

Sentences and clauses in Jan Neruda's Tales of the Lesser Quarter: a stylometric analysis

MARTIN SCHACHERL

Jihočeská univerzita v Českých Budějovicích
ORCID: <https://orcid.org/0000-0002-2719-8075>
E-mail: schacherl@pf.jcu.cz

Abstract: The study provides a quantitative syntactic analysis of Jan Neruda's *Povídky malostranské* (*Tales of the Lesser Quarter*) and seeks to empirically verify the traditional literary-historical characteristics of his style. Using stylometric indicators, it examines sentence length, the ratio of simple sentences to compound sentences, the relationship between the narrator and the characters, the degree of parataxis and hypotaxis, and the horizontal structure of the text. The results demonstrate the stability of Neruda's syntactic model across the collection. Parataxis is the dominant principle, while hypotaxis serves more as a means of semantic differentiation. The differences between simple sentences and compound sentences are mainly quantitative in nature. The horizontal division of the text does not manifest itself as a factor of syntactic fragmentation, but rather as a means of compositional organisation and thematic clarity. The stylometric data thus corrects the idea of the fragmentary nature of Neruda's syntax and confirms its structural coherence.

Keywords: Jan Neruda; Prague Tales from the Lesser Quarter; quantitative stylistics; syntax; sentence and clause structure; authorial style; 19th-century Czech prose



Bohemistika, 2026 (3) © The Author(s). Published by: Adam Mickiewicz University in Poznań, Open Access article, distributed under the terms of the CC Attribution-NonCommercial-NoDerivatives 4.0 International licence (BY-NC-ND, <https://creativecommons.org/licenses/by-nc-nd/4.0/>)

This study focuses on the language and style of *Povídky malostranské* [*Tales of the Lesser Quarter*] by Czech writer Jan Neruda (1834–1891), specifically on the syntactic level of the text with regard to simple sentences and compound sentences. Its aim is to objectify traditional assessments of the poet's language and style and to empirically verify the degree of stylistic (in)consistency of the collection by comparing quantifiable linguistic parameters, i.e. situations where stylistic features in the text function comprehensively – either in harmony with each other (the style of the text is consistent) or in contrast (the style of the text is unintentionally or intentionally inconsistent). Stylometric analysis aims to verify claims about the author's individual style and to assess the extent to which they are based on verifiable features of the syntactic structure of the text. The basic research question of the study is whether the syntactic structure of *Povídky malostranské* creates a stable authorial model of sentence construction, or whether the individual texts show more significant stylistic variability. The analytical starting point is the quantification of selected linguistic features of Neruda's composition; the study does not attempt a comprehensive analysis of sentences and clauses, but focuses on selected aspects of the topic under investigation within the limits of contemporary stylistic research. Quantitative stylistics methods were used to achieve the set objectives; the QuitaUp application was used to capture selected text features, which allows the calculation of stylometric indicators that quantitatively express selected text features. Frequency analyses of parameters for simple sentences and compound sentences were performed manually by the author.

The collection *Povídky malostranské* (1878; hereinafter PM) belongs to the canonical works of the golden fund of Czech literature, to texts with time-tested aesthetic quality and a permanent presence in the national educational tradition. It consists of thirteen prose works in which Jan Neruda thematises the environment of Prague's Malá Strana in the first half of the 19th century as a socially and linguistically differentiated microcosm. The narrative focuses on everyday life in a bourgeois environment and on distinct characters, whose charac-

teristics are developed through their actions and individualised language, which is clearly reflected in the sentence and clause structure of the text. Compositionally, the collection alternates between relatively closed novella structures and episodic, mosaic-like compositions. The texts were written gradually and were first published in magazines (*Národní listy*, *Lumír*, *Květy*, *Podřipan* in 1875–1877); the older short story *Týden v tichém domě* (A Week in a Quiet House, 1867) was also included in the book edition.

The language and style of Neruda's prose have been the subject of systematic attention from literary critics and historians since the beginning of its reception, mainly as part of broader literary-historical syntheses (Novák, 1910, 1919; F. X. Šalda, 1934; Polák, 1947; Arbes, 1952; Haman, 1968, 2014; Křivánek, 1983; Mocná, 2012), less often as a separately defined subject of stylistic or linguistic analysis (Stich, 1974; Hausenblas, 1977, 1984; Bartůňková, 1985; Macurová, 1991; Králík, 1995; Plecháč & Flaišman, 2017). The symptomatic nature and interpretative relevance of selected aspects of Neruda's work have been repeatedly discussed, but mainly at the level of evaluative judgements rather than systematic quantitative analysis. Emblematic in this respect is the statement by F. X. Šalda (1937, p. 122), characterising the form of Neruda's prose as "mostly negative, destructive" and "unconstructive", giving the "impression of a pile of gravel", because his prose "is not epic, it is not malleable, it is not constructive". Aleš Haman (1968, pp. 33–34) takes the opposite view, arguing that syntax is "the dominant component of the textual layer" and at the same time "the bearer of the basic principle – the conflict between subject and object – on which the poetic quality of Neruda's prose is built". These extreme examples illustrate, on the one hand, the polarity in the assessment of the significance of the poet's sentences and clauses and, on the other hand, their broad interpretative potential in the interpretation of Neruda's prose.

Aleš Haman identifies two characteristic features of Neruda's composition: *paratacticity* and *parentheticality*. He interprets paratacticity as "free composition", i.e. a tendency to divide sentences into

smaller segments (often using dashes and semicolons). According to Haman, in these cases the expected boundaries between sentences and clauses are blurred: "where we would expect first-degree breaks (at the end of sentences and clauses), the break is often less pronounced (compare the frequent use of semicolons, which is completely different from the structure of classical periods in Mácha or Palacký), and conversely, breaks are often less prevalent than breaks in places of greater intonational significance". Neruda's tendency towards paratactic diction is also supported by Oldřich Králík, who describes Neruda's narrative style as a "paragraph technique" (Králík, 1995, p. 294): the sentence is not a fully autonomous unit, but is functionally subordinate to a higher compositional unit, the paragraph.

According to Haman, the second distinctive feature of Neruda's composition is his tendency towards parentheticality. Haman describes this tendency towards parentheticality as "the most important and enduring syntactic feature of Neruda's sentences"; in places, he argues, this feature is so prominent that it "gives the impression of an insertion even where this syntactic phenomenon is clearly not the case" (2014, p. 41). He characterises this tendency as a disruption of the immediate syntactic context through parenthetical insertions and inserted constructions. The parenthetical disruption of linear connection does not weaken the cohesion of the text, but rather leads to its internal differentiation and a denser semantic structure; in the context of *Povídky malostranské*, this approach is interpreted as one of the means of creating a multi-layered, functionally contrasting picture of reality (Haman, 1958, p. 55). The boundaries of sentence units are thus weakened in favour of a distinct internal structure, which is reflected in a fragmented intonation line. The cause of this effect lies in the syntactic structure based on the frequent use of parentheses, apposition constructions and free attributive combinations, which – both in terms of punctuation and their position in the linear chain – initiate the need for more pronounced horizontal division. Neruda's expression therefore gives the impression of being interrupted; Haman sees its source in the effort to continuously refine and deepen the content of the message by

inserting new aspects and details into the basic semantic idea. The result is a “combined multiple deepening of the basic idea” (Haman, 2014, pp. 41–42). A similar emphasis on the breakdown of a continuous storyline can be found in Arne Novák, who believes that Neruda remained a “master of distinctive detail” (1919, p. 35), as he consistently breaks down the narrative into a sequential list of static motifs instead of developing them continuously.

