

**UNIVERZITA PALACKÉHO V OLOMOUCI**

Filozofická fakulta

Katedra asijských studií

**MAGISTERSKÁ DIPLOMOVÁ PRÁCA**

**Aplikácia Menzerath-Altmanovho zákona na súčasnú hovorenú  
čínštinu na úrovni výpovede**

An Application of the Menzerath-Altmann Law to Contemporary Spoken  
Chinese on the Language Level - Statement

OLOMOUC 2013

Bc. Jana Ščigulinská

Vedúci práce: Mgr. Ondřej Kučera

## Zadanie Diplomovej práce



Čestne prehlasujem, že som magisterskú diplomovú prácu s názvom *Aplikácia Menzerath-Altmanovho zákona na súčasnú hovorenú čínštinu na úrovni výpovede* vypracovala samostatne a použila som len uvedené pramene a literatúru.

V Olomouci, dňa 22.04.2013

-----  
Bc. Jana Ščigulinská

## **Anotácia**

Cieľom tejto diplomovej práce je overiť platnosť Menzerath-Altmanovho zákona na súčasnej hovorenej čínštine, s dôrazom na vzájomný vzťah výpovede a prízvukového taktu meraného v slabikách. Výskum prebiehal v dvoch fázach, pričom v prvej fáze bol pomocou kvantitatívnej metódy skúmaný vzájomný vzťah konštruktú a konštituentu na troch jazykových úrovniach. Pôvodná hypotéza stanovuje, že vzájomné vzťahy jazykových jednotiek hovorenej čínštiny podliehajú zákonitostiam MALu na všetkých troch jazykových úrovniach. Druhá fáza výskumu sa zamerala na jednu konkrétnu jazykovú úroveň, v rámci ktorej sa overovala hypotéza, na základe ktorej platí, že vplyv umelého rytmu hudobného žánru dokáže ovplyvniť prirodzený rytmus reči. Dáta získané z oboch častí výskumu sú zobrazované pomocou tabuliek a grafov. Výsledky výskumu súvisiace s oboma hypotézami sú zhrnuté závere práce.

**Kľúčové slová:** kvantitatívna lingvistika, Menzerath-Altmanov zákon, hovorená čínština, segmentácia, jazykové jednotky, výpoveď.

Touto cestou by som chcela poďakovať Mgr. Ondrejovi Kučerovi za vedenie, odborný dohľad a cenné rady, ktorými prispel k vypracovaniu tejto diplomovej práce. Taktiež vrelo ďakujem Mgr. Martine Benešovej, Ph.D. za neoceniteľné konzultácie a rady v oblasti lingvistiky a matematiky, vďaka ktorým mohla táto práca vzniknúť.

# Obsah

Zoznam tabuliek.....	7
Zoznam grafov .....	8
Edičná poznámka .....	10
Úvod.....	11
1. Lingvistika .....	14
1.2 Matematická lingvistika.....	14
1.2.1 Kvantitatívna lingvistika.....	15
1.2.2 Algebraická lingvistika .....	17
2. Menzerath-Almannov zákon.....	19
2.1 Vývoj Menzerath-Altmannovho zákona.....	19
2.2 Matematické a slovné vyjadrenie MAL.....	21
3. Segmentácia hovoreného textu .....	25
3.1 Hovorený text.....	26
3.2 Textové jednotky.....	27
3.2.1 Hláska.....	28
3.2.2 Slabika.....	28
3.2.3 Prízvukový takt .....	29
3.2.4 Výpoveď .....	30
3.2.5 Replika .....	32
3.3 Jazykové úrovne v rámci čínštiny .....	33
3.3.1 Segment.....	34
3.3.2 Slabika.....	34
3.3.3 Rozdiely v názvosloví.....	35
3.4 Rytmus reči .....	35
3.5 Definovanie jednotiek .....	36
4. Problematika vhodnej transkripcie.....	38
4.1 Fonetické problémy a porovnanie pinyin a českej transkripcie.....	39
4.1.1 Finálna posícia i vs. e.....	39
4.1.2 Iniciálne aspirované spoluhlásky .....	40
4.1.3 Iniciálne palatálne hlásky.....	40
4.1.4 Iniciálna spoluhláska q.....	41
4.1.5 Trojhlásky /iou/, /uei/ .....	41
5. Výsledky a diskusia .....	44
5.1 U1: Replika meraná vo výpovediach, ktoré sú merané v priemernom počte ich prízvukových taktov .....	45
5.2 U2: Výpoveď meraná v prízvukových taktach, ktoré sú merané v priemernom počte ich slabík.....	54
5.3 U3: Prízvukový takt meraný v slabikách, ktoré sú merané v priemernom počte ich hlások .....	67
Záver .....	73
Resumé.....	76
Zoznam literatúry .....	77

