Towards the Automatic Analysis of Czech Verse[1]

Robert Ibrahim
Petr Plecháč

The objective of the following paper is to present a project of software development that should be suitable for metrical (-rhythmical) analysis of Czech verse.[2]

In Czech and Slovak verse theory the formal approaches appeared in the 1960s.[3] Research of that period focused mainly on statistical analysis[4] using contemporary machines (punched cards and their automatic analysis). The idea of a fully computerized description and scansion of verse also appeared (Sgallová 1964). For both internal and external reasons the effort was abandoned in the late 1960s.

After 1989 the idea of using a computer for purposes of researching versification was resumed for various reasons (political, technological). The main task, however, was to make the manually analysed material publicly available and to make it possible to search through such material using computer programs (cf. Sgallová 1999). Thanks to the great effort of Miroslav Červenka and Květa Sgallová a complete database exists with metrical annotation of Czech literary verse written between 1795 and 1825; this database is accessible via the internet (the so-called Thesaurus of Czech meters at the website of the Institute of Czech Literature of the Czech Academy of Sciences). Most of the 19th century material, however, is unavailable in this way. During their research the scholars just mentioned processed a huge amount of data, but the results of their analysis are not all accessible in one place (they were published separately or remain unpublished).

It is self-evident that research of 19th-century Czech poetry must be based on thorough knowledge of verse theory and that such knowledge must be based on analysis of data that are as extensive as possible in order to be reliable. It should be recalled that even Jan Mukařovský (1891–1975) – the icon of the Czech structuralism – based his thesis on a small number of scanned lines

1 The research presented in this paper is kindly supported by Grant from Czech Science Foundation (P406/11/1825).

2 We are grateful to Vladimír Petkevič, Petr Pořízka, and Dan y altýnek for helpful comments in the preparation of this paper.

3 The foundations of such approaches come from the interwar era – they can be found in the works of structuralist scholars inspired by Russian theories and groundbreaking studies by B. V. Tomashevskii.

4 Cf. Levý (1964 a, b, 1971), Červenka 1965, Červenka–Sgallová 1967, Červenka 1971, Štraus 2002.

(usually 100), which led to many inaccuracies and errors.

The huge amount of unprocessed data from the 19th century (or data processed in the way discussed above) implies the question whether one person's lifetime is enough to process it. And it is evident that a different solution needs to be found. This involves the utilization of computer technology – in the development of a program capable of automatically identifying a poem's meter. This is where we differ from previous attempts at analysis of Czech verse, where the computer was used only to process manually analysed data.

In the first stage we will make the program's metrical identification of Czech syllabotonic 19th-century verse as precise as possible and we will try to finalize (in this respect at least) the project started by Miroslav Červenka and Květa Sgallová. In the next stages we would like to extend the program functions to the automatic analysis of rhythm, rhyme types and schemes, and stanzas. Moreover, we would like to extend the data to (a) other types of verse (Czech quantitative verse) and possibly free verse, (b) poetry of other historical periods (before and after the 19th century).

Let us proceed to a description of the program (KVĚTA 2.1).

PROSODIC ANALYSIS

The initial subroutine of prosodic analysis is based on the notion that in Czech the first syllable of every polysyllabic unit (monosyllabic prepositions are treated as part of a following unit) and some monosyllables are stressed. Thus, in the first step the program identifies particular syllables and marks them as being (1) the first syllable of a polysyllabic unit (marked as "1"), (2) a non-initial syllable of a polysyllabic unit (marked as "0"), (3) a monosyllable (marked as "+"), or (4) a monosyllabic preposition.[5] In the last case the syllable following the preposition is marked as being the non-initial syllable of a polysyllabic unit ("0") and the preposition itself is marked as the initial syllable ("1") of a polysyllable. For example the line

Za trochu lásky šel bych světa kraj

[P][xx][Xx][X][x][Xx][X][6]

5 Particular problems (that is ambiguity of some prepositions and nouns) are examined in cooperation with the Institute of Theoretical and Computational Linguistics, Faculty of Arts, Charles University.

6 In the actual rhythmical notation "X" stands for a stressed syllable, "x" for a non-stressed syllable and "P" for the stress of a monosyllabic preposition. Square brackets represent word-unit boundaries. Thus [x] represents an unstressed monosyllable, [X] a stressed monosyllable, [Xx] a disyllable, and [Xxx] a trisyllable. Sequences containing a monosyllabic preposition are respectively marked as [P][x], [P][xx], or [P][xxx].

is marked as the sequence “10010++10+”. From now on we will refer to such sequences as “primary strings”.

METRICAL ANALYSIS

The metrical analysis subroutine is based on the generative rules of Czech syllabotonic verse proposed by Miroslav Červenka (2006). They were simplified and implemented by the following algorithms ($n \in N^0$):

(1) A line is iambic (I) if no odd position except the first[7] is occupied by the stress of a polysyllabic unit .

(2) A line is trochaic (T) if no even position is occupied by the stress of a polysyllabic unit.

(3) A line is dactylic (D) if

(a) no $3n+3^{rd}$ position is occupied by the stress of a polysyllabic unit and

(b) no $3n+2^{nd}$ position is occupied by the stress of a unit consisting of three or more syllables and

(c) every stress occupying a $3n+2^{nd}$ position is preceded by a monosyllabic unit

.

(4) A line is dactylic with anacrusis (Da) if

(a) the first position is occupied by a monosyllabic unit and

(b) no $3n+4^{th}$ position is occupied by the stress of a polysyllabic unit and

(c) no $3n+3^{rd}$ position is occupied by the stress of a unit consisting of three or more syllables and

(d) every stress occupying a $3n+3^{rd}$ position is preceded by a monosyllabic unit

(5) A line is dactylo-trochaic (DT) if a “virtual syllable”[8] can be inserted into the line after some of the $3n+2^{nd}$ positions (at least once) in order to meet the conditions specified in (3).

(6) A line is dactylo-trochaic with anacrusis (DTa) if the “virtual syllable” can be inserted into the line after some of the $3n+3^{rd}$ positions (at least once) in order to meet the conditions specified in (4).

---

7 By this exemption we allow so called “dactylic incipits” (Xxx...) which represent a statistically significant part of the realisation of iambic verse in Czech poetry.

8 “Virtual syllable” is a purely theoretical construct. The realization of the positions is analysed from the beginning of the line. The order of the position after which the “virtual syllable” is inserted is specified not only by the sum of preceding “non-virtual positions”, but also by the sum of preceding “virtual positions”.

The algorithms just listed do not subsume the following rules of Červenka:

K4(d) The strong position of [a dactylic] meter is stressed.
K7(t) The first position of a trochee is stressed.
K8 The weak positions that follow the last strong one are unstressed.
KO2 (d) If the strong position [of a dactylic meter] is not stressed, the preceding weak position is not stressed either.
KO3 (i,t) The weak position [of iambs and trochees] can be stressed only when the following strong position is also stressed.

Rules K4(d) and K7(t) are descriptive rather than prescriptive. They do not mirror the constraints that are imposed on speech by meter, but the situation that arises from the nature of the Czech language itself (see Plecháč 2008). Hence subsuming them into algorithms would be redundant. Červenka himself calls KO3(i,t) questionable since in binary meters it rules out the configurations of a stressed monosyllable occupying the weak position followed by an unstressed one ([X][x]). According to Červenka, the occurrence of such configurations in verse is less frequent than it is in prose, but they stand on the boundary between metricality and unmetricality. The output of the program achieved so far has confirmed the hypothesis based on language probability, namely that stresses in a polysyllable are frequent enough to distinguish binary meters from one another. The conditions stated in (1) and (2) – a stressed monosyllable occupying the weak position of a binary meter (I,T) is allowed no matter what follows – are therefore sufficient to distinguish the iamb from the trochee, hence KO3 need not be subsumed into the program. The planned incorporation of rhythmical analysis into the program should allow us to verify or relativize its impact by comparing the frequency of given configurations in verse and in prosaic speech. By analogy, we omit KO2 as well – a simple condition that monosyllables may occupy weak positions seems to be sufficient also in ternary meters (D, Da). Finally we also omit K8. Distinguishing iambic verse from trochaic verse is sufficiently provided by internal positions in the line and in distinguishing binary meters from ternary meters so that K8 is not relevant – in the case of ternary meters its application is blocked by a rule stating "when the final position of dactylic meters is occupied by two weak positions, unstressed syllables are **preferred**" (Červenka 2006: 118; the emphasis is ours). However, we will see later that a technical barrier, which does not allow us now automatically distinguish stressed monosyllables from unstressed monosyllables, may result in our inability to recognize distinctions among some rare types of dactylo-trochee.

All the algorithms listed so far are sufficient to distinguish (most of the) meters from one another, but not to identify them correctly. Poets, of course, do not produce purely metrical lines only, but also lines violating given constraints. The most frequent unmetrical type is "prepositional transaccentation" — that is, treating the sequence of a monosyllabic preposition and another unit as having the stress located not on the first syllable (preposition) but on the second one. Thus in the next step the program takes the primary strings as the input, relocates the first prepositional stress in the line to a subsequent syllable, and repeats this process by ignoring the prepositions thus found up to the end of the line thus creating alternative strings (I). In the same fashion this relocates the stress of a polysyllabic unit (incidental prepositions included) to the immediately preceding monosyllable (if there are any) (II) – that represents another frequent type. All these alternative strings are analyzed by the algorithms listed above. In the last step (III) primary strings are analyzed once again with toleration of one polysyllabic unit's stress occupying one of the forbidden positions.

Let us demonstrate the preliminary results. The following tables contain the results of the analysis of a quatrain from Karel Hynek Mácha's "Máj" – a poem in which iambic constraints are frequently violated in various ways. The row "zero" shows the primary strings and rows I–III show the alternative strings just mentioned. Elements of the strings that violate constraints of a given meter (columns I/T/D/Da; DT and DTa are not yet included) are emphasized. The first string in a column that does not violate the constraints of a given meter or violates it only once (row III) is underlined.

Zhasla měsíce světlá moc,
[Xx][Xxx][Xx][X]

	I	T	D	Da
0	10**1**0010+	10100**1**0+	10**1**00**1**0+	**1**0**1**00**1**0+
I	10**1**0010+	10100**1**0+	10**1**00**1**0+	**1**0**1**00**1**0+
II	10**1**0010+	10100**1**0+	10**1**00**1**0+	**1**0**1**00**1**0+
III	10**1**0010+	10100**1**0+	10**1**00**1**0+	**1**0**1**00**1**0+

output→I/T

i hvězdný svit a kol a kol
[x][Xx][X][x][X][x][X]

	I	T	D	Da
0	+10+++++	+**1**0+++++	+10+++++	+10+++++
I	+10+++++	+**1**0+++++	+10+++++	+10+++++
II	100+++++	100+++++	100+++++	**1**00+++++
III	+10+++++	+**1**0+++++	+10+++++	+10+++++

output→I/T/D/Da

je pouhé temno, širý dol
[x][Xx][Xx][Xx][X]

	I	T	D	Da
0	+101010+	+**1**0**1**0**1**0+	+1010**1**0+	+10**1**0**1**0+
I	+101010+	+**1**0**1**0**1**0+	+1010**1**0+	+10**1**0**1**0+
II	1001010+	100**1**0**1**0+	10010**1**0+	**1**00**1**0**1**0+
III	+101010+	+**1**0**1**0**1**0+	+1010**1**0+	+10**1**0**1**0+

output→I/D

co hrob daleký zívá.
[x][X][Xxx][Xx]

	I	T	D	Da
0	++**1**0010	++100**1**0	++**1**00**1**0	++**1**00**1**0
I	++**1**0010	++100**1**0	++**1**00**1**0	++**1**00**1**0
II	<u>+100010</u>	+**1**000**1**0	+**1**000**1**0	+1000**1**0
III	++**1**0010	<u>++100**1**0</u>	++**1**00**1**0	++**1**00**1**0

output→I/T

Although every line was correctly recognized as iambic, three quarters were also marked as T, one half as D, and one line as Da. Although the configuration represented by the second line can occur also in trochees, it is without doubt more iambic than trochaic because it does not violate the constraints of the former one at all while constraints of the latter are violated by the stress of a polysyllabic unit occupying a weak position ("hvězdný" [Xx]). On that basis we follow our Moscow colleagues and establish the magnitude "metrical index" (MI)[9] that quantifies the degree of metricality in a line according to particular meters from 0 to 100. The calculation of MI is as follows:[10]

- If a line is marked as I/T/D/Da/DT/Dta in any step, the meter is assigned the value of MI 100.
- Every stress of a disyllabic unit according to 3b/3c and 4c/4d except for the second position[11] in a line lowers MI by 20.
- If the line is marked as far as step I ("prepositional transaccentation"), MI is lowered by 40
- If the line was marked as far as step II (transaccentation to the foregoing monosyllable), MI is lowered by 60
- If the line was marked as far as step III (one stress of a polysyllabic unit occupying the weak position is tolerated), MI is lowered by 80
- Incidental negative values are annulled

The meter with the highest arithmetic mean of MI, on the condition that the value is higher than 50 and MI of no line is zero for a given meter, is the (virtual) final output of the program. Thus analysis of Mácha's quatrain discussed above gives the following results:

---

9 See Pilshchikov–Starostin 2008.

10 Our calculation differs from that of Moscow scholars and is considered to be only the first approximation. Values in (2–5) were established de facto intuitively and presented for discussion.

11 This way so called "iambic incipits" in dactylic verse are allowed.

	I	T	D	Da
Zhasla měsíce světlá moc,	20	20	0	0
i hvězdný svit a kol a kol	100	40	100	100
je pouhé temno, širý dol	100	0	20	0
co hrob daleký zívá.	40	20	0	0
arithmetic mean	65	20	30	25

Relatively low variation of final values does not result from the program's ineffectiveness – it mirrors the irregularities of Mácha's verse. Let us present for comparison, the analysis of Jaroslav Vrchlický's "regular" iambs:

	I	T	D	Da
Rád slýchám z jara bouřné větry váti. [X][Xx][Xx][Xx][Xx][Xx]	100	0	0	0
Má duše hne se, jak by chtěla bráti [x][Xx][X][x][X][x][Xx][Xx]	100	0	80	20
se s labutěmi, které z jihu spějí, [x] [Xxxx][Xx][Xx][Xx]	100	0	0	0
a jak by chtěla v jejich vírném reji [X][x][x][Xx][Xx][Xx][Xx]	100	0	0	0
nad step, jež dřímá, nad moře, jež duní, [P][x][x][Xx][P][xx][x][Xx]	100	0	60	0
se vznést a volat: Nesu jaro s vůní, [x][X][x][Xx][Xx][Xx][Xx]	100	0	0	0
arithmetic mean	100	0	23,3	3,3

MULTIMETRICAL LINES AND MULTIMETRICAL POEMS

All the algorithms listed above indicate that, according to our approach, there are no lines which do not correspond to a metrical pattern even in the strictest version (step "zero" in which transaccentations are not considered)[12] and that there are lines which correspond to more than one metrical pattern to an equal degree. The latter will be referred to as "multimetrical lines" and – as was just indicated

[12] If the stress follows any 2nd position in T, any 3rd position in D and DT results from a combination of both. It is possible to stress the $2a+3b+1^{st}$ position in at least one of these meters ($n, a, b \in N^0$), that is, any position except the second can be stressed, because for any odd number x, the equation $x = 2a+3b+1$ (where $a = (x+1)/2$; $b = 0$) can be applied. For any even number y greater than 2, the equation $y= 2a+3b+1$ (where $a = (x-4)/2$; $b = 1$) can be applied. The second position can be stressed in meters which differ from the three meters mentioned above by adding a weak position at the beginning of the pattern (I, Da, DTa).

in the case of Mácha – their final interpretation is determined by their context (the second line "i hvězdný svit a kol a kol" fits I, D, and Da to a full extent – its final iambic interpretation is determined by MI in the other lines).

Let us consider, however, a poem in which all lines are multimetrical. Especially short iambic and dactylic meters are involved here; the iamb and dactyl are both distinguished from the trochee in the very beginning of a line while the iamb is not distinguished from the dactyl only in the internal positions of the line (i.e. positions that belong neither to the first, nor to the last "foot"). Such "metrical ambiguity" at the beginning of the line is caused by the fact that dactylic incipits in iamb or iambic incipits in dactyl verse are allowed. Such ambiguities can be resolved by a rather improbable situation where the second position is occupied by the stress of a unit consisting of three or more syllables (in such a case a dactylic interpretation is excluded; see 3b). The end of a line is also "metrically ambiguous"; such an ambiguity affects not only the iamb and the dactyl, but also the trochee. Configurations X and Xxx can conclude a masculine iambic line, a masculine trochaic line and the acatalectic dactylic line; and configurations Xx can conclude the feminine variant of all three types. Short meters lack the distinctive internal line positions and thus the whole poem often does not contain a configuration which could determine the meter explicitly. For example, number eight of Vítězslav Hálek's "Večerní písně"

Kdyby ten slavíček	[Xx][x][Xxx]
s svou družkou přebýval,	[x][Xx][Xxx]
on by tak žalostně	[X][x][x][Xxx]
o lásce nezpíval.	[P][xx][Xxx]
Kdyby to srdéčko	[Xx][x][Xxx]
s Tebou noc probdělo,	[Xx][X][Xxx]
však by tak bolestně	[X][x][x][Xxx]
v těch ňadrech neznělo	[x][Xx][Xxx]

fits the pattern of masculine iambic trimeter and the pattern of acatalectic dactylic dimeter equally. Multimetricality of all lines leads to the multimetricality of the poem:[13] it is iambic and dactylic at the same time.

Almost all the iambic, trochaic, and dactylic lines are theoretically multimetrical according to the dactylo-trochee as well. Hálek's poem quoted above could also be interpreted as DT3m (DTT) (masculine dactylo-trochaic trimeter consisting of one dactylic "foot" and two trochaic "feet"). Any T[n]m (masculine

---

[13] Cf.: "Let us now imagine or find a real example of what Petr Rudnev called a "transitional metric form": a poem consisting of the lines of both meters where one meter predominates [Rudnev 1972: 227]. If some lines in such a poem are bimetrical, this becomes their rhythmic as well as metric feature. The probability of such a line to be amphibrachic or iambic is equal." (Pilshchikov–Starostin 2011)

trochaic n-meter) could be interpreted as DT[n–1]m with a final dactylic "foot"; any D[n]f (feminine dactylic n-meter) could be interpreted as DT[n]f with a final trochaic "foot". In such a case, a combined metrical interpretation would, however, be based on the fulfillment of positions that are not distinctive in given meters (i.e., the line ending in T/D, and line beginning and line ending in I/D). This could be solved by introducing a rule excluding dactylo-trochees that contain (1) a ternary substring (SWW) only at the beginning and/or at the end of a metrical pattern, or (2) a binary substring (SW) only at the end of a metrical pattern. In such a case, however, we would exclude also the dactylo-trochaic types that are without doubt part of a reader's metrical perception, that is, DT, TD, and DTD. In them the lack of distinctive positions or their low number leads to real multi-metricality according to dactylo-trochee (ternary alternation is competing with binary alternation and vice versa in at least one half of the line). Another type is represented by poems where the exclusion of dactylo-trochaic interpretation leads to unjustified polymetry. Out of context we would interpret Eliška Krásnohorská's lines "pomněnce vtisklo za znamení / o prvním, prvním políbení" ([Xxx][Xx][P][xxx]/[P][xx][Xx][Xxxx]) as iambic, but it would be unreasonable to do so since they occur in the vicinity of the lines "připloulo jaro, připloulo z jihu / jalo se básnit o lásce knihu" ([Xxx][Xx][Xxx][Xx]/[Xx][x][Xx][P][xx][Xx]). Instead of a monometric poem combining feminine and masculine endings in the same meter (DT4m(DTDT)/DT4f(DTDT)) we would, in such a case, interpret the poem as a polymetric combination of dactylo-trochee and iamb (I4f/DT4f(DTDT)), which is, as we said, unreasonable.[14]

The last "superfluous ambiguity" concerns the ending of dactylo-trochaic lines. Should we, for example, consider Jan Neruda's line "když k Vám vesel hledím, zlatá vy Kuřátka" ([X][x][Xx][Xx][Xx][x][Xxx]) as DT5a(TTTDD) (acatalectic dactylo-trochaic pentameter consisting of three trochaic "feet" and two dactylic "feet"), or as DT6m(TTTDTT)? Should we consider the sequences [Xx][Xxx][Xxxx] as DT4m(TDDD), DT4m(TDDT), or DT4f(TDTT)? We propose to give preference to a meter ending with a complete "foot" (DT5a, DT4f), rather than to a meter ending with an incomplete "foot" (DT6m, DT4m). The second type could, however, possibly contain the sequence [Xx][Xxx][Xxx][X] which, according to Červenka's rules, excludes the acatalectic interpretation ("Weak positions that follow the last strong position are unstressed" (Červenka 2006: 117). As we mentioned earlier, we are still not capable of distinguishing stressed monosyllables automatically from unstressed monosyllables. On the other hand, it is easy to suppose that these types occur