It follows from the above characteristics that the paratactic and parenthetical structure of the sentence became established in contemporary and later reflections as a key feature of Neruda’s composition. These characteristics, which have so far been formulated mainly in interpretative terms, form the starting point for the analysis presented in this study. Its aim is to empirically refine and comparatively verify traditional claims about Neruda’s syntax by quantifying selected identifiers of his singular style at the linguistic level examined in *Povídky malostranské*. The study seeks to objectify intuitively named stylistic dominants by systematically measuring the following indicators: linear sentence and clause length, the ratio of hypotaxis to parataxis, the distribution of the narrator’s and characters’ speech, the degree of activity and descriptiveness of the text, the frequency of parentheses, and the size and internal structure of paragraphs. The analysis examines how the declared tendency towards paratacticity (free composition) and parentheticality (decomposability, fragmentation) is reflected in the specific construction of simple sentences and clauses, as well as in the hierarchisation of syntactic relationships.

The comparisons also take into account the genesis of individual Little Quarter stories, their thematic, genre and stylistic differentiation, and examine the possible linguistic conditioning of the phenomena under investigation, i.e. whether changes in syntactic structure are linked to the time of the text’s creation, the type of material or the narrative perspective. Special attention is paid to the possibility that the selected parameters of the composition reflect elements of journalistic style, in other words, the “fusion of fiction and journalism” (Mocná, 2012, p. 32) that is often associated with Neruda’s work. The analysis

therefore does not only look at the isolated values of individual indicators, but also their mutual relationships and configurations within specific texts. The aim is to demonstrate whether and how the traditionally cited features of Neruda’s composition – a tendency towards paratacticity and a propensity for parentheticality – are systematically reflected in the syntactic structure of simple sentences and compound sentences, whether they are actually and regularly realised at the linguistic level and whether they are quantitatively verifiable and statistically measurable, or whether they remain predominantly an interpretative construct of the literary-critical tradition; or whether they show developmental dynamics or stylistic variability within the corpus.

Analysis of selected stylometric indicators¹

Lemmata N = 21740–3071–2215–1639–3182–1956–1722–1501–649–3683–4885–3243–19996; V = 4703–1086–891–685–1130–793–729–648–344–1309–1669–1137–4029; TTR = 0,216–0,354–0,402–0,418–0,355–0,405–0,423–0,423–0,53–0,355–0,342–0,351–0,201; 1) Hapaxy = 2887–712–624–474–744–535–487–452–237–868–1113–727–2445; 2) H = 9,651–8,723–8,544–8,303–8,677–8,49–8,476–8,296–7,747–8,828–9,051–8,701–9,389; 3) VD = 5,284–6,179–5,512–5,886–5,05–5,731–5,615–6,338–5,465–6,142–5,098–5,245–4,631; 4) Q = 0,624–0,558–0,622–0,571–0,637–0,571–0,58–0,586–0,565–0,539–0,621–0,613–0,696; 5) D = 0,376–0,442–0,378–0,429–0,363–0,429,42–0,414–0,435–0,461–0,379–0,387–0,304; 6) TC = 0,0357–0,0845–0,101–0,0721–0,034–0,0667–0,0638–0,0892–0–0,0031–0,0193–0,0379–0,0328.

¹ The QuitaUp application was used to calculate stylometric indicators that quantitatively express selected text characteristics. The specific outputs are: word token frequency (N) expresses the total number of all tokens in a given text (a token is a graphic word, i.e. a sequence of characters separated by spaces), word type frequency (V) expresses the absolute frequency of all types in a given text (a type is an abstract unit and a token is its concrete realisation. The number of types therefore corresponds to the number of different words in the text) and the ratio of types to tokens (TTR) expresses the concentration of the vocabulary of a given text (this is a basic index of so-called lexical richness). Language characteristics monitored: 1) Frequency of hapaxes (legomenon), i.e. words that appear only once in the text (the ratio of hapaxes to tokens expresses how large a proportion of the text is made up of hapaxes. This value depends on the length of the text – the longer the text, the

The collection *Povídky malostranské* comprises thirteen short stories with a total length of 208,446 tokens.² The longest texts are *Týden v tichém domě* (21,740 tokens) and *Figurky* (19,996 tokens), which are significantly longer than the other stories in the collection. The third longest story is *Jak to přišlo* (4,885 tokens). The shortest text is *U Tří lilií* (649 tokens), followed by the stories *Jak si nakouřil pan Vorel pěnovku* (1,501 tokens) and *O měkkém srdci paní Rusky* (1,639 tokens). The average length of a short story in PM is 5,345 tokens. The collection thus includes short and medium-length prose, with a considerable range between individual texts. However, there is no direct correlation between the length of the text and its final composition, nor between its length and its inclusion in the collection. The short sto-

smaller the proportion of hapaxes), 2) Entropy (H) is generally understood as a measure of the uncertainty of a system. We understand entropy as a value expressing the degree of diversity of the vocabulary – the higher the entropy value, the more diversified (i.e. less concentrated) the vocabulary is, so a high entropy value is a sign of great lexical richness. The entropy value depends on the length of the text, 3) Verb distance (VD) is the arithmetic mean of the number of tokens between two consecutive verbs in the text, excluding auxiliary words, 4) Activity (Q) expresses the degree of action in the text, in contrast to descriptiveness. It is the ratio of verbs to the sum of verbs and adjectives that occur in the text. 5) Descriptiveness (D) expresses the degree of descriptiveness of the text. It is thus the inverse value of activity (see above). 6) Thematic concentration (TC) expresses the degree to which the text focuses on the main theme or themes, with the central theme being detected using so-called thematic words. In addition to the TC value, QuitaUp also displays individual thematic words and their weights (TW). Comparisons were made for case-sensitive (distinguishes between upper and lower case letters), case-insensitive (does not distinguish between upper and lower case letters) and lemma forms. Due to the limited scope of the study, only the results for lemma are given.

² *Týden v tichém domě* [A week in a quiet house]; *Pan Ryšánek a pan Schlegl* [Mr Ryšánek and Mr Schlegl]; *Přivedla žebráka na mizinu* [She Brought the Beggar to Ruin]; *O měkkém srdci paní Rusky* [About Mrs Rusky's Soft Heart]; *Večerní šplechty* [Evening Gossip]; *Doktor Kazisvět* [Doctor Kazisvět]; *Hastrman* [Hastrman]; *Jak si nakouřil pan Vorel pěnovku* [How Mr Vorel Smoked a Foam Pipe]; *U Tří lilií* [At the Three Lilies]; *Svatováclavská mše* [St. Wenceslas Mass]; *Jak to přišlo* [How It Happened]; *Psáno o letošních dušičkách* [Written About This Year's All Souls' Day]; *Figurky* [Figurines].

ries were originally published in magazines, and their length was apparently not determined by any uniform editorial restrictions. Nor does it show any relationship to the chronology of their creation. The longest texts frame the entire collection: *Týden v tichém domě* (1867) is the oldest text and was added to the book edition later,³ while *Figurky* (1877) is the youngest text. The difference between the two prose works may have contributed to the final order of the stories within the cycle; however, it cannot be interpreted as a factor determining the genesis or final form of the individual texts. Arne Novák (1919, p. 39) speaks of the compact nature of the collection, which is for him “a unified work, arranged cyclically”.

In order to objectify stylistic differences, we therefore use quantitative analysis of syntactic parameters. The first quantified stylistic indicator is linear sentence length (VD). This parameter represents a potential stylistic characteristic: statistical data on sentence length allow us to partially characterise the text, the author and the stylistic domain. In Czech, there is considerable variability in sentence length, which reflects the systemic possibilities of the language, the thematic focus of the text, the norms of the relevant stylistic sphere, and the individual author's disposition. The informative value of this indicator is limited, as the statistical characteristic of sentence length always captures only one aspect of style. The average VD value in the corpus is 5.552 tokens. The highest value is achieved by *Jak si nakouřil pan Vorel pěnovku* (6.338 tokens), the lowest by *Figurky* (4.631 tokens); the difference between the extreme values is 1.707 tokens. Texts with higher VD values also include *Pan Ryšánek a pan Schlegl* (6.179 tokens) and *Svatováclavská mše* (6.142 tokens). On the other hand, lower values, besides *Figurky*, are shown by *Večerní šplechty* (5,050 tokens), *Jak to přišlo* (5,098 tokens) and *Psáno o letošních dušičkách* (5,245 tokens).

³ Mocná deduces from Neruda's correspondence that Neruda first considered including this story in the PM cycle, even though he “still does not feel it is part of the Malá Strana cycle (this is evidenced by the fact that, from a genetic point of view, this prose is simply not a Malá Strana story)” (Mocná, 2012, p. 52).

A comparison of linear sentence length shows that SL is a relatively stable style indicator in the analysed corpus; the deviations from the average do not indicate a developmental trend. In the case of *Figurky*, the lower SL value is probably related to their genre stylisation (notes, diary and epistolary character). The short story *Jak si nakouřil pan Vorel pěnovku* is atypical in this respect: the deviation from the average here is caused by a higher frequency of linearly longer simple sentences, reflecting the author's specific intention. The differences found thus result more from the specifics of individual texts than from a systematic change in Neruda's sentence structure.

The summary data on VD can be supplemented by a detailed quantification of sentences and clauses in one of the short stories, specifically in the text with the highest VD value, *Jak si nakouřil pan Vorel pěnovku*⁴, and for verification of the results also by comparison with the short story *Hastrman*⁵ from the same period and of similar length.

⁴ The story shows a slight predominance of compound sentences over simple sentences (47 simple sentences; 40.87%, and 68 compound sentences; 59.13%). The average length of a simple sentence is 8.78 graphic words in the narrator's section (narrator's speech) and 3.2 graphic words in the characters' section (character speech), with a sentence equivalent of 2.5 graphic words in character speech. Within compound sentences, coordinate constructions predominate over subordinate ones (79.41% compound sentence; 20.59% complex sentence). The average linear length of a subordinate compound sentence is 20 graphic words in narrator's speech with an average number of 2.91 sentences; in character speech it is 13.5 graphic words and 2.5 sentences. Coordinate clauses have an average length of 17.14 graphic words and 2.78 sentences in narrator's speech; in character speech, they have an average length of 9.5 graphic words and 2.25 sentences. The text is characterised by a predominance of parataxis, most often in two- to four-sentence units. Subordinate clauses are less frequent, but on average longer than coordinate clauses. The average length of a simple sentence remains relatively high in relation to the length of a compound sentence: a simple sentence reaches 46.95% of the length of a compound sentence (simple sentence 8.78 words; sentence 18.7 words). In the character range, this ratio decreases to 21.77% (simple sentence 3.2 words; sentence 11.5 words).

⁵ The analysis yields different results: here, simple sentences slightly predominate over compound sentences (94 simple sentences; 56.29%, resp. 73 compound

A comparison of the two analysed short stories confirms that the frequency distribution of simple sentences and compound sentences is not invariant throughout the entire corpus, but rather depends on the text. Compound sentences predominate in the first short story (59.13%), while simple sentences predominate in the second (56.29%); but in both cases, the representation of both syntactic types is relatively balanced. A stable feature is the predominance of coordinate clauses over subordinate clauses (79.41% compound sentence; resp. 64.38% compound sentence), i.e. the systematic dominance of parataxis regardless of the overall proportion of clauses in the text. The average linear length of clauses and the average number of sentences in a clause are close in both stories, which indicates the relative stability of the syntactic structure of multi-sentence units. Although subordinate clauses are less frequent, they systematically have a higher average length than coordinate clauses. The most significant difference between the texts is in the ratio of the length of simple sentences to the length of compound sentences: in the first story, simple sentences account for 46.95% of the length of compound sentences, while in the second, they account for 26.82%. While the first text shows a low linear distance between the two syntactic units, the sec-

sentences; 43.71%). The average length of a simple sentence is 7.43 graphic words in narrator's speech; 3.84 graphic words in character speech, and 2.08 graphic words in sentence equivalents. Here, too, coordinate constructions predominate over subordinate ones (64.38% compound sentence; 35.62% complex sentence). The average length of a subordinate clause is 17.25 graphic words per 2.65 sentences in narrator's speech; 13.33 graphic words and 3 sentences in character speech. Coordinate clauses have an average of 14.98 graphic words and 2.71 sentences in narrator's speech; 16 graphic words and 2 sentences in character speech. Parataxis also predominates in this story, most often in two-sentence units; subordinate clauses are less frequent and on average longer than coordinate clauses. The average length of a simple sentence is significantly shorter than that of a compound sentence: a simple sentence is 26.82% of the length of a compound sentence (simple sentence 5.64 words; sentence 15.39 words), in the range of characters 20.75% (simple sentence 3.84 words; sentence 14.67 words). In the short story *Hastrman*, there is a more pronounced contrast between the linear length of simple sentences and compound sentences.

ond works with a more pronounced length contrast between simple sentences and compound sentences.