## Zoznam tabuliek

Tabuľka 1: Zhrnutie foneticky problematických aspektov .....	42
Tabuľka 2: U1: konštrukt $x_1$ – dĺžka repliky (v počte jej výpovedí), $z_1$ – frekvencia konštruktov, konštituenty $y_1$ – priemerná dĺžka výpovede (v počte prízvukových taktov), výberový súbor A.....	46
Tabuľka 3: U1: konštrukt $x_1$ – dĺžka repliky (v počte jej výpovedí), $z_1$ – frekvencia konštruktov, konštituenty $y_1$ – priemerná dĺžka výpovede (v počte prízvukových taktov), výberový súbor B, verzia 1 a 2. ....	46
Tabuľka 4: Prehľad parametrov <sup>b</sup> MAL a koeficientov determinácie $R^2$ na U1.....	52
Tabuľka 5: U2: konštrukt $x_2$ – dĺžka výpovede (v počte jej prízvukových taktov), $z_2$ – frekvencia konštruktov, konštituenty $y_2$ – priemerná dĺžka prízvukového taktu (v počte slabík), výberový súbor A. ....	56
Tabuľka 6: U2: konštrukt $x_2$ – dĺžka výpovede (v počte jej prízvukových taktov), $z_2$ – frekvencia konštruktov, konštituenty $y_2$ – priemerná dĺžka prízvukového taktu (v počte slabík), výberový súbor B, verzia 1 a 2. ....	56
Tabuľka 7: U2: konštrukt $x_2$ – dĺžka výpovede (v počte jej prízvukových taktov), $z_2$ – frekvencia konštruktov, konštituenty $y_2$ – priemerná dĺžka prízvukového taktu (v počte slabík), výberový súbor C, verzia 1.....	57
Tabuľka 8: U2: konštrukt $x_2$ – dĺžka kolónu (v počte jeho segmentov), $z_2$ – frekvencia konštruktov, konštituenty $y_2$ – priemerná dĺžka segmentu (v počte slabík), výberový súbor C, verzia 2. ....	57
Tabuľka 9: Prehľad parametrov <sup>b</sup> MAL a koeficientov determinácie $R^2$ na U2.....	65
Tabuľka 10: U3: konštrukt $x_3$ – dĺžka prízvukového taktu (v počte jeho slabík), $z_3$ – frekvencia konštruktov, konštituenty $y_3$ – priemerná dĺžka slabiky (v počte jej hlások), výberový súbor A. ....	67
Tabuľka 11: U3: konštrukt $x_3$ – dĺžka prízvukového taktu (v počte jeho slabík), $z_3$ – frekvencia konštruktov, konštituenty $y_3$ – priemerná dĺžka slabiky (v počte jej hlások), výberový súbor B. ....	68
Tabuľka 12: Prehľad parametrov <sup>b</sup> MAL a koeficientov determinácie $R^2$ na U3.....	72



## Zoznam grafov

Graf 1.A: U1 (replika vs. výpoveď) – súbor A: vizualizácia súboru dát v tabuľke 2, pomocou programu Microsoft Excel. ....	48
Graf 1.Ar: U1 (replika vs. výpoveď) – súbor A: vizualizácia súboru dát v tabuľke 2, pomocou programu R.....	48
Graf 1.B1: U1 (replika vs. výpoveď) – súbor B verzia 1: vizualizácia súboru dát v tabuľke 3, pomocou programu Microsoft Excel. ....	50
Graf 1.B1r: U1 (replika vs. výpoveď) – súbor B verzia 1: vizualizácia súboru dát v tabuľke 3, pomocou programu R. ....	50
Graf 1.B2: U1 (replika vs. výpoveď) – súbor B verzia 2: vizualizácia súboru dát v tabuľke 3 , pomocou programu Microsoft Excel. ....	51
Graf 1.B2r: U1 (replika vs. výpoveď) – súbor B verzia 2: vizualizácia súboru dát v tabuľke 3, pomocou programu R. ....	51
Graf 2.A: U2 (výpoveď vs. prízvukový takt) – súbor A: vizualizácia súboru dát v tabuľke 5, pomocou programu Microsoft Excel. ....	58
Graf 2.Ar: U2 (výpoveď vs. prízvukový takt) – súbor A: vizualizácia súboru dát v tabuľke 2, pomocou programu R. ....	59
Graf 2.B1: U2 (výpoveď vs. prízvukový takt) – súbor B verzia 1: vizualizácia súboru dát v tabuľke 6, pomocou programu Microsoft Excel. ....	60
Graf 2.B1r: U2 (výpoveď vs. prízvukový takt) – súbor B verzia 1: vizualizácia súboru dát v tabuľke 6 pomocou programu R. ....	60
Graf 2.B2: U2 (výpoveď vs. prízvukový takt) – súbor B verzia 2: vizualizácia súboru dát v tabuľke 6, pomocou programu Microsoft Excel. ....	61
Graf 2.B2r: U2 (výpoveď vs. prízvukový takt) – súbor B verzia 2: vizualizácia súboru dát v tabuľke 6 pomocou programu R. ....	61
Graf 2.C1: U2 (výpoveď vs. prízvukový takt) – súbor C verzia 1: vizualizácia súboru dát v tabuľke 7, pomocou programu Microsoft Excel. ....	63
Graf 2.C1r: U2 (výpoveď vs. prízvukový takt)– súbor C verzia 1: vizualizácia súboru dát v tabuľke 7, pomocou programu R. ....	63
Graf 2.C2: U2 (výpoveď vs. prízvukový takt) – súbor C verzia 2: vizualizácia súboru dát v tabuľke 8, pomocou programu Microsoft Excel. ....	64
Graf 2.C2r: U2 (výpoveď vs. prízvukový takt)– súbor C verzia 2: vizualizácia súboru dát v tabuľke 8 pomocou programu R. ....	64

Graf 3.A: U3 (prízvukový takt vs. slabika) – súbor A: vizualizácia súboru dát v tabuľke 10, pomocou programu Microsoft Excel. ....	69
Graf 3.Ar: U3 (prízvukový takt vs. slabika) – súbor A: vizualizácia súboru dát v tabuľke 10, pomocou programu R.....	69
Graf 3.B: U3 (prízvukový takt vs. slabika) – súