a function of non-random (i.e. significant) repetition of one or more sounds. Every phoneme can contribute to euphony either by its special position in the line or by its mere repetition. Some sounds may contribute to the euphony of the strophe (e.g. those in the rhyme position) but need not have a euphonic value in the line. The same holds for a combination of two or more sounds in a certain order.

In general, a sound can have a euphonic value only if it occurs in the line at least twice. Since neither vowels nor consonants can alone fill the whole line, we shall consider separately the number of vowels (V) and that of consonants (C). If there are V places for vowels, then a given vowel has 2^V possibilities of appearing there in different combinations (positions). But this fact must be scrutinized for each vowel separately because for all of them the probability of their occurrence is different. A sound occuring seldom in the language has a smaller probability, but the smaller the probability, the greater is the euphonic values of its appearance. We simplify the problem and consider the occurrence at any position as independent from previous occurrences. Then a sound can occur in the given position with probability p and fail to occur with probability $1 - p = q$. Hence the probability that a vowel occupies x positions out of V possible ones is given as

$$(1) \qquad P_x = \binom{V}{x} p^x q^{V-x},$$

i.e. binomially. The same holds for consonants replacing V by C. Since we need not only the probability of the given occurrence but also all the extreme ones, we compute (cf. Altmann 1966, Wimmer G. et al. 2003; Strauss, Fan, Altmann 2008: 45f)

$$(2) \qquad P(X \geq x_i) = \sum_{x=x_i}^{V} \binom{V}{x} p^x q^{V-x},$$

where x_i is the empirically observed number of the vowel i in the line (replace V by C with consonants). Setting the significance level at $\alpha = 0.05$ we consider a sound as having a euphonic value only if $P(X \geq x) < 0.05$. In order to express the euphonic weight of the given probability, we set up the indicator

$$(3) \qquad E(i) = \begin{cases} 100[\alpha - P(X \geq x_i)], & if\ \alpha > P(X \geq x_i) \\ 0 \quad otherwise \end{cases},$$

and compute it for each sound of the line. In order to obtain the euphonic value of the line, we compute the mean euphony $E(i)$ of those phonemes which have a positive euphony (> 0). Hence for a line we obtain

$$(4) \qquad \bar{E}(line) = \begin{cases} \frac{1}{k}\sum_{i=1}^{k} E_i, & if\ k \geq 1 \\ 0 & otherwise \end{cases}$$

where k is the number of sounds with significant euphony (not all sounds of the line). This indicator can be used for studying the course of euphony in the poem.

Let the number of verses in the poem be n. Then the extent of euphony in the poem can be obtained from the formula

$$(5) \qquad E(poem) = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{k}\sum_{i=1}^{k} E_{ij} = \frac{1}{n}\sum_{j=1}^{n}\bar{E}(line)_j$$

which simply adds the euphonic results in all lines and divides the sum by the number of lines.

Obtaining the expected relative frequencies p of individual sounds is a problem of its own. If one analyzes the work of a poet who lived in 19th century, one cannot use the frequencies obtained from a modern corpus. Since in language there are no populations (cf. Orlov, Boroda, Nadarejšvili 1982); one cannot improve the facts taking a corpus of texts from the 19th century. For some centuries (or languages) there were even no corpora. One can approximate the quasi-population by choosing the best available texts, viz. the works of the given author but if (s)he wrote both prosaic and poetic works, mixing them up may strongly change the (expected) proportion. Hence the best approximation to the expected relative frequencies is given by the works of the same author written in the same text sort. However, since we consider vowels and consonants separately, we must use the conditional probabilities, i.e. $p_{vowel} = f_{vowel}/N_V$.

Our aim is to study the euphony in the poetic work of the Slovak poetess Eva Bachletová. To this end we computed the frequencies of sounds in all her poems we

want to analyze and set up a table of observed frequencies from which we obtained the expected relative frequencies as f_x/N where N is either N_V or N_C for vowels and consonants respectively. However, we first reorganized the Slovak sound inventory from the euphonic point of view which differs from the inventory presented in http://www.ui.sav.sk/speech/sampa_sk.htm (retrieved April 5, 2011).