The variable parameter remains primarily the frequency ratio between simple sentences and compound sentences, while the dominance of parataxis, the average number of sentences in a compound sentence, and the differences between the narrator's and characters' speech remain stable. In the case of the short story *Hastrman*, the quantitative centre of gravity is in simple sentences, which are linearly shorter and more frequent. In the short story *Jak si nakouřil pan Vorel pěnovku*, on the other hand, compound sentences predominate, linearly longer and mostly coordinate, not broken down into a higher number of sentences (on average less than three sentences, similar to the short story *Hastrman*). The variability therefore appears to be a functional modulation within a stable syntactic model, rather than a manifestation of a different authorial style. Neruda's composition is capable of changing the distribution of basic sentence units while maintaining the same organisational principles. This finding is confirmed by frequency analyses of other short stories (e.g. in the short story *U Tří lilií*, the ratio of simple sentences to compound sentences is 49.11% to 50.89%). Neruda's syntactic structure has a dual character: flexibility in the distribution of basic sentence units and, at the same time, stability in their organisational principles.

The statistical data obtained on VD can be further compared with values from works of similar time period and genre, especially prose by representatives of the same literary generation: Vítězslav Hálek (1835–1874), Jakub Arbes (1840–1914) and Karolína Světlá (1830–1899). Such a comparison allows us to place Neruda's sentence structure in the broader context of the syntactic norms of artistic prose of a single poetic generation.⁶ The average VD value for Neruda

(5.552 tokens) is higher than for Hálek's short stories (*Muzikantská Liduška* 4.162; *Na statku a v chaloupce* 4.523 tokens). Karolína Světlá's values are at a similar level (*Černý Petříček* 5.423 tokens), while those for the novel *Kříž u potoka* are slightly lower (4.945 tokens). The average VD for Jakub Arbes, on the other hand, is higher (*Svatý Xaverius* 6.359; *Ukřižovaná* 6.312; *Newtonův mozek* 6.466 tokens) and approaches the extreme value of Neruda's short story *Jak si nakouřil pan Vorel pěnovku* (6.338 tokens).

Compared to Hálek, whose prose tends towards linearly shorter sentences and more pronounced segmentation (characterised by linearly shorter simple sentences and less developed compound sentences, which in places resemble the dynamics of spoken diction), Neruda's sentences appear more compact and are more often constructed on a paratactic chaining of relatively autonomous utterances. In relation to Arbes, Neruda's VD ranges in the middle values: Arbes has a higher frequency of simple sentences, works with linearly longer sentences, often expanded by the accumulation of developing sentence elements. Neruda, on the other hand, achieves a more balanced ratio between the linear range of the sentence and syntactic complexity, so that his sentence structure does not appear significantly condensed or excessively branched. The closest to Neruda's values is the prose of Karolína Světlá, in which, however, hypotactic modelling of the utterance plays a more significant role. The comparison thus shows that Neruda's sentences represent a middle type within the generational production: they do not exhibit extreme values in terms of linear sentence length or significant syntactic complexity expressed by the number of sentences or the number of dependencies in compound sentences. Rather, a characteristic feature is structural balance and relative stability across the Malá Strana stories. The linear length

⁶ The comparison included Hálek's short stories *Muzikantská Liduška* (1861, 9,310 tokens) and *Na statku a v chaloupce* (Nsach, 1871, 30,990 tokens), the prose work *Černý Petříček* (ČP, 1871, 28,705 tokens) and the novel *Kříž u potoka* (KuP, 1868, 83,957 tokens) by Karolína Světlá, and Arbes's novellas *Svatý Xaverius* (SX,

1873, 23,095 tokens), *Ukřižovaná* (U, 1876, 37,094 tokens) and *Newtonův mozek* (NM, 1877, 21,071 tokens). The compared texts are more extensive than the PM collection, but they correspond to the 1860s and 1870s, i.e. the period of their creation.

of sentences in Neruda's work is thus not a significant individualising feature, either in the diachronic development of the corpus under study or in relation to his contemporaries.⁷

The linear length of a sentence is often linked to the distribution of speech between the narrator and the characters. The highest proportion of character speech is found in *Večerní šplechty* (68.16%), *Týden v tichém domě* (46.59%) and *Figurky* (34.35%), while the lowest is in *Svatováclavská mše* (0.57%), *U Tří lilií* (4.93%), *Mr Ryšánek and Mr Schlegl* (6.15%) and *Jak si nakouřil pan Vorel pěnovku* (7.79%). The highest difference between the VD values in the narrator's range (narrator's speech) and the characters' range (character speech) is found in *Svatováclavská mše* (6.142 / 3.5 tokens), *Doktor Kazisvět* (5.731 /

⁷ The comparative perspective of the presented analysis can be extended by the results of the study *Sentences and Clauses in J. Arbes, A. Jirásek, and J. Zeyer* (Schacherl, 2022, pp. 337–352), which quantifies selected parameters of sentence structure in three authors of the second half of the 19th century. The study includes the texts *Il Divino Boemo* (novella, 11,273 words) by Jakub Arbes, *Balady z Rokoka* (ballad, 4,548 words) by Alois Jirásek, and the text *Inultus* from the cycle *Tři legendy o krucifixu* (legend, 5,963 words) by Julius Zeyer. The inclusion of this data allows the results obtained for Neruda to be placed in the broader context of contemporary prose, transcending generational affiliations, and to be compared with different authorial and genre models without the need to expand the analytical corpus itself. Although the differences between individual authors appear less pronounced in quantitative terms, they are significant in terms of singular style and allow us to capture the distinctive features of their syntactic organisation, with a possible overlap into more general characteristics of authorial style. A comparison with the singular style of *Povídky malostranské* shows that Jan Neruda's sentences and clauses occupy middle values within the quantitative linguistic characteristics studied, both in terms of frequency and linear length. With his predominance of parataxis, Neruda is similar to Alois Jirásek and Julius Zeyer, while the higher frequency of simple sentences and their linear length show a similarity to the syntactic model of Jakub Arbes, although Neruda is functionally flexible in this respect. From a comparative perspective, Neruda's syntactic profile thus represents a balanced type in quantitative characteristics, oscillating between a paratactic tendency and a more developed simple sentence, thus moving between the more extreme poles represented by Arbes's hypotactic expansion and Zeyer's and Jirásek's paratactic dominance.

4.225 tokens) and *Pan Ryšánek a pan Schlegl* (6.179 / 4.844 tokens). Conversely, the lowest differences are found in *Figurky* (4.631 / 4.633 tokens), *Večerní šplechty* (5.050 / 4.940 tokens), *Přivedla žebráka na mizinu* (5,512 / 5,361 tokens) and *Týden v tichém domě* (5,284 / 5,120 tokens); the linear length of the sentence is almost equal to the values in the narrator's range in the characters' speech. *Figurky* are the only text in the collection in which the VD in character speech is slightly higher than in narrator's speech. The differences found are mostly small and do not indicate a systematic trend across the collection. Therefore, it is not possible to prove a direct dependence between the linear sentence length and the vertical structure of the text; the degree of representation of the characters' speech does not in itself determine the differences in verb distance values between narrator's speech and character speech. VD thus does not appear to be a parameter determined by the distribution of communication bands.

The values of activity (Q) and descriptiveness (D) usually correlate with the average sentence length; however, this correlation does not apply unconditionally in *Povídky malostranské*. The highest activity value is shown by *Figurky* (Q = 0.696), which also has the lowest VD. This is followed by *Večerní šplechty* (Q = 0.637), *Týden v tichém domě* (Q = 0.624) and *Přivedla žebráka na mizinu* (Q = 0.622). The highest descriptivity values are achieved by *Svatováclavská mše* (D = 0.461), *Pan Ryšánek a pan Schlegl* (D = 0.442) and *U Tří lilií* (D = 0.435). The highest activity values are shown by two mixed-genre texts together with the short story *Večerní šplechty*, where the dynamics are reinforced by inserted texts and a higher proportion of direct speech. It is noteworthy that the probable authorial intention diverges from the linguistic structure of the text: Neruda apparently sought a different stylistic effect than that suggested by the quantitative results based on the stylometric indicators monitored.⁸ The plot of the short story *Přivedla*

⁸ He told me that he was writing something special. That it would be a story completely without any plot, without any intrigue, and that only the characters or figures, as he put it at the time, would be and should be the main thing (Quis, 1908, p. 135).

žebráka na mizinu may be surprising in this context; however, here too, Neruda supports the dynamics of the text with the characteristic representation of parts of speech and a longer average sentence length in the characters' speeches. The short stories *Jak si nakouřil pan Vorel pěnovku* and *Hastrman*, analysed in more detail, fall within the range of average values across the collection (Vor: Q = 0.586; D = 0.414; Ha: Q = 0.580; D = 0.420). The data show that the activity and descriptiveness of the text are not realised exclusively through the differentiation of average sentence length across the collection. Neruda's syntactic structure is also characterised by a relatively low contrast between the linear length of simple sentences and compound sentences, so that the differences in Q and D values cannot be reduced to a mere variation in the length parameter. It is more appropriate to interpret these phenomena as the result of a more complex configuration of syntactic and lexical factors. Activity and descriptiveness are therefore not isolated quantitative parameters, but contribute to the complex organisation of the semantic structure of the text.

Aleš Haman explains the relatively lower values of activity (Q) and higher values of descriptivity (D) in Neruda's work with the concept of so-called internal action; F. X. Šalda speaks in a similar sense of internal discursivity. According to Haman, "on the surface, not much happens in Neruda's prose – in contrast to the external dramatic action of Světlá's stories; but beneath this seemingly motionless surface, hidden dramas unfold and a hidden life pulsates intensely. In a single scene, Neruda manages to concentrate the diversity and multiplicity of his fictional world" (Haman, 2014, pp. 191–192). Stylometric data indirectly support this characterisation: lower Q and higher D do not mean an absence of dynamics, but rather its shift from the level of external action to the level of internal semantic differentiation.

A comparison with authors of the May generation confirms this picture. In the short stories of Vítězslav Hálek, the values of Q reach 0.785 (ML) and 0.787 (Nsach), while D is 0.215 (ML) and 0.213 (Nsach). Compared to Neruda's texts, they are characterised by significantly higher activity and lower descriptiveness, which also corre-

sponds to a shorter average sentence length. For Karolína Světlá, the values are VD 4.945 (KuP) and 5.423 (ČP), Q 0.712 (KuP) and 0.651 (ČP), and D 0.288 (KuP) and 0.359 (ČP). The sentence length is slightly shorter than in Neruda's work, but the more significant difference is in the parameters of activity and descriptivity: the texts are more active and less descriptive. Jakub Arbes's short novel (SX 23,095 tokens, U 37,094 tokens, NM 21,071 tokens), comparable in length to the longest Little Quarter story, has values of VD 6.359 (SX), 6.312 (U), 6.466 (NM), Q 0.565 (SX), 0.547 (U), 0.543 (NM) and D 0.435 (SX), 0.453 (U), 0.457 (NM). Arbes's sentence is comparable in length to the highest values in Neruda's work, or slightly exceeds them; however, in terms of the Q and D parameters, his texts are less active and at the same time slightly more descriptive.

The comparison thus shows that Neruda's position within the generational context is specific: his syntactic structure combines relatively longer sentences with little contrast between the linear length of simple sentences and compound sentences with a balanced, central position on the axis of activity and descriptiveness. A lower Q value here therefore does not imply staticity or a lack of action, but rather signals that the dynamics of the text are not primarily based on an action sequence of events. The plot impulse shifts from the level of external action to the realm of semantic relationships, situational condensation, and subtle shifts in characterisation. Stylometric indicators thus quantitatively capture what literary-historical interpretation refers to as "internal plot": not a dramatic escalation of external action, but an intensification of the semantic structure within a relatively stable syntactic framework.

While the analysis so far has followed the macrosyntactic parameters of sentence structure, the next part focuses on the means of internal structuring of the utterance, which literary history associates with the characteristic "looseness" of Neruda's syntax. The second parameter observed, which literary-historical reflection repeatedly associates with Neruda's prose, is parentheticality. Scholarly interpretation links the decomposability or fragmentation of Neruda's syntax to

the fact that, in addition to basic graphic symbols (full stop, question mark, exclamation mark), the author also makes significant use of the dash and semicolon. The dash (–) does not only serve as a separator, but also signals an interruption or retardation of the flow of speech, delimits parenthetical expressions and opens up space for the insertion of secondary segments into a linearly constructed sentence structure. The graphic segmentation of the text thus becomes a visible correlate of the parenthetical principle, i.e. the internal loosening of tight syntactic bonds. The semicolon (;) is a punctuation mark that stands between a comma and a full stop: it indicates a greater degree of separation than a comma, without making the statement fully independent like a full stop. It is used mainly in longer sentences between coordinately connected main clauses or when separating closely related statements where it is necessary to preserve their semantic coherence and at the same time clearly differentiate them. Its functional load thus corresponds to a syntactic model combining coordination with internal segmentation and can be interpreted as another level of realisation of the parenthetical tendency.

The occurrence of sentences and clauses divided by a dash or semicolon, which has been repeatedly pointed out by literary history and criticism, occurs 12 times in the short story *Jak si nakouřil pan Vorel pěnovku* (13.8% of all sentences and clauses). Of this number, 9 occurrences are dashes (10.35%) and 3 are semicolons (3.45%). Four cases are simple sentences (1.88% of all simple sentences), eight are compound sentences (5.44% of all compound sentences), specifically one subordinate clause (7.1% of all subordinate clauses) and seven coordinate clauses (3.5% of all coordinate clauses). Here, the dash separates a separately highlighted sentence element in the preposition (*Druhý, třetí den – no však to ještě půjde!*) and postposition (*Staříčkou vdovu odvezli před čtvrt rokem do Košíř a teď – nač jen ten krám!*); it can introduce a separate statement within a sentence (*Tak krásně nakouřenou pěnovku jsem ještě neviděl – podívejte se!*) or incorporate an interjection into a sentence (*Sklenky kolovaly – paní hospodská měla v celé zásobě své jen dvě sklenky vinní –, ale žádná nezavadila o pana*

Vorla). The semicolon is used to separate closely related statements less emphatically, both in relation to a compound sentence and a subordinate clause (*Bylo tomu právě také tak asi čtvrt roku, co se pan Vorel byl přistěhoval do Ostruhové ulice; přicházel odněkud z venkova*), as well as between two clauses (*Nevěděli o něm ničeho, než že je synem mlynáře; snad by jim byl sám pověděl víc, a to ochotně, ale neptali se ho*). Constructions with dashes and semicolons are used in both the narrator's and the characters' speech. Their frequency is relatively low, which confirms that they are a rather rare device, albeit stylistically striking.

Verification of this data in the short story *Hastrman* yielded structurally comparable results. A total of 18 instances of the observed horizontal division were recorded in the text (10.87% of all simple sentences and compound sentences), of which 16 were realised with a dash (9.58%) and two with a semicolon (1.2%). The use of the semicolon corresponds to the principle already documented in the previously analysed short story, i.e. a less emphatic separation of closely related sentences while maintaining their semantic and continuity. Neruda also uses the semicolon repeatedly within the same compound sentence (*Kdo šel právě Ostruhovou ulicí, zůstal stát; kdo stál právě tak podomácku venku, zavolal honem něco dovnitř domů; z krámů vybíhali lidé*), thereby strengthening its internal segmentation while maintaining the coordinate relationship between the individual parts and linearly extending the statement. Here, too, it is a means that is limited in frequency and stylistically marked. In the case of the dash, as in the previous text, two basic functions prevail: the delimitation of an insertion (*My děti jsme sobě vypravovaly, že když u Šajvlů – tak se nazývali ti páně Rybářovi příbuzní – se myje podlaha, posypávají ji místo pískem samým natlučeným cukrem*) and the signalling of a pause in speech. At the same time, the dash allows the speech to be segmented into several rhythmically separated sections, both in the narrator's section (*Neviděl nic určitého, vše se třpytilo jako rozpuštěný dým, vše se vlnilo – až k samému oknu – až do jeho oka – moře – moře!*), and in the characters' speech (*Pak zašeptal, dívaje se upřeně k*

Petřinu: „*Nemá to ceny – obálky –*“). A comparison with the previous text confirms, in this case too, the low but stable presence of these devices, whose function is primarily to extend the linear length of the statement, increasing its internal segmentation and rhythmic structure, but not changing the basic syntactic model of its construction.

Neruda uses both of the observed means of horizontal division in accordance with the features of his diction, as defined above. The semicolon separates mainly coordinate, semantically closely related sentences; graphically, it thus marks a higher, internally structured whole that transcends the boundaries of individual sentences without completely separating the statements. It thus creates a segment with a higher degree of internal cohesion than would be implied by separation with a full stop, and at the same time with a more pronounced break than that provided by a comma. The dash signals an interruption or delay in the flow of speech, marks a pause between characters' lines, and delimits parenthetical remarks within a sentence. Its function is primarily to highlight internal segmentation and the rhythmic structure of speech. The frequency of both devices remains relatively low, with semicolons being rare. Quantitative data thus suggest that although their functional characteristics correspond to interpretations by literary critics and historians, their position in the overall organisation of the text is not dominant. However, their stylistic conspicuousness may lead to an interpretative emphasis on their significance.

The more pronounced horizontal division of the text in *Povídky malostranské* does not appear in the analysed corpus as a factor of fragmentation or syntactic incoherence, but rather as a means of continuous organisation and internal differentiation of the statement. This effect is particularly noticeable at the level of paragraph composition. The situation can be illustrated by the example of the story *Jak si nakouřil pan Vorel pěnovku* (1,501 tokens), which is divided into 24 paragraphs; the average length of a paragraph is 4.79 sentences. The short story *Hastrman* (1,722 tokens) contains 43 paragraphs with an average length of 3.89 sentences. The higher number of shorter paragraphs in the second text thus corresponds to a finer segmentation of the nar-

rative for a comparable text length. The validity of these findings can be reinforced by extending the comparison to include the shortest short story, *U Tří lilí* (649 tokens), and the longest text in the collection, *Týden v tichém domě* (21,740 tokens). The former is structured into 56 paragraphs, the latter into 654 paragraphs, with an average length of 3.5 and 2.45 sentences, respectively. The length of the paragraphs correlates with the average linear length of the sentences in the individual stories; the most pronounced paragraph segmentation is evident in the story *Týden v tichém domě*. Another characteristic feature of the collection is the tendency to begin a paragraph with a simple sentence, which often serves as a thematic sentence: it names the main theme of the paragraph or defines it in more detail. This approach enhances the rhythmic clarity of the text — the thematic sentence at the beginning of the paragraph creates a clear entry point for the thematic section, which may be followed by a more syntactically complex development. In terms of its structure, this principle is reminiscent of the concept of *paragraph writing*, i.e. the strict compositional organisation of paragraphs in the Anglo-Saxon essay tradition. The introductory sentence is developed by supporting sentences, which further specify the main idea and elaborate on it with arguments; the concluding sentence of the paragraph then often serves as a summary or transitional formulation. The established structure of the paragraph increases the comprehensibility of the text: the main ideas are presented in a timely, clear and concise manner. This technique is probably related to Neruda's long career as a journalist. The horizontal division does not disrupt the continuity of the text, but rather gives it an organised rhythm and greater compositional variability while maintaining structural cohesion. In the short story *Jak si nakouřil pan Vorel pěnovku*, 12 of the 24 paragraphs (50%) begin with a simple sentence, and in the short story *Hastrman*, 26 of the 43 paragraphs (60.47%) begin with a simple sentence.

In addition to syntactic parameters, stylistic stability can also be observed at the level of lexical diversity. Entropy values, i.e. the index of lexical richness, demonstrate the lexical stability of Neruda's prose

in *Povídky malostranské*. This is confirmed by similar values in the framework stories *Týden v tichém domě* (1867; H = 10.965–10.741–9.651) and *Figurky* (1877; H = 10.691–10.397–9.389), the two most extensive texts in the collection, which show very similar results. Comparable values are also achieved by stories of similar length from 1875 and 1876: *Mr Ryšánek and Mr Schlegl* (H = 9.620–9.470–8.723; 3,071 tokens) and *Psáno o letošních dušičkách* (H = 9.674–9.479–8.701; 3,243 tokens). Analogous stability is also evident in the shorter texts *O měkkém srdci paní Rusky* (H = 9.114–8.969–8.303; 1,639 tokens) and *Jak si nakouřil pan Vorel pěnovku* (H = 8.979–8.849–8.296; 1,501 tokens). The lexical richness index therefore shows no significant fluctuations either in diachronic comparison or when taking into account the genre, thematic or stylistic specifics of individual texts, which testifies to the stability of lexical diversity across the entire corpus.

A comparison of the linguistic characteristics of short stories identified by literary history as having thematic or compositional similarities did not reveal any systematic linguistic conditioning. A comparison of texts with predominantly journalistic features (*Pan Ryšánek a pan Schlegl, Přivedla žebráka na mizinu, O měkkém srdci paní Rusky*) and short stories with a predominantly fictional character (*Doktor Kazisvět, Hastrman, Jak si nakouřil pan Vorel pěnovku*) does not show significant differences in the quantitative indicators of sentence structure. Even in texts with autobiographical features (*Večerní šplechty, U Tří lilií, Svatováclavská mše*), no symptomatic similarities or deviations were found within the analysed parameters. In contrast, stylometric analyses of thematic concentration link certain groups of texts on the basis of frequency similarities of so-called thematic words. The short stories *Týden v tichém domě, Večerní šplechty* and *Figurky* are linked by a predominance of frequent adverbs (*už, tak, pak*), non-meaningful verbs (*být, mít*) and forms of address (*pan, pane*). The second group consists of the short stories *Doktor Kazisvět* and *Přivedla žebráka na mizinu*, where the thematic axis is created by a lexicon associated with the central character (*doktor, Kazisvět, pan,*

Vojtíšek). The stylistic indicators of Neruda's syntactic structure do not show any significant diachronic changes. In the analysed corpus, the first and last texts (*Týden v tichém domě, Figurky*) differ significantly in length; However, in terms of the language parameters studied, only *Figurky* stands out more significantly, with a lower average linear sentence length, a higher linear length of utterances in the character range compared to the narrator range, and an increased representation of direct speech. The short story *Jak si nakouřil pan Vorel pěnovku* has the highest average linear sentence length and, at the same time, a less pronounced horizontal division of the text into a smaller number of linearly longer paragraphs, which is reflected in a higher degree of narrative continuity. The linguistic structure of the collection is thus characterised by a high degree of stability and relative autonomy from thematic, genre and autobiographical factors. The stylistic identity of the collection is thus shaped primarily by the internal organisation of linguistic means, rather than by the thematic or compositional differentiation of individual texts.

Discussion

The analytical findings can be interpreted in the broader context of the literary-historical evaluation of Neruda's prose. The stylistic dominants of Neruda's composition – paratacticity and parentheticality – are understood in the present analysis as empirically verifiable syntactical phenomena. These phenomena can be operationalised through selected stylometric identifiers, in particular verb distance, parameters of text activity and descriptiveness, quantification of parenthetical expressions, the degree of interruption of immediate syntactic links, and the extent of internal segmentation of sentence units. The results of stylometric analyses show that Jan Neruda's prose is characterised by a slight predominance of compound sentences over simple sentences in terms of sentence structure. Within compound sentences, coordinate constructions predominate over subordinate constructions, which are less common, but their average linear length is greater than that of coordinate sentences; hypotactic constructions

are therefore not structurally marginal, but functionally reserved for more complex units of meaning. Paratactic compound sentences most often consist of two to four sentences, which confirms the tendency towards relatively compact syntactic segmentation without significant expansion into multiple hierarchical structures. These features apply regardless of the length of the individual stories, their chronological order, or the degree of representation of the characters. The prominent position of clauses and the predominance of parataxis are therefore not text-dependent deviations, but rather a systemic feature of the author's syntactic organisation. The results of the analysis do not suggest that Neruda breaks down his diction into a larger number of small sentence units.

The average linear length of a simple sentence is approximately half the length of a compound sentence; however, the difference between the two units is relatively small. Simple sentences exhibit relatively high lexical saturation and are functionally similar to shorter compound sentences. The contrast between simple sentences and compound sentences is therefore mainly quantitative, not fundamentally structural. This is also reflected in the low range of values for sentence linear length within the corpus, which does not indicate any developmental trend or systematic differentiation between texts. An analysis of the narrator's and characters' registers did not reveal a direct correlation between the proportion of direct speech and the linear length of sentences. The differences between the two registers are minor in most cases; only in some texts is there a more significant shortening of syntactic units in the characters' register, which is reflected in an increase in text activity values. Even a high proportion of dialogue passages does not automatically lead to a reduction in the linear length of the sentence. The vertical division of the text is therefore not a determining factor in the basic parameters of sentence structure; the syntactic model remains structurally consistent in both bands, so that the differences between the bands are not of a systemic stylistic nature.

The use of dashes and semicolons represents a specific level of syntactic organisation. Sentences and clauses divided by these punctuation marks constitute a relatively low proportion of all syntactic units (approximately 10–14% in the analysed texts); it is therefore a means that is limited in frequency but stylistically striking and symptomatic. The dash is significantly more frequent than the semicolon in *Povídky malostranské* and serves to delimit parenthetical expressions, signal a pause in speech, and segment utterances. The semicolon separates closely related sentences and signals a medium degree of syntactic and semantic autonomy. Both devices act as a factor prolonging the linear length of the sentence and at the same time reinforce the dominance of parataxis; they also represent an important tool for internal articulation and rhythmisation of the utterance within a predominantly paratactic model. The horizontal division of the text, realised through paragraphs, shows a clear tendency in Jan Neruda's work towards relatively short, compositionally compact units. These are often initiated by a simple sentence or a linearly shorter compound sentence, which serves as the thematic core and explicitly formulates the basic statement or initial idea of the following section. As a rule, the paragraph thus has a stable internal structure: the introductory thematic sentence introduces the main theme, the middle part consists of sentences that develop and support the argument or narrative, and the conclusion serves a summarising or transitional function, linking the given section with the following text. This compositional regularity reinforces the rhythmic organisation of the statement and contributes to its segmented clarity. The stability and functional coherence of the paragraph is in many ways reminiscent of the model of the thematically oriented paragraph known from the Anglo-Saxon tradition of *academic writing*, in particular its emphasis on the early explicitation of the main idea and its systematic development. In the context of Neruda's prose work, this phenomenon can also be interpreted as a structural reflection of his long-term journalistic practice, which favours semantic economy, thematic transparency and reader-friendly text organisation. The paragraph technique is thus not only a formal

compositional device, but also an important factor in textual cohesion and comprehensibility: key ideas are exposed early on, formulated explicitly and then systematically developed, which supports interpretative stability and the reader's comfort.

