

Červenka, Miroslav; Sgallová, Květa

## Na cestě k pravděpodobnostnímu modelu českého verše

In: *Teorie verše. II, Sborník druhé brněnské versologické konference, 18.-20. října 1966*. Palas, Karel (editor); Levý, Jiří (editor). Vyd. 1. Brno: Universita J.E. Purkyně, 1968, pp. 67-72

Stable URL (handle): <https://hdl.handle.net/11222.digilib/120152>

Access Date: 01. 03. 2024

Version: 20220831

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

# NA CESTĚ K PRAVDĚPODOBNOSTNÍMU MODELU ČESKÉHO VERŠE

MIROSLAV ČERVENKA — KVĚTA SGALLOVÁ (Praha)

Tato úvaha<sup>1</sup> vznikla v souvislosti s rozsáhlou materiálovou prací, kterou podniká tým pro statistický výzkum verše v Ústavu pro českou literaturu.<sup>2</sup> Je jasné, že v okamžiku hodnocení výsledků statistického popisu konkrétních veršových útvarů vyvstane otázka příznakových a nepříznakových vlastností veršované řeči i jejích individuálních realizací, a spolu s tím i otázka pozadí, na němž se tyto vlastnosti rýsují.

1. Základním materiálem statistického výzkumu veršové stavby bude zjištění těchto distribucí:

- a) distribuce příslušných jazykových jednotek v mluvě neveršované;
- b) pravděpodobné distribuce týchž jednotek ve verši — vypočtené jako apriorní pravděpodobnostní model na základě distribuce (a) a metrických norem verše;
- c) skutečný úzus, tj. distribuce rytmicky relevantních jazykových jednotek v určitém veršovém typu, období, básnické škole, individuálním díle a jeho poměr k modelu (b).

Odpověď na otázku (a) dává lingvistická statistika; my chceme na tomto místě podat program vytvoření teoretického modelu (b). Obecné předpoklady k tomu byly už dány v pracích Kolmogorovovy školy a u nás J. Levého.

2. Vycházíme z této metrické normy českého sylabotónického verše:

1.1. konstanta — nepřítomnost akcentu víceslabičného slova na rytmicky nedůležitých slabikách verše (tj. sudých v trocheji, lichých v jambu).

1.2. tendence — nepřítomnost akcentu jednoslabičného slova na rytmicky nedůležitých slabikách verše.

1.3. konstanta — nepřítomnost více než i-členné řady akcentovaných monosylab za sebou v kterémkoli místě verše.

2. konstanta — přítomnost akcentu na počáteční rytmicky důležitá slabice verše (tj. na 1. slabice verše v trocheji, na 2. nebo 1. slabice verše v jambu).

3. tendence — přítomnost akcentu na rytmicky důležitých slabikách verše (lichých v trocheji, sudých v jambu; netýká se ovšem počátečních slabik, kde běží o konstantu).

---

<sup>1</sup> Její plné znění vyšlo anglicky ve sborníku *Prague Studies in Mathematical Linguistics II.*, Praha 1967.

<sup>2</sup> Zprávu o něm viz v *České literatuře* 13, 1965, str. 541n.



5. Jak ukázala předběžná zkouška, celkový počet možných sestav taktů, realizujících jen jediný typ, už u šestislabičných veršů dosahuje několika desítek. Bez využití automatického počítače by tento úkol pro delší verše byl těžko zvládnutelný.

Bylo už naznačeno, že v dalším stadiu dojde ke srovnání teoretického modelu s empiricky zjištěnou distribucí rytmických typů v individuálním díle, básnické škole, v celém vývoji novodobé české poezie a tak ke stanovení povahy úzu, jeho blízkosti či vzdálenosti od jazykových předpokladů.

Zvláštní pozornost při empirickém výzkumu bude potřeba věnovat těm typům, jež lze pokládat za příznakové nejen pro malou jazykovou pravděpodobnost jejich výskytu, nýbrž i pro inherentní vlastnosti jejich zvukově rytmické kontury.

## 6. Prvky modelu a výpočet pravděpodobnostních údajů

### 1. *Element*

1.1. Jsou dány tři nedefinované jednotky, které označujeme **a**, **b**, **c** a nazýváme elementy (**e**).

Interpretujeme tyto elementy takto:

- a** — přízvučná slabika víceslabičného slova, event. přízvučná předložka,
- b** — přízvučné jednoslabičné slovo,
- c** — jakákoli nepřízvučná slabika.

1.2. **e** může existovat jen jako součást segmentu (viz 2.).

1.3. Každý **e** zaujímá jednu pozici v řadě (viz 3.1.).

### 2. *Segment*

2.1. Segment (**s**) definujeme jako jednočlenný až sedmičlenný řetěz elementů, který vyhovuje těmto podmínkám:

- a) V **s** musí být jeden a jen jeden **e** odlišný od **c**;
- b) po **a** musí následovat alespoň jedno **c**;
- c) před **a** nebo **b** smí předcházet nanejvýše jedno **e**.

Segmenty interpretujeme jako typy přízvukových taktů.

2.2. Tím je dán repertoár všech **s**.

2.3. Je dána relativní četnost každého **s** v množině všech **s** (a to na základě empirických zjištění).

### 3. *Řada, rozměr, typ*

3.1. Řada (**R**) je definována jako řetěz segmentů, který neobsahuje více než 3 **b** těsně za sebou.

Mluvíme-li dále o sudých a lichých elementech v řadě apod., rozumíme tím příslušnou pozici, počítáno zleva.

3.2. Definujeme nyní některé podmnožiny všech řad.

3.2.1.  $P_j$  definujeme jako množinu všech řad, kde žádným lichým elementem kromě prvního není **a**.

Tuto množinu interpretujeme jako jambické verše.

$P_v$  definujeme jako množinu všech řad, které jsou prvky  $P_j$  a kde druhý element není **c**.

V naší interpretaci jde o první variantu jambu („čistý“ jamb).

$P_v$  definujeme jako množinu všech řad, které jsou prvky  $P_j$  a jako druhý element mají **c**.

V naší interpretaci jde o druhou variantu jambu (jamb s „daktylským“ začátkem).

$P_i$  definujeme jako množinu všech řad, kde žádným sudým elementem není **a** a první element není **e**.

V naší interpretaci jde o verše trochejské.

Pozn.: Nevymezujeme tu jiné množiny (např. verše daktylské).

3.2.2. Dále budeme pracovat s množinami  $S_i$  pro  $5 \leq i \leq 12$ , definovanými jako řady o **i** elementech.

3.3. Průnik  $S_i \triangle P_k$ , kde **k** je za **t**, **u**, **v**, budeme nazývat rozměr a značit  $M_{ik}$ . Můžeme tedy některé řady třídit podle toho, zda patří např. do  $M_{6u}$  (první varianta šestislabičného jambu),  $M_{8t}$  (osmislabičný trochej) atd.

3.4. Dále pracujeme s rozklady množin  $M_{ig}$  (kde  $5 \leq i \leq 12$ , **g** je za **u**, **v**), definovanými na základě toho, na kterých **s u d ý c h** pozicích v dané řadě je element **e**. Každé posloupnosti sudých slabik řady z  $M_{ig}$  přiřadíme jednoznačně přirozené číslo **m** a jednotlivé prvky rozkladu označujeme jako  $M_{igm}$ . Podobně pro množiny  $M_{it}$  pracujeme s rozklady definovanými na základě toho, na kterých **l i c h ý c h** pozicích v dané řadě je element **e**.

Interpretujeme prvky těchto rozkladů jako typy rytmické realizace metra (srov. str. 68).

Pozn.: Počet **s v R** není nijak přímo vymezen; je dán pouze nepřímě jako důsledek omezení, kterým musí vyhovovat daný rozměr.

#### 4. Úkoly pro výpočet pravděpodobnostního modelu

4.1.1. Pro každý  $M_{ik}$  stanovit množinu všech **R** jako kombinací **s** podle repertoáru **s** (2.2.) a definice příslušného  $M_{ik}$  (3.2. a 3.3.). Počet takových **R** není znám předem.

4.1.2. Pro každou **R** stanovit teoretickou pravděpodobnost na základě 2.3.

$$p(R_i) = p_1(s_1) \cdot p_2(s_2) \cdot \dots \cdot p_n(s_n),$$

kde  $s_i$  jsou segmenty řady.<sup>4</sup>

<sup>4</sup> Pavel Novák v recenzi sborníku Prague Studies in Mathematical Linguistics 2, 1967, v němž vyšla širší anglická verze této studie (recenze byla otištěna v čas. Kybernetika 4, 1968, str. 183), právem upozornil na nemožnost srovnávání pravděpodobností jednotlivých  $R_n$  v jistém  $M_{ikm}$ , jestliže  $n$  má pro různá  $R$  proměnlivou hodnotu. Po konzultaci s dr. Zítkem z Matematického ústavu ČSAV jsme dospěli k této závěru:

Hodnoty  $p(s_i)$  nejsou navzájem nezávislé: výběr každého  $s_i$  je omezen jednak metrickou normou, jednak už učiněnými volbami  $s_1, s_2, \dots, s_{i-1}$ . Např. pro  $M_{9t}$  (devítislabičný trochej), jsou-li  $s_1, s_2, s_3, s_4$  určeny jako dvouslabičné takty, může být  $s_5$  výlučně jednoslabičný takt, a tedy  $p(s_5) = 1$ ; podobně je omezen i výběr  $s_4$ , neboť po předchozí sestavě taktů už mohou následovat jedině takty  $xxx, xx/x, x/xx$ , atd.

Jak je vidět už z příkladu, uvederá námitka komplikuje proponovaný výpočet, je však možné se s ní vyrovnat. Při násobení pravděpodobností pro stanovení  $p(R)$  je nutno pravděpodobnost každého  $s_i$  přepočítat vzhledem k základu, jímž je soubor všech segmentů v daném metru v dané pozici použitelných; v uvedené realizaci metra  $M_{9t}$  bude tedy  $p(s_1)$  nižší než  $p(s_2)$ ,  $p(s_2)$  nižší než  $p(s_3)$  atd., a to v míře dané repertoárem taktů v příslušné pozici vyloučených. Tuto modifikaci vyjadřujeme prozatím připojením indexu k symbolu pravděpodobnosti každého  $s_i$ .

Potom  $p(M_{ik}) = 1$ ; tedy teoretické pravděpodobnosti každé  $R_i$  i každého typu  $M_{ikm}$  se rovnají jejich teoretickým relativním četnostem. Odpadají tím výpočty uvedené v anglické verzi sub 4.4.1. a 4.4.2. Za těchto podmínek nelze ovšem na základě navržených výpočtů získat srovnávací charakteristiku jednotlivých meter z hlediska jejich „obtížnosti“, jak jsme původně předpokládali.

4.2. Stanovit teoretickou pravděpodobnost každého  $M_{ik}$  sumací pravděpodobností všech  $\mathbf{R}$ , vyhovujících jeho definici:

$$p(M_{ik}) = \sum_{i=1}^q p(\mathbf{R}_i),$$

kde  $q$  je počet různých řad v  $M_{ik}$ .

4.3.1. Množiny všech  $\mathbf{R}$ , vyhovujících definicím jednotlivých  $M_{ik}$ , rozdělit v podmnožiny  $M_{ikm}$  podle definic 3.4.

4.3.2. Pro každý  $M_{ikm}$  stanovit jeho teoretickou pravděpodobnost sumací pravděpodobností všech  $\mathbf{R}$ , jež vyhovují jeho definici

$$p(M_{ikm}) = \sum_{i=1}^r p(\mathbf{R}_i),$$

kde  $r$  je počet různých řad v  $M_{ikm}$ .

---

Základní údaj pravděpodobnostního modelu, tj zjištění teoretických relativních četností jednotlivých typů  $M_{ikm}$  v rámci daného metra, však můžeme stanovit.

Poznamenáváme ještě, že náš původní výpočet  $p(\mathbf{R}_i)$  vycházel z běžného postupu, jakého se dosud v matematické teorii verše používalo (např. ve známé *Kondratovově* práci *Matematika i poezija*, u nás v pracích J. Levého), a že tedy uvedená oprava, podnětená Novákovou připomínkou a umožněná pomocí dr. Zítka, může mít jistý význam pro řešení daného problému v širším měřítku. Proto také jsme se rozhodli zatížit korekturu stati touto obsáhlou poznámkou.

## ON THE WAY TO A PROBABILITY MODEL OF THE CZECH VERSE

This paper has been written in connection with a study based on an extensive collection of material and undertaken by a group for the statistical investigation of verse at the ČSAV Institute of Czech Literature in Prague. The authors deal with the distributions they have found out to be the fundamental material for the statistical investigation of verse structure, with the metrical norm of the Czech syllabic-accentual verse, and with a probability model of the Czech verse. A more detailed formulation of their considerations (in English) was published in Prague Studies in Mathematical Linguistics II, 1967.