In search of a euphony unit: A case study in Czech 1830s poetry*

MICHAL MÍSTECKÝ, RADEK ČECH & PETR PLECHÁČ

ABSTRACT: The study examines selected samples of Czech 1830s poetry production through the prism of a quantitative conception of euphony. Stemming from Jan Mukařovský's reflections on the topic, it tries to strengthen the notion through the creation of exact figures with intersubjective validity. To this end, the count of this property devised by Gabriel Altmann is utilized and innovated by a new unit – the consonant set – which endeavours to grasp the phonetic reality of language more effectively than the casually used concept of sound. The analysis proves the solid character of the new unit, shows interrelations between the two calculations, and proposes a few interpretations of the euphony situation in Czech poetry of the 1830s. Moreover, it demonstrates effective ways of displaying data results by means of scatter plots and cluster analyses. As a spinoff of the core idea, the paper also explores Mukařovský's high evaluation of Mácha's use of euphony.

Key words: euphony, consonant, consonant set, 1830s, Czech poetry

1. Introduction

Euphony can be considered one of the most familiar notions in literary studies focusing on the characteristics of poetry. However, as is common for the majority of "basic" notions of this kind (e.g. *word*, *sentence*, *text* in linguistics), an all-encompassing definition of the phenomenon has not been introduced yet, despite considerable discussions among several prominent Czech linguists, including Mukařovský (1948, 1976), Červenka (2002), and Čermák (2007). Identical problems appear in the *Princeton Encyclopedia of Poetry and Poetics* (Perrine 1972: 258), where a theoretical debate, restricted to a mere declaration of the emotional effect created by euphony and its subsequent distinction from cacophony, gives way to tendentious statements on the appropriateness of the euphonious in a poetic text, and ends in an unfounded statement that euphoniousness is higher in poetry than in ordinary speech; nor does its fourth edition (Carper 2012) seem to take any further steps, as it repeatedly exemplifies the mysterious link between sound and sense, failing to arrive at any positive statement about the concept studied. Such an evasive way of thinking is not fruitful when it comes to the provision of a basis for research methodology.

Czech studies focused on this problem usually present a set of examples, classifications, and selected possible solutions. In Mukařovský, for example, euphony made important appearances at least twice: the first within the context of the essentials of poetic language (Mukařovský 1976), and the second in a case-study of Karel Hynek Mácha's *Máj* (1948; for the first time in 1928); however, neither appearance offers an elaboration on the straightforward characteristic of the device. Furthermore, the two papers even differ from each other in the points they emphasise. In his general study on poetic language, Mukařovský states that euphony "occurs more often in such a way that a certain speech

* While working upon this paper, Petr Plecháč was supported by the Czech Science Foundation, project GA17-01723S (Stylometric Analysis of Poetic Texts).

sound is repeated many times or that an entire cluster of speech sounds is repeated once or, in some cases, many times either in the same or in a somewhat altered pattern" (Mukařovský 1976: 26–27), which is – with all due respect paid to the fact that he is trying to consider the entire set of features – a vague statement that seems to approach broad sound instrumentation more than circumscription of the notion in question. When analysing *Máj*, Mukařovský continues to blur the differences between the two devices, trying to incorporate in the definition the metre of the poem as well. This was repeatedly criticised by a disciple of his, Červenka (2002), who understood euphony as a subtype of sound instrumentation, but despite insisting on extensive corpus-based research, he literally rejected the possibility of delimiting the notion quantitatively. However, Čermák (2007), a well-known translator of Old English poems into Czech, proposed sound instrumentation as a suitable solution when one deals with various patterns of Anglo-Saxon alliteration, thus dismantling one of the purest types of euphony into an intuitive stretch of quasi-similar sounds.

A rather different approach to the problem was introduced by Altmann (1966), who suggested operationalizing the notion strictly and formally as a function of a non-random (i.e., significant) repetition of one or more sounds in a verse (see Section 3). Obviously, this approach does not solve the problem of the general status of the phenomenon, and can even be considered too reductive. However, it enables not only an intersubjectively focused description, but also hypothesis testing (Popescu, Mačutek & Altmann 2009; Čech, Popescu & Altmann 2014; Plecháč & Říha 2014). In this study, we generally proceed using Altmann's euphony research methods.

The aim of this paper is to introduce an innovation in the quantitative research on euphony. Specifically, we define a so-called consonant set (for details, see Section 2) as a unit which is considered to be a carrier of euphony in a poem (up to now, only the speech sound as a unit has been used). Further, this unit is used in a methodological procedure proposed by Altmann (1966) and, finally, a comparison of measurement of euphonies based on individual sounds on the one hand, and on consonant sets on the other, is performed. The article is organized as follows: first, the unit of consonant set is introduced and the reasons for its employment explained; second, the methodology of the computation and the analysed language material are presented; third, a theoretical analysis of the comparison of the two counts is performed, and some general statements are drawn from the counted tests; and fourth, the material is interpreted on the basis of the two methods, and the outcome is confronted with Mukařovský's aforementioned presuppositions. The study concludes with an overview of findings and a sketch of issues for further research.

2. Consonant set

As concerns possible units for analysis, the conceptions of euphony elaborated upon by Mukařovský (1948, 1976) and Červenka (2002) are rigorously divided. As shown in the above definition, Mukařovský is rather inconsistent in his statements, considering both sounds and groups of sounds as possible carriers of euphony; however, no delimitation of such groups is given, which already rules out placing a euphonic occurrence within an accountable frame. The reasoning of Červenka is even more controversial, as his proposal of the phoneme as the basic unit of euphony – though it is at least stricter than Mukařovský's lenience – appears to ignore its very nature: a phoneme is, to wit, not a sound in the pure sense of the word, but a functionally-charged, systemically-working and cerebrally-proceeded intersection of the articulatory chaos of a voice travelling through the human vocal folds;[1] on the other hand, euphony is an utterly phonetic matter, and should thus not be studied with tools designed within the domain of phonology. In the *Princeton Encyclopedia*, Perrine (1972) uses the concept of sounds, even providing a list of the supposedly most euphonic ones (*l*, *m*, *n*, *r*), with the accompanying criterion of ease of their pronunciation. However, as he adds an arrangement of sounds and the metre of a poem as other essentials for a euphony-production, he seems to come to the same standstill as Mukařovský: having compiled too many features of the studied object prevents him from wording a usable attempt at its definition. The same holds for Carper (2012), whose fuzziness has already been mentioned in Section 1.

Despite a lack of a clear solution of the unit issue, the aforementioned papers have fuelled the discussion in the quantitatively-oriented research as well. It began to seem appropriate to ask whether, in the course of euphony research, groups of sounds should not be taken as separate units, as sounds do not exist in words separately, but – especially in consonantal languages such as Czech or Slovak – flock into clusters that may share certain phonetic features (e.g. assimilation or assibilation). For example, in the cluster [pr], which is frequently seen in Czech words, the phonetic qualities of [p] and [r] differ from the situations where the two sounds appear in other clusters, or on their own, in between two vowels. This intuitive predisposition of the research may be backed by the fact that what is considered to be a cluster in one language is to be approached as a single sound in another; for instance, Czech considers [tʃ] a single sound, but a different language may show it as a cluster of [t] and [ʃ], respectively.

The blurring of borders between sounds and clusters has inspired our aim at undertaking euphony research on a basis free of these discrepancies. In the present paper, an attempt is made at establishing a new euphony unit, which will be called, for lack of a more suitable term, a *consonant set*. The consonant set basically encompasses all consonants that are to be found in between two vowels in a line, at its beginning, or at its ending. Thus, if there is a line "přišel poslíček" ["a messenger came"], it has six consonant sets, [př], [š], [lp], [sl], [č], and [k], respectively; the corresponding English verse – "a messenger came" – comprises, in Received Pronunciation, [m], [s], [nʤ], [k], [m]. As for multiple occurrences, a good example is Mácha's line "bloudila blankytnými pásky" ["she was roving through azure rays"], where there is a repetition of the [bl] set. The advantage of such an approach is in the ease and straightforwardness of its operation – the researcher does not have to bother with what to consider a "natural" combination of consonants and what merely a coincidental accumulation of sounds, as both are placed at the same level.

[1] This characteristic of phoneme endeavours to incorporate, within a small space, its main features from the structural point of view; the notion, multifarious in its uses and connotations, is paid due attention in, e.g. Skarnitzl, Šturm & Volín (2016).

3. Methodology and language material

As mentioned above, Altmann's (1966) approach[2] to euphony restricted the phenomenon to counting probabilities of repetitive occurrences of sounds in a verse, basing this idea upon the Shklovsky-based psychological effect of foregrounding (cf. Shklovsky 1991). Mathematically taken, a euphony of a selected unit in a line is the sum of the probabilities subtracted from the statistics-given level of significance, which is, in similar cases, stated at $\alpha = 0.05$, meaning that whenever there is less than five-percent probability that a sound makes more than one appearance in a line, this sound is considered euphonic.