The length of the individual stories – from very short prose to more extensive compositions – does not imply a fundamental structural change in sentence construction. The variability in text length therefore does not imply a structural change in sentence construction. The parameters of activity (Q) and descriptivity (D) range within a relatively balanced band of values; their higher values are linked to specific texts, not to systemic differentiation in the corpus as a whole. The relationship between Q/D and linear sentence length is not direct: a higher degree of descriptivity does not automatically mean longer syntactic constructions. The dynamics of the text are thus realised through semantic differentiation and internal segmentation rather than syntactic expansion, which supports the conclusion that the basic syntactic model is stable.

A general comparison with the prose of authors of a similar generation (e.g. Vítězslav Hálek, Jakub Arbes, Karolína Světlá) suggests that Neruda's sentences and clauses occupy a middle position in terms of the quantitative characteristics studied. They show neither significant hypotactic expansion nor a tendency towards reduced, short sentences. The specific stylistic dominants of Neruda's prose – a tendency towards paratactic and parenthetical structures – are thus not manifested in extreme quantitative values, but rather in structural organisation within a relatively stable range.

In summary, it can be said that the syntactic structure of *Povídky malostranské* is based on the stable dominance of parataxis, a relatively balanced ratio of the linear length of simple sentences and compound sentences, a frequency-limited but stylistically marked horizontal division of the text, and structural continuity across texts of varying length and compositional organisation. Quantitative data thus define Neruda's sentence organisation as systemically consistent and modifiable within individual texts, but without significant fluctua-

tions in basic syntactic parameters. These findings also relativise interpretative stereotypes that characterise Neruda's syntax as significantly fragmentary or loose. Empirical data rather support the conclusion about the stability and internal discipline of his syntactic model, which is evident across the entire set of analysed texts.

Conclusion

The presented analysis allows us to summarise the main findings regarding the syntactic organisation of the collection. Stylo-metric analysis of sentences and clauses in Jan Neruda's *Povídky malostranské* (*Tales of the Lesser Quarter*) demonstrates the existence of a stable syntactic model that is consistently applied throughout the entire collection. The dominant organisational principle is parataxis, while hypotaxis appears as a means of semantic differentiation rather than as the basic structural framework of the text. The difference between a simple sentence and a compound sentence is mainly a difference in scope rather than a qualitatively different syntactic strategy.

The basic parameters of sentence structure remain largely invariant regardless of the length of the text, the degree of dialogue or the thematic variability of individual prose works. The indicators monitored – linear sentence length, degree of verbosity and nominality, verbal distance and frequency of means of higher punctuation articulation – vary within a relatively narrow range of values and do not indicate any developmental trend or significant stylistic differentiation between texts. Punctuation devices, especially the dash, do not function as a tool of syntactic expansion, but rather as a means of internal segmentation and rhythmisation of paratactically organised utterances.

The analysed corpus does not show signs of diachronic experimentation in the area of sentence organisation, but represents the realisation of an already established authorial model. The variability between individual stories is mainly evident at the thematic and compositional level, not in the basic model of syntactic construction. Stylo-metric data here serve only as an illustrative supplement to interpretative judgements, rather than as their empirical support and corrective.

Neruda's sentence organisation thus appears to be a structurally consistent, paratactically oriented and balanced system that is resistant to intra-corpus variability.

Literature

- ARBES, J. (1952). *O Janu Nerudovi* [On Jan Neruda]. Prague: Melantrich.
- BARTŮŇKOVÁ, J. (1985). Funkce stylizované verbální komunikace v povídkách Jana Nerudy (K 150. výročí narození) [The Function of Stylised Verbal Communication in Jan Neruda's Short Stories] (On the 150th Anniversary of His Birth). *Naše řeč*, 68, 124–135.
- CVRČEK, V., ČECH, R., & KUBÁT, M. (2020). QuitaUp – a tool for quantitative stylometric analysis. Czech National Corpus and University of Ostrava. Available from: <https://korpus.cz/quitaup>.
- HAMAN, A. (1958). *Neruda prozaik* [Neruda as a Prose Writer]. Prague: Odeon.
- HAMAN, A. (2014). *Tři stálice moderní české prózy: Neruda, Čapek, Kundera* [Three Pillars of Modern Czech Prose: Neruda, Čapek, Kundera]. Prague: Karolinum.
- HAUSENBLAS, K. (1977). Řeč o řeči v Nerudově povídce *U tří lilíí* [Language about Language in Neruda's Short Story *At the Three Lilies*]. *Slovo a slovesnost*, 38(4), 336–339.
- HAUSENBLAS, K. (1984). Nad jazykem a stylem prózy Jana Nerudy [On the Language and Style of Jan Neruda's Prose]. *Literární měsíčník*, 1(1), 68–73.
- KRÁLÍK, O. (1995). Nerudovy Povídky malostranské [Neruda's *Tales of the Lesser Quarter*]. In *Liberated Words* (pp. 256–307), edited by Jiří Opelík. Prague: Torst.
- KŘIVÁNEK, V. (1983). *Jan Neruda*. Prague: Horizont.
- MACUROVÁ, Alena. (1991). Reflexe smíchu a smích jako prostředek reflexe člověka a světa (Poznámka k Nerudovým Povídkám malostranským) [Reflection on Laughter and Laughter as a Means of Reflecting Man and the World (A Note on Neruda's *Tales of the Lesser Quarter*)]. In *Proudý český umělecký tvorby 19. století. Smích v umění* (pp. 101–106). Prague.
- MOCNÁ, D. (2012). *Záludný svět Povídek malostranských* [The Insidious World of *Tales of the Lesser Quarter*]. Prague: Academia.
- NOVÁK, A. (1910). *Jan Neruda*. Prague: SVU Mánes.
- NOVÁK, A. (1919). *Studie o Janu Nerudovi* [Studies on Jan Neruda]. Prague: F. Topič.
- PLECHÁČ, P., & FLAIŠMAN, J. (2017). Problém Barák–Neruda z pohledu současné stylometrie [The Barák–Neruda Problem from the Perspective of Contemporary Stylometry]. *Česká literatura*, 65(5), 743–769.

- POLÁK, K. (1947). *Nerudovy Povídky malostranské* [Neruda's *Tales of the Lesser Town*]. Prague: J. R. Vilímek.
- QUIS, L. (1908). *O Janu Nerudovi*. Prague.
- SCHACHERL, M. (2022). Sentence Structure in the Works of J. Arbes, A. Jirásek, and J. Zeyer. *Bohemistika*, 22(3), 337–352.
- STICH, A. (1974). Jan Neruda jako teoretik a praktik jazykové kultury [Jan Neruda as a Theorist and Practitioner of Linguistic Culture]. *Naše řeč*, 57(5), 225–242.
- ŠALDA, F. X. (1934). Neruda poněkud nekonvenční [Neruda Somewhat Unconventional]. *Šaldův zápisník*, 6(10–11), 320–345.