Since short and long vowels do not differ acoustically, we consider [a] and [a:] etc. euphonically as identical and obtain the vowels [a,e,i,o,u]; the vowel written as {ä} is pronounced as [e] both in the language of the poetess (personal communication) and in the mother tongue of one of the authors. There are four diphthongs [ia, ie, iu, uo]; the diphthong written as {ov, ou} is interpreted by two sounds [o] and [v]. Consonants {l,ĺ,r,ŕ} and their syllabic forms are unified and represented by the short forms [l, r]. The three variants of {n}, i.e. {n,N\,N} are unified in [n]. The final {-j}, in SAMPA written as {i_^} is identified with {j}. The SAMPA {G} is identical with {ch}, phonetically [x], and the SAMPA {F} has been eliminated because it represents the nasalisation of [a]. Thus we obtain the sounds and their frequencies as presented in Table 1. It is to be emphasized that we do not perform phonemic but rather euphonic analysis. The sample consists of 8515 sounds.

Table 1
Frequencies of sounds in E. Bachletová's poems

Sound	Freq.	Rel. freq.	Rel.freq. category	Sound	Freq.	Rel. frequency	Rel.freq. category
a	913	0,10722255	0,25396384	n	307	0,03605402	0,06239837
e	792	0,09301233	0,22030598	s	444	0,05214328	0,09024390
o	777	0,09125073	0,21613352	z	167	0,01961245	0,03394309
i	673	0,07903699	0,18720445	l	154	0,01808573	0,03130081
u	270	0,03170875	0,07510431	r	375	0,04403993	0,07621951
ia	38	0,00446271	0,01057024	c	193	0,02266588	0,03922764
ie	116	0,01362302	0,03226704	J\ = ď	74	0,00869055	0,01504065
iu	1	0,00011744	0,00027816	S = š	93	0,0109219	0,01890244
uo	15	0,0017616	0,00417246	Z = ž	93	0,0109219	0,01890244
p	252	0,02959483	0,05121951	tS = č	90	0,01056958	0,01829268
b	182	0,02137405	0,03699187	dZ=dž	8	0,00093952	0,00162602
f	73	0,00857311	0,0148374	L = ľ	141	0,01655901	0,02865854
v	283	0,03323547	0,05752033	j	197	0,02313564	0,04004065
w	54	0,00634175	0,01097561	J = ň	205	0,02407516	0,04166667
m	374	0,04392249	0,07601626	k	262	0,03076923	0,05325203
t	355	0,04169113	0,07215447	g	24	0,00281856	0,00487805
d	212	0,02489724	0,04308943	h	133	0,0156195	0,02703252
ts	85	0,00998238	0,01727642	x	83	0,0097475	0,01686992
dz	7	0,00082208	0,00142276				

Rel.freq. category = conditional relative frequency within the category V or C.

For the sake of illustration we show the analysis of the first line of the poem *Aby spriesvitnela.* Orthographically we have

Nemám rada bielu

(eu)phonically.we may write

[ňemam rada bielu].

Since only sounds occurring twice or more can be taken into account, we have two candidates: [m] and [a]. There are 7 consonants in the line, hence $C = 7$; the conditional probability of [m] is 0,07215447 as shown in Table 1, hence we compute

$$P([m] \geq 2) = \sum_{i=2}^{7} \binom{7}{i} 0,07215447^{\ i}(1 - 0,07215447\ )^{7-i} = 0,09389934.$$

Since the result is greater than 0,05, the euphonic weight is zero.

Further, we have [a] three times, $V = 6$, hence

$$P([a] \geq 3) = \sum_{i=3}^{6} \binom{6}{i} 0,25396384^{i}(1-0,25396384)^{6-i} = 0,17575387$$

again, $E(a) = 0$. In the third line of the poem we find the first significant frequency, namely with [ň]. There are 7 consonants and twice [ň], hence

$$P([\check{n}] \geq 2) = \sum_{i=2}^{7} \binom{7}{i} 0,04166667^{i}(1-0,04166667)^{7-i} = 0,0317008$$

yielding

E([ň]) = 100(0,05 – 0,0317008) = 1,8299.

The results are presented in Table 2 together with the orthographical text of the poem.

For easier checking, the number of consonants (C) and vowels (V) in the line is given, too.

Table 2
Euphonic values of sounds in the poem *Aby spriesvitnela*

Text	**C**	**V**	**Euphonies**
Nemám rada bielu	7	6	
dnes je prízrakom chladu	13	7	
znecitlivenia	7	5	[ň] = 1,8299
konečného verdiktu	10	7	
nad človekom	7	4	
nad pocitom	6	4	
nad láskou.	6	3	
Dnes je tu iná biela	8	6	
biela obrazovky	7	6	[b] = 2,4616
počítača	4	4	[č] = 4,8041
tam nahadzujeme	7	6	
svoje vnemy	6	4	
čiernymi linkami	8	6	[i] = 3,6665
rýchlo a bezpečne	8	6	
kreslíme životy	8	6	
slovami,	4	3	
ktoré navždy	7	4	
zmenili bielu	6	5	
a odviedli nás	6	5	
od základných farieb	10	6	
bytia.	2	2	
A možno stačí jedna	9	7	
nenapísaná veta	7	7	
aby „novodobá“	5	6	[b] = 3,7301
biela spriesvitnela.	10	6	[l] = 1,2696, [ie] = 3,5678
Lebo čistá – biela krehkosť	13	8	
prichádza potichu...	7	6	[p] = 0,3620, [x] = 4,4351

The above table yields a number of research possibilities: (1) Is there some regularity in the values of euphony in the course of the poem? (2) What part of euphonies is made up by the sounds of the poem title? (3) How to compare statistically the euphony of two poems? (4) Is there some historical evolution of euphony in the work of the given author?

Here we shall consider only the overall euphony of the poem. Using formula (5) we add all mean euphonies of lines and divide by the number of lines (n = 27). We obtain

$$\bar{E}(poem) = [1{,}8299 + 2{,}4616 + 4{,}8041 + 3{,}6665 + 3{,}7301 + (1{,}2696 + 3{,}5678)/2 + (0{,}3620 + 4{,}4351)/2]/27 = 21{,}3095/27 = 0{,}7892,$$

which represents the mean euphony of the poem per line. As can be seen in Table 2, many lines have E(line) = 0. The euphonies occurring in the same line are averaged, not added.

The variance of the poem's euphony which will be used in comparisons can be computed in different ways according to the aspect of comparison. We restrict ourselves to the following procedure. We compute the mean squared deviations of mean line euphonies from $\bar{E}$ (*poem*), i.e.

$$(6) \qquad Var(E) = \frac{1}{n-1}\sum_{i=1}^{n}[\bar{E}(line)_i - \bar{E}(poem)]^2$$

yielding in the above case

$$Var(E) = [20(0\text{-}0{,}7892)^2 + (1{,}8299 - 0{,}7892)^2 + (2{,}4616 - 0{,}7892)^2 + (4{,}8041 - 0{,}7892)^2 + (3{,}6665 - 0{,}7892)^2 + (3{,}7301 - 0{,}7892)^2 + (2{,}4187 - 0{,}7892)^2 + (2{,}3986 - 0{,}7892)^2]/26 = 2{,}1011,$$

and this variance can be used in the asymptotic normal test for comparing the mean euphonies of two poems, namely as

$$(7) \qquad u = \frac{|\bar{E}(poem)_1 - \bar{E}(poem)_2|}{\sqrt{\frac{Var(E_1)}{n_1} + \frac{Var(E_2)}{n_2}}}.$$

The division of $Var(E)$ by the number of lines must be performed because we need $Var(\bar{E}(poem))$. For example, the difference between the poems *Aby spriesvitnela*.and *Iba neha* yields

$$u = \frac{|0.7894 - 0.8698|}{\sqrt{\frac{2.1011}{27} + \frac{2.5877}{54}}} = 0.2267$$

which is not significant

Table 3
Values of euphony in poems by E. Bachletová

Poem	**#Lines**	**#Euphonies**	$\overline{E}(poem)$	***Var*(*E*)**
Aby spriesvitnela	27	9	0,7892	2,1011
Bez rozlúčky	16	6	0,7316	1,8172
Čakáme šťastie	13	9	1,1597	1,3579
Čakanie na boží jas	29	5	0,4194	1,3195
Čas pre nádych vône	18	10	1,2001	2,6953
Dielo Stvoriteľa	44	19	0,7815	2,0019
Dnešný luxus	12	5	1,3363	3,4899
Do večnosti beží čas	18	5	0,5501	1,3292
Ešte raz	7	5	2,2522	3,2378
Hľadanie odpovedí	24	11	0,9947	2,1287
Iba neha	54	19	0,7784	2,2743
Iba v modlitbe	5	7	1,8158	1,3932
Iba život	14	29	2,6472	1,5591
Ihly na nebi	21	7	1,8567	1,0482
Istota	9	2	0,6610	1,9158
Každodennosť	8	6	1,9020	4,3369
Keď dohorí deň	14	10	1,6372	2,0889
Kým ich máme	16	3	0,4581	1,9220
Malé modlitby	11	21	2,4515	3,4998
Mladé oči	7	2	0,6855	1,4395
Malý ošiaľ	27	12	0,8494	2,1483
Moje určenie	52	17	0,4907	1,1333
Nado mnou ty sám	10	4	0,9285	2,7892
Náš chrám	23	13	1,1775	2,9273
Naše dejiny	7	5	1,0391	2,0358
Naše mamy	14	4	0,9628	2,8418
Naše svetlo	28	17	1,5164	3,3362
Návraty	8	4	0,9654	2,7769
Neha domova	9	6	1,4756	4,0364
Neopusť ma	6	5	1,8638	2,3396
Nepoznateľné	51	19	0,6472	1,5851
Otázka	6	5	1,1409	2,1372
Podobnosť bytia	12	3	0,2917	0,3513
Precitnutie	13	6	0,8343	1,9732
Prvotný sen	27	15	1,1919	2,4199
Rozdelená bytosť	26	8	0,4548	1,0741
Rozťatá prítomnosť	36	8	0,5446	1,7323
Smútok	9	3	0,6482	1,2754
Som iná	21	5	0,4125	1,2397
Spájania	14	4	0,1713	0,2392
Stály smútok pre šesť písmen	48	15	0,4819	0,9211

Tak málo úsmevu	20	11	1,1373	2,5246
Tiché verše	12	2	0,3162	1,0939
To všetko je dar	24	8	0,6882	1,9045
Večerná ruža	15	10	1,4516	3,7202
Večerné ticho	19	39	2,6204	1,2604
Vo večnosti slobodná	31	9	0,8760	2,6253
Vrátili sa	12	4	0,5027	1,1785
Vyznania	26	5	0,3935	1,3004
Z neba do neba	40	13	0,6763	1,7015
Zasľúbenie jasu	12	9	1,8127	2,5282
Zázrak	6	1	0,5695	1,9458
Zbytočné srdce	11	5	0,7625	1,3176

Looking at the numbers in Table 3 we can state that euphony is a quite irregular phenomenon. It is created ad hoc, sometimes subconsciously, sometimes consciously and in many cases in dependence on the meanings of words which are more important than their phonic structuring. In rhymed poetry it has a greater importance because, at least subconsciously, the poet searches for phonic agreements and tries to place the pertinent words at the end of the lines.

It can be shown that the extent of euphony, $\overline{E}(poem)$, does not depend on the number of lines in the poem – at least not for Bachletová – and the same holds for the number of euphonic expressions in relation to the number of lines. The line euphonies of Bachletová lie in the interval <0,3162; 2,6472>, i.e., rather in the lower part of euphony which may move in the interval <0; 5>. This state is most probably caused by the small length of her verses containing many times only one word. According to personal communication, she writes her poems "in one go" and makes corrections only in exceptional cases.

If we observe the phones and the sums of their euphonic values in all poems as presented in Table 4, we can state that no phonetic/phonemic order can be discerned. The most striking is the fact that within the line, vowels do not play any special euphonic role. There is no preference for voiced or voiceless consonants or for a special place of articulation. Some sounds do not even appear in a euphonic role in the complete work of the author, In order to detect the preferences of the poetess, other poets should be analyzed using the same method. Of course, it would be possible to study the iconic origin (cf. Koch 2005) of some words in which euphonic sounds occur; but the data in Table 4 represent the overall euphonic weights occurring in all words of all poems. The poetess avoids iconism, rhythm – which is also an imitation of a natural phenomenon – and even the modern poetic form expressed by rhyme. She does not express her ideas materially but rather philosophically, not caring for any formal restrictions. Of course, a thorough study of the poems from another point of view could reveal iconic components which are present in all languages, but they do not seem to be present in these poems. The individual sounds having euphonic weight do not have submorphemic meaning, which can be found in many English word beginnings (cf. Lvova 2011) and is object of intensive investigation. "Phonetic meaning" seems to be foreign to these poems.

As a matter of fact a more thorough computation of euphony would be possible. Separating vowels and consonants one could obtain e.g. all vowels that occur at least twice (= euphonic vowels) and compute the cumulative probability of the multinomial distribution in which non-euphonic vowels (= occurring only once) would be joined in a common class. The same could be done separately for consonants. However, such a procedure is not only considerably more complex but it would not yield "better" results (cf. Wimmer et al. 2003).

2. Verse alliteration

Alliterations at the beginning of verses (Skinner version) may appear without the conscious will of the poet. They may be caused by perseveration, formal strengthening, spontaneity, or by grammatical necessity. Nevertheless, the poet can create them consciously, too. Since it is not (always) possible to know the cause, our problem is merely to state whether the status quo can be considered random or whether there is some tendency. Since sounds must be repeated, their repetition is necessary but it may be random. The longer a poem the more frequently all sounds occur and alliteration must arise automatically.

In order to test the existence of a tendency we use the same method as above. However, here we shall use the non-conditional frequencies

Consider again the poem *Aby spriesvitnela* in which the initial sounds of verses are as follows

[N,d,z,k,n,n,n,d,b,p,t,s,tS,r,k,s,k,z,a,o,b,a,N,a,b,L,p]

Combining identical sounds together we obtain the frequencies

t	tS	r	L	o	N	d	z	p	s	k	n	b	a
1	1	1	1	1	2	2	2	2	2	3	3	3	3

Since the poem has 27 lines, we use formula (2) replacing V by 27 and the individual *p*'s by the relative frequencies in the second column of Table 1. We obtain only one euphonic tendency, namely with [b] yielding *EU*(b) = 3,0523, hence $\overline{E}(poem) =$ 3,0523/27 = 0,1131. That means, in this poem the alliteration of lines is very low.

Performing this investigation for 54 poems we obtain the results presented in Table 4

Table 4
Verse alliterations in poems by E.Bachletová

Poem	Lines	Significantly alliterated sounds	$\bar{E}$
Aby spriesvitnela	27	b	0,1131
Bez rozlúčky	16	b,ž	0,2239
Čakáme šťastie	13	f,S	0,6617
Čakanie na boží jas	29	a,n,p	0,3825
Čas pre nádych vône	18	-	0
Dielo Stvoriteľa	44	d,j,p,n	0,3261
Dnešný luxus	12	-	0
Do večnosti beží čas	18	u,L,b	0,591
Ešte raz	7	p	0,7024
Hľadanie odpovedí	24	g	0,1996
Iba neha	54	a,t,tS	0,1887
Iba v modlitbe	5	-	-
Iba život	14	z	0,1433
Ihly na nebi	21	n,J	0,3829
Istota	9	u	0,2088
Každodennosť	8	u	0,3152
Keď dohorí oheň	13	p	0,339
Kým ich máme	16	h	0,1543
Malé modlitby	11	J	0,4545
Malý ošiaľ	27	a	0,1078
Mladé oči	7	t	0,2609
Moje určenie	52	a,f,v,k	0,1766
Nado mnou Ty sám	10	p,d	0,4819
Náš chrám	23	a,p,v	0,2088
Naše dejiny	7	ď=J\	0,6923
Naše mamy	14	-	-
Naše svetlo	28	j,k,d	0,4153
Návraty	8	d	0,4287
Neha domova	9	k	0,2276
Neopusť ma	6	J=ň	5,0000
Nepoznateľné	51	n,L	0,1051

Otázka	6	-	0
Podobnosť bytia	29	z	0,1074
Precitnutie	13	b,h	0,4042
Prvotny sen	27	v,tS,z,	0,4444
Rozdelená bytosť	26	Z,J	0,167
Rozťatá prítomnosť	36	g,J(=ň),p,Z	0,482
Smútok	9	-	0
Som iná	21	s	0,2196
Spájanie	14	j	0,3306
Stály smútok pre šesť pismen	48	a,k,f,J(=ň)	0,3978
Tak málo úsmevu	20	s	0,2346
Tiché verše	12	b	0,1987
To všetko je dar	24	p,t,z	0,3774
Večerná ruža	15	j,p,	0,2996
Večerné ticho	19	f	0,2603
Vo večnosti slobodná	31	tS,J(=ň)	0,1372
Vrátili sa	12	f	0,3785
Vyznanie	26	t,Z	0,3585
Z neba do neba	40	d,p,x	0,2393
Zasľúbenie jasu	12	-	0
Zázrak	6	k,s	0,8582
Zbytočné srdce	11	d	0,1876

Aliteration plays a still more irrelevant role in Bachletová poems than the general line euphony. Though here, the alliteration attained both of its extreme values, i.e. 0 in *Naše mamy* and 5,0 in *Neopusť ma*, the rest of the values is very small. The mean alliteration of 64 poems is $\overline{E} = 0{,}3764$ with variance of the mean $V(\overline{E}) = 0{,}0088$

3. Conclusions

The results show that euphony is present in poems even if the author does not create it consciously. It awards to poetry a special phonic colour which need not be present in other text sorts. However, extensive investigation is necessary taking into account both different text sorts and languages. Here we can formulate at least some hypotheses to be tested:

(a) Poetic texts are more euphonic than prosaic ones. Here a test for the comparison of texts must be developed.

(b) The more sounds there are in a language, the more clearly euphony can be presented. The reverse hypothesis would not be quite correct because the computations of euphony is based on sound frequencies and the greater the sound inventory, the smaller are the relative frequencies of sounds and the greater is the chance to obtain small cumulative binomial probabilities.

(c) The measurement of euphony depends also on the aspect we choose. Languages having many (or all).words ending with a vowel and having very few consonant clusters have a more melodious sounding (e.g. Italian, Hawaiian) than those full of consonant clusters, e.g. the Czech sentence "Strč prst skrz krk" will surely not evoke euphonic impressions even if [r] occurs in it frequently; but fortunately, there are no such sentences in texts. Thus, euphony is just a concept having many possible definitions and ways of computation.

References

Altmann, G. (1966). The measurement of euphony. In: *Teorie verše I, 259-261.* Brno: Universita J.E. Purkyně.

Knauer, K. (1969). Die Analyse von Feinstrukturen im sprachlichen Kunstwerk. In: Kreutzer, V, Gunzenhäuser, H. (eds.), *Mathematik und Dichtung: 193-210.* München: Nymphenburger.

Koch, W. (2005). *The iconic roots of language. Essayes on the non-arbitrary origins of human communication*. Lüdenscheid: RAM

Lvova, N.L (2011). Determining phonetic symbolism of the text. In: Kelih, E. et al. (eds.), *Issues in Quantitative Linguistics, Vol 2, 94-103.*

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

Skinner, **B.F**. (1939). The alliteration in Shakespeare's sonnets: A study of literary behaviour. *The Psychological Record 3, 186-192.*

Skinner, B.F. (1941). A quantitative estimate of certain types of sound patterning in poetry. *The American Journal of Psychology 54, 64-79.*

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

Wimmer, G., Altmann, G., Hřebíček, L., Ondrejovič, S., Wimmerová, S. (2003). *Úvod do analýzy textov*. Bratislava: Veda.