The preceding may be formularized as –

$$(1) \qquad E = 0.05 - \sum_{x=x_1}^{n} \binom{n}{x} p^x q^{n-x} \,,$$

where x stands for the frequency of a sound in a line and n for the number of positions in which it is theoretically possible to place the sound (it thus equals a number of consonants which occur in a line); $\binom{n}{x}$, the so-called binomial coefficient, then takes into consideration all placements of the particular x-member group in a n number of positions, thus equalling the total of existing combinations; p is the probability of the occurrence of the sound, or, which is the same, its relative frequency (the ratio of the frequency of the sound in the given corpus and the total of all sounds – $p = \frac{f(x)}{\Sigma f}$), and q the probability of other cases, i.e. – $q = 1 - p$. This procedure accounts for all imaginable situations of the co-occurrences of a sound in a line. If the euphony count yields negative numbers, such a co-occurrence of a sound is treated as non-euphonic.

In this study, Altmann's (1966) mathematical background is kept, with the addition of the innovative approach of a consonant set, as defined in Section 2. Although this unit swap may seem to be a cosmetic change in the already established system of measurement, its intention is to reflect the articulatory nature of language more realistically, of which the limitative employment of the sound is probably incapable.

To clarify this counting process, we provide an example. At the beginning of Mácha's *Máj*, there is the line –

bloudila blankytnými pásky ["she was roving through azure rays"],

which contains 9 consonant sets – [bl], [ďʼ], [l], [bl], [ŋk], [tn], [m], [p], and [sk]; out of these, [bl] repeats twice. Formula (1) will thus concretize[3] into

$$E_{[bl]} = 0.05 - \sum_{x=2}^{9} \binom{9}{x} 0.002929^{x} (1 - 0.002929)^{9-x}$$
$$= 0.05 - [\binom{9}{2} 0.002929^{2} (1 - 0.002929)^{7} + \binom{9}{3} 0.002929^{3} (1 - 0.002929)^{6} + \binom{9}{4} 0.002929^{4} (1 - 0.002929)^{5} + \binom{9}{5} 0.002929^{5} (1 - 0.002929)^{4} + \binom{9}{6} 0.002929^{6} (1 - 0.002929)^{3} + \binom{9}{7} 0.002929^{7} (1 - 0.002929)^{2} + \binom{9}{8} 0.002929^{8} (1 - 0.002929)^{1} + \binom{9}{9} 0.002929^{9}] = 0.0497.$$

[2] This approach is also described by Čech, Popescu & Altmann (2014), which is a more easily accessible study.

[3] The probability p of the occurrence of the [bl] set, determined from the corpus used in this study, is 0.002929.

Since the calculation yields a positive number, the occurrence of this consonant set is considered euphonic in this line.

In order to provide a coherent frame for euphonic measurements, a group of poets from the 1830s has been chosen. The corpus includes, namely, ten samples from the given period, totalling 27,113 euphonically analysed lines (see Table 1). The transcription is carried out on the basis of the PhoEBE (Phonetics in Eight-Bits Encoding) phonetic alphabet, which is used by the Institute for Czech Literature's Versification Research Group (Plecháč 2018).

Table 1: An overview of the study corpus

Author	Poems	Number of Lines
Josef Jaroslav Langer	Selanky (1830)	1,703
Martin Alexander Přibil	Národní písně obsahu rozmanitého (1830)	950
Jan Kollár	Slávy dcera (1832)	9,130
František Jaroslav Kamenický	Písně v národním českém duchu (1833)	1,876
Antonín Jaroslav Puchmajer	Fialky (1833)	3,505
Josef Vlastimil Kamarýt	Pomněnky (1834)	2,463
Vojtěch Nejedlý	Otokar (1835)	4,312
František Matouš Klácel	Lyrické básně (1836)	1,169
Karel Hynek Mácha	Máj (1836)	824
František Jaromír Rubeš	Deklamovánky a písně (1837)	1,181
	TOTAL	**27,113**

As for the investigation, the corpus was taken as a whole, meaning that it was used to count the relative frequencies of both individual consonants and sets. This procedure was followed by calculating the consonantal euphony[4] (hereinafter E_C) of every line and the consonant-set euphony (hereinafter E_S) of the same, thus –

$$(2) \quad E_{C(line)} = \sum_{i=1}^{n} E_C,$$

$$(3) \quad E_{S(line)} = \sum_{i=1}^{n} E_S,$$

where n is the number of the euphonic units in the line.

Afterwards, the averages of line euphonic values of both E_C and E_S were counted for each author. In order to make the numbers easier to proceed, the probability counts were multiplied by 100:

$$(4) \quad E_{C(author)} = \frac{\sum_{i=1}^{n} E_{C(line)}}{m} 100,$$

$$(5) \quad E_{S(author)} = \frac{\sum_{i=1}^{n} E_{S(line)}}{m} 100,$$

where m is the sum of all lines in an author's sample.

[4] Vowel euphony is not taken into consideration in this study.

Finally, the percentage of euphonic verses for both consonants E_{PC} and sets E_{PS} in the given sample is counted:

$$(6) \quad E_{PC(author)} = \frac{e_c}{m} 100,$$

$$(7) \quad E_{PS\,(author)} = \frac{e_s}{m} 100,$$

where e_C and e_S are the numbers of respective euphonic lines, and m is the sum of all lines in an author's sample.

All the results will be commented upon in Section 5.

4. Results

First of all, it needs to be decided whether the consonant set can be considered a "standard" language unit, i.e., whether it possesses the properties shared by all of the units. It has been proved (Altmann 2005) that the rank-frequency distribution of "well-established" language items (e.g. words, lemmas, morphemes, sounds, and n-grams) follows a "regular" distribution. In this study, we used a generalized form of the power-law distribution, specifically

$$(8) \quad y = ax^{-b} e^{-cx},$$

where x is a rank, y is a frequency, and a, b, c are parameters. The power law distribution accounts for the fact that there are few frequently occurring items, and a very high number of those that make appearances very rarely. This is interpreted as a result of the so-called diversification process in language. Finally, this model is a special case of a general mathematical formula expressing many language laws (cf. Wimmer & Altmann 2005).

Fitting the function (8) to the data, we obtain not only parameter values, but also a degree of a correspondence between the model and the data. This degree is expressed by the determination coefficient R^2 – if the model fits the data sufficiently well, values of $R^2 > 0.9$. Therefore, the higher the R^2, the better the model fits the data (cf. Mačutek & Wimmer 2013 for a discussion of several goodness-of-fit measures applied in quantitative linguistics). In this paper, the *NLREG – Nonlinear Regression and Curve Fitting*[5] software is used for the computation.

The graph (Figure 1) demonstrates that the distribution of consonant sets – which are 4,719 in total – can be approximated with the function (8) with parameters $a = 38520.65$, $b = 0.1479$, $c = 0.0545$, and $R^2 > 0.99$. As for the consonants, we obtain parameters $a = 98099.53$, $b = 0.0396$, $c = 0.0826$, and $R^2 > 0.98$ (cf. Figure 2).

Bearing the fit in mind, one may take it a step further and count the occurrences of both euphonies in the corpus and compare them (see Table 2). This has shown that the total of 15% of lines possess E_C, whereas almost 33% demonstrate signs of E_S. The fact that E_S

5 See <http://www.nlreg.com/index.htm>.

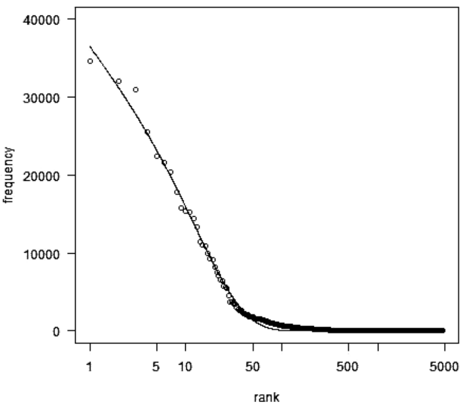

Figure 1: The rank-frequency distribution of consonant sets; the line represents the fit of the data to the function (8)

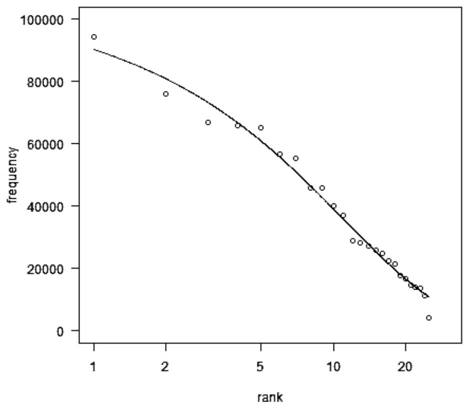

Figure 2: The rank-frequency distribution of consonants, the line represents the fit of the data to the function (8)

verses considerably surpass the number of E_C ones is essential, since the high percentage raises the practical utility of the new euphonic tool, as the classical, consonant-based count is sometimes too severe to yield comparable results. On the other hand, of the counted euphonic verses, only 11% are euphonic in both ways (so-called intersecting verses), this indicating that the two approaches probably take into consideration different properties of language. Given this, it is justifiable to consider both procedures acceptable, as each treats the line from a specific point of view; concerning usage, the qualitative argumentation of the "phonetic realism" of the consonant sets may be also taken into consideration as their advantage.

Table 2: A summary of E_C and E_S lines: their totals, participations in the total of lines, and the percentages of lines that are euphonic in both ways

	Total	Percentage Participation in All Lines	Percentage of Intersecting Lines
E_C Lines	4122	15.20	37.26
E_S Lines	8849	32.64	17.35
		Total of Intersecting Lines	1536 (11.84%)

To support the idea that the two euphonies are independent of each other, the data of the intersecting verses were subjected to further testing. First, their distributions were (via Shapiro–Wilk Test) proved not to be normal, which increases the probability of a presence of mechanism infusing the verses with euphony (e.g. the author's intentions). Consequently, the Kendall τ Correlation Coefficient was used to count whether the two sets of values correspond with each other. This was numerated as $\tau = 0.06$ ($p < 0.001$), which indicates a virtual non-existence of the link between the counts. This is additional proof that the two formulae see the phenomenon from two different angles.

As for the euphonic line values, the averages are 2.95 (for E_C) and 3.74 (for E_S); this distinction was exposed to Mann–Whitney Test (which is nonparametric, and thus usable for the non-normal data), and proved to be significant ($z = -8.12$; $p < 0.001$). This means that the traditional, i.e. sound-based, euphony count is likely to score less than the consonant-set calculations. This may be linked to the fact that there are many more types of consonant sets (4,719 in total) than of consonants (25 in total) in a given sample, this decreasing the probability of the occurrence of several identical ones in a line (and increasing, simultaneously, the value of its E_S – cf. Figure 6). However, as there are many E_S lines that contain, at the same time, high figures in the count, consonant-set euphony seems to be a more effective tool than the consonant count. Nonetheless, additional sound mathematics would be needed to confirm this intuition.

5. Interpretation

The results obtained from the corpus are interpreted by means of a series of graphs (Figures 3, 4, 5, and 6). The first set focuses on the comparison of the two euphony types separately, whereas the second one brings the two counts together, examining them on the basis of the percentages of euphonic lines and the euphonic values, respectively. In all cases, cluster analysis[6] is carried out, and commentaries are provided for the results of the investigation.

The graphs enable the researcher to scrutinize the situation closely. First, in both euphony counts, there are some prominent names that distinguish themselves from the others: in the case of E_C (see Figure 3), this is Klácel, with regard to the percent proportion of euphonic verses, and Mácha, with regard to the average E_C value in a line, whilst E_S (see Figure 4) yields Mácha and Langer. Within the scope of the aforementioned qualitative opinions on euphony, it may be concluded that Mácha's position is significant in both the counts, which is a feasible quantitative rendition of Mukařovský's words about the poet's complex treatment of the studied phenomenon. However, the fact that Mácha, when it came to the balance between value and percentage, gave priority to consonant-set euphony over the simple consonant calculation – doing this, supposedly, due to his creative intuition – contributes to the intricate sophistication of his poetry's stylistic features (polymetric lines, variegated rhyme schemes, metaphorical eccentricities). It is important to bear in mind this combination of factors when Langer, the highest-scoring poet in E_S, is analysed: here, the elevated number is mostly caused by chorus-like line repetitions, grammatical rhymes, discrepancies in line lengths, and frequent epizeuxes; this is caused by the folklore-imitation nature of his poetry.

Second, it is necessary to explain Klácel's increased percentage of E_C lines. His Antiquity-modelled use of verse, with preference given to long structures, is the most probable fuelling force behind this number; however, as the E_C values never score considerably (in eighth place out of the ten samples), it is an example of an extensive, yet not very efficient use of euphony. On the other hand, Mácha's handling of the issue is founded

---

[6] Cluster analysis is utilized in Plecháč & Kolár (2017), and its theoretical background is explained, for instance, in Volín (2007), or in Koščová, Čech & Mačutek (2016). In the present paper, the method of k-means is utilized, and the number of clusters – three – was chosen on the basis of introspection.

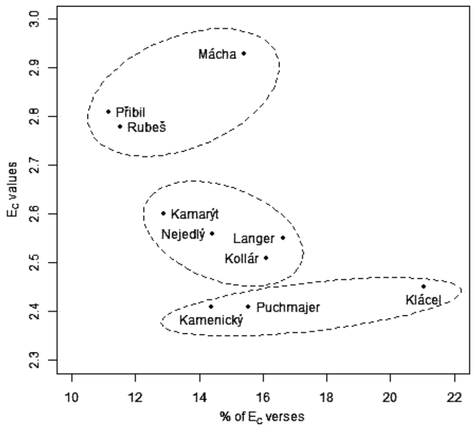

Figure 3: Average E_C values compared to the percentage of E_C verses

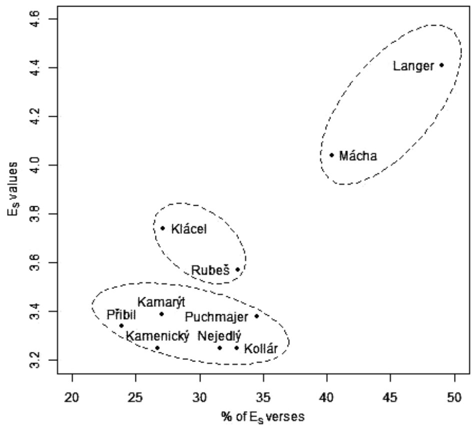

Figure 4: Average E_S values compared to the percentage of E_S verses

upon the infrequent employment of rare consonant sounds, which may have a surprising effect on the reader. The two treatments of euphony thus exploit the broadly defined notions of "quantity" (Klácel) and "quality" (Mácha).

Next, particular attention was devoted to Kollár's quantitative *Introductory Canto* ("Předzpěv"), as the length of the dactylic hexameter promised an escalated euphony score; however, this was not verified (as only 16 percent of these lines show any consonant euphony), with a possible interpretation that the careful choice of long and short syllables blocked any of the poet's other efforts towards sound instrumentation. An unusual verse pattern may thus be an impeding factor in developing euphony.

As for cluster analysis, the poets are grouped according to the euphony values, not the percentages. In both figures, Mácha occurs in the cluster with the highest numbers, which points out his outstanding position among the samples. Moreover, the composition of the two middle clusters is always completely different, which indicates that the counts are founded upon distinct bases. Throughout the clustering, only Puchmajer and Kamenický remain in the same, low-scoring group.

The graphs in Figures 5 and 6 combine the euphonic counts, thus assessing the success of the poets in confrontation of values and percentages. The results in the first graph corroborate the tendency of Klácel to score very well at E_C, as well as Mácha's competition with Langer regarding the efficiency of euphonic values. However, it appears that the case of Klácel is more complicated, as it also acquired eminent E_S values, which were already indicated in Figure 4. These findings, therefore, yield a sort of "double reading" of his poetry: whereas in the case of E_C, Klácel fills many of his lines with multiple, but frequent consonants, but from the E_S perspective, he uses a strategy close to the one of Mácha, producing a limited number of exceptionally efficient euphonic verses. The general implications of this particular outcome have yet to be studied further.

Seen from the perspective of clusters, which follow the same logic as in Figures 3 and 4, the Mácha–Langer duo occupies the highest position, confirming their specific

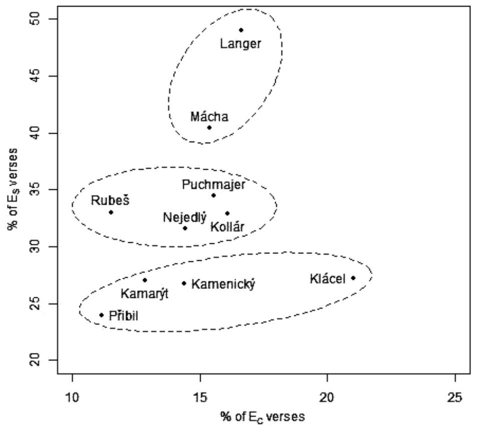

Figure 5: Average E_C and E_S percentages compared to each other

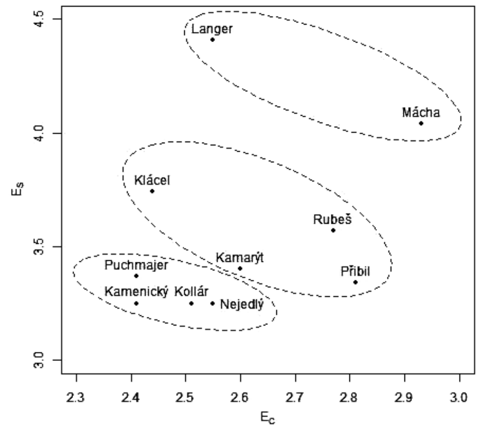

Figure 6: Average E_C and E_S values compared to each other

standing in the counts. It is interesting that in these cross-euphony scatterplots, there is not a single pair that would be classified into the same grouping in both diagrams – this may mean that each of the four counts should be studied separately. The study of poetic idiosyncrasy can possibly start from those authors that either score very high, or who participate in the least numerous clusters.