[14] With such types of dactylo-trochee excluded, the program would actually not mark the poem as polymetric, but would not reveal the meter at all. The algorithm based on the arithmetical means of MI can discern, of course, only monometric poems. We will try to solve the question of polymetry in the future in connection with questions about rhyme.

with extremely low frequency (if they occur at all). Hence we apply the last rule as an algorithm whose results can be verified manually. Thus the algorithms for dactylo-trochees are as follows:

(1) Preference is given to the meter that can be applied to the entire poem (monometry).
(2) Preference is given (a) to the meter in which alternation of strong and weak positions is regular or (b) to the meter containing at least as many consecutive substrings in which strong and weak positions alternate with a different frequency than in a meter chosen by (a) because it contains consecutive substrings in which strong and weak positions alternate with the same frequency as in the meter chosen by (a).
(3) Precedence is given to a meter ending with a complete "foot".

It would be premature to consider whether analysis of a larger corpus of data will produce different results (in regard both to the general characteristics of Czech verse and to the individual characteristics of authors, generations...) than the ones achieved by Miroslav Červenka and Květa Sgallová (or other scholars). What we can already state is that the development of a program may result not only in important discoveries about the history of Czech verse,[15] but it can also influence its theory. We should find which configurations represent the dactylo-trochee in Czech poetry and which do not. We also should reinterpret the rules of Czech syllabotonic verse.[16]

15 Automatic analysis differs from manual analysis by the lack of pre-selection (which in the latter is based on a scholar's experience resulting from a proportional occurrence of particular types). Thus we can expect several "surprises".

16 The disadvantage of the algorithm presented here is that it does not take account of the line's length and that it does not appreciate agreements of metrical and rhythmical pattern only penalizing discrepancies (e.g. a decasyllabic trochaic line which is entirely regular in terms of alternation of stressed and unstressed syllables except for one unmetrical stress (XxXxXxxXxx) is assigned the very same MI as a short trochaic line in which no strong position is occupied by a stressed syllable and no weak position is occupied by an unstressed one (xXx). Nowadays a new algorithm which takes account of both these issues is being examined. Every syllable is assigned the value h+ varying from 0 to 1 according to the degree of agreement with metrical pattern (1 for stressed syllable occupying strong position / an unstressed syllable occupying a weak position, 0.5 for an unstressed syllable occupying a strong position, 0 for a stressed syllable occupying a weak position) and the value h- varying from 0 to 1 according to the degree of metricality (1 for metrical realization, lower values for stressed syllables occupying weak positions). MI of a m-syllabic line is calculated as the arithmetic mean of h+ of every syllable multiplied by the product of h- of every syllable. The MI of a poem consisting of n lines is then calculated as the arithmetic mean of MIs of particular lines.

References

Červenka, Miroslav (1965). Nový projekt statistického rozboru verše. *Česká literatura 13, 541–544.*

Červenka, Miroslav (1971). *Statistické obrazy verše.* Praha: ÚČSL.

Červenka, Miroslav (2006). *Kapitoly českém verši.* Praha: Karolinum.

Červenka, Miroslav; Sgallová, Květa (1967). On a Probabilistic Model of the Czech Verse. In: L. Doležel, P. Sgall, M. Těšitelová, J. Vachek (eds.), *Prague Studies in Mathematical Linguistics* 2*, 105–120*. Praha: Academia.

Levý, Jiří (1964a). Matematický a experimentální rozbor verše. *Česká literatura 12, 181–213.*

Levý, Jiří (1964b). Předběžné poznámky k informační analýze verše. *Slovenská literatúra 9, 15–37.*

Levý, Jiří (1971). Matematické aspekty teorie verše. In: M. Červenka (ed.), *Bude literární věda exaktní vědou?, 264-288*. Praha: ČS.

Pilshchikov, Igor; Starostin, Anatoli (2011). Automated Analysis of Poetic Texts and the Problem of Verse Meter. In: Ch. Küper (ed.), *Current Trends in Metrical Analysis,* 133-140. Bern et al.: Peter Lang.

Plecháč, Petr (2008). Česká versifikace a generativní metrika. *Aluze 11, 86–93.*

Sgallová, Květa (1964). Vy užití moderní techniky při rozboru verše. *Česká literatura 12, 158-165.*

Sgallová, Květa (1999). Thesaurus českých meter. *Česká literatura 47, 286–289.*

Štraus, František (2002). *Základy informačnej analýzy verša.* Bratislava: UK