6. Conclusions

The goal of the article was twofold: first, to prove that consonant-set euphony is an independent and viable method that yields relevant figures about the sonic structure of a line – this has been proved by its conformity to the power-law distribution, and a series of tests that distinguished it from the traditional euphony formula; second, to interpret, very preliminarily, the situation in 1830s Czech poetry, focusing particularly on the prominence of Karel Hynek Mácha. In this case, the answer to the simple question, "Is Mácha's *Máj* significantly euphonic compared to other poems in the corpus?" has been proved to depend on the chosen perspective, as these have provided different results (cf. the figures above). Finally, the Mácha–Langer pairing has been established as euphonically the richest when all circumstances are taken into consideration, the main distinction between the "winners" being found in Mácha's elevated euphony values and Langer's escalating percentages of euphonic verses. Mácha, therefore, places more emphasis on the shocking effect produced by the euphonic uses of rare sounds upon the reader, whereas Langer underlines frequent euphonic occurrences of sounds which are more or less expectable. On the other hand, Klácel, whose numbers are idiosyncratic, uses different strategies in E_C and E_S.

As for cluster analysis, four poets should receive more attention from literary scholars, as the top results of Mácha, Přibil, Rubeš, and Langer should be explained using solid, history-based interpretations. It appears, for instance, that Mácha scores high in all clusterings, whereas Přibil is strong in E_C and weak in E_S; for Langer, the situation is the opposite, as he excels in E_S and lacks in E_C. A provisional conclusion may thus be that Mácha appears to have found a level ground for the treatment of euphony, but more analyses will be needed before this statement is supported with enough evidence.

REFERENCES

Altmann, G. (1966): The measurement of euphony. In: J. Levý (ed.), *Teorie verše 1 / Theory of Verse 1 / Teorija sticha 1: Sborník brněnské versologické konference, 13.–16. května 1964* [Verse Theory 1: Proceedings from the Brno Conference on Verse Theory, 13–16 May 1964]. Brno: Universita J. E. Purkyně, 208–209.

Altmann, G. (2005): Diversification processes. In: R. Köhler, G. Altmann & R. G. Piotrowski (eds.), *Quantitative Linguistik: Ein internationales Handbuch* / *Quantitative Linguistics: An International Handbook*. Berlin, New York: Walter de Gruyter, 646–659.

Carper, T. (2012): Euphony. In: R. Greene, S. Cushman, C. Cavanagh, J. Ramazani & P. Rouzer (eds.), *The Princeton Encyclopedia of Poetry and Poetics*. Fourth edition. Princeton: Princeton University Press, 464–465.

Čech, R., Popescu, I.-I. & Altmann, G. (2014): *Metody kvantitativní analýzy (nejen) básnických textů* [Methods of Quantitative Analysis of (Not Only) Poetic Texts]. Olomouc: Univerzita Palackého. Available online at: <http://www.cechradek.cz/publ/2014_Cech_etal_Metody_Q_analyzy_basnik-kych_textu.pdf>.

Čermák, J. (2007):‚Shromáždím hluboký ze slov svých hvozd…‘: k možnostem překladu anglické středověké poezie [“I will gather deep woods of my words”: On the translation possibilities of English medieval poetry]. In: R. Tomášková (ed.), *Translatologica Ostraviensia: Sborník z konference Den s překladem* [Translatologica Ostraviensia: Proceedings from the Conference on the Day of Translation]. Ostrava: Filozofická fakulta Ostravské univerzity, 29–36.

Červenka, M. (2002): Hlásková instrumentace [Sound instrumentation]. In: M. Červenka, M. Jankovič, M. Kubínová & M. Langerová (eds.), *Zvuk, význam, obraz: Poetika literárního díla 20. století* [Sound, Sense, Image: The Poetic Principles of a 20th-Century Work of Fiction]. Praha: Torst, 9–54.

Koščová, M., Čech, R. & Mačutek, J. (2016): Shluková analýza [Cluster analysis]. In: P. Karlík, M. Nekula & J. Pleskalová (eds.), *Nový encyklopedický slovník češtiny* [The New Encyclopedia of the Czech Language]. Praha: NLN, 1606–1610. Available online at: <https://www.czechency.org/slovnik/ SHLUKOVÁ ANALÝZA>.

Mačutek, J. & Wimmer, G. (2013): Evaluating goodness-of-fit of discrete distribution models in quantitative linguistics. *Journal of Quantitative Linguistics*, 20 (3), s. 227–240.

Mukařovský, J. (1948): Máchův Máj: Studie estetická [Mácha’s Máj: An esthetic study]. In: J. Mukařovský, *Studie z české poetiky III* [Studies in Czech Poetics III]. Praha: Svoboda, 7–202.

Mukařovský, J. (1976): *On Poetic Language*. London, New York: Yale University Press.

Perrine, L. (1972): Euphony. In: A. Preminger, F. J. Warnke & O. B. Hardison, Jr. (eds.), *Princeton Encyclopedia of Poetry and Poetics*. Princeton: Princeton University Press, 258.

Plecháč, P. (2018): *PhoEBE*. Prague: Versification Research Group, Institute of Czech Literature, Czech Academy of Sciences. Available online: <http://versologie.cz/v2/web_content/phoebe.php?lang=en>.

Plecháč, P. & Kolár, R. (2017): *Kapitoly z korpusové versologie* [Chapters from Corpus Verse Studies]. Praha: Akropolis.

Plecháč, P. & Říha, J. (2014): Measuring euphony. In: G. V. Vekshin (ed.), *Methodology and Practices of Russian Formalism: Brikovskij sbornik 2*. Moskva: Azbukovnik, 194–199.

Popescu, I.-I., Mačutek, J. & Altmann, G. (2009): *Aspects of Word Frequencies* (= Studies in Quantitative Linguistics 3). Lüdenscheid: RAM-Verlag.

Shklovsky, V. (1991): *Theory of Prose*. Normal: Dalkey Archive Press.

Skarnitzl, R., Šturm, P. & Volín, J. (2016): *Zvuková báze řečové komunikace: Fonetický a fonologický popis řeči* [The Sound Basis of Speech Communication: A Phonetic and Phonological Description of Speech]. Praha: Karolinum.

Volín, J. (2007): *Statistické metody ve fonetickém výzkumu* [Statistical Methods in Phonetic Research]. Praha: Epocha.

Wimmer, G. & Altmann, G. (2005): Towards a unified derivation of some linguistic laws. In: R. Köhler, G. Altmann & R. G. Piotrowski (eds.), *Quantitative Linguistik: Ein internationales Handbuch / Quantitative Linguistics: An International Handbook*. Berlin, New York: Walter de Gruyter, 646–659.

RÉSUMÉ

À la recherche d'une unité de l'euphonie : Une étude de cas de la poésie tchèque des années 1830

Cet article étudie des échantillons de la poésie tchèque des années 1830 à partir d'une conception quantitative de l'euphonie. Se fondant sur les réflexions de Mukařovský sur ce sujet, il tente de rendre cette notion plus efficace, en fournissant des chiffres exacts à valeur intersubjective. La recherche de ce phénomène repose sur la formule proposée par Gabriel Altmann, qui considère l'euphonie comme une répétition signifiante d'une unité dans un vers. Ce travail considère que cette unité cherchée est un ensemble consonantique, c'est-à-dire un regroupement de consonnes entre deux voyelles, au début d'un vers, ou à la fin. Cette innovation essaie de tenir compte de la réalité phonétique de la langue de façon plus efficace que le son, utilisé d'habitude. L'analyse confirme le caractère raisonnable de l'unité nouvelle, montre les relations mutuelles entre les deux calculs, et esquisse quelques interprétations de la situation dans la poésie tchèque des années 1830. La comparaison statistique montre que les deux approches sont indépendantes l'une de l'autre, mesurant des propriétés différentes du texte. Les poètes Karel Hynek Mácha, représentant éminent du romantisme tchèque, et Josef Jaroslav Langer, son contemporain à orientation plutôt folklorique, sont les auteurs dont les vers sont les plus euphoniques ; tandis que dans le cas de Langer, on relève des pourcentages élevées de vers euphoniques, Mácha se focalise sur la hauteur de la valeur euphonique d'un nombre limité de vers.

Department of Czech Language, University of Ostrava
Reální 5, 701 03 Ostrava, Czech Republic
<mmistecky@seznam.cz>

Department of Czech Language, University of Ostrava
Reální 5, 701 03 Ostrava, Czech Republic
<cechradek@gmail.com>

Institute of Czech Literature, Czech Academy of Sciences
Na Florenci 1420/3, 110 00 Praha 1, Czech Republic
<plechac@ucl.cas.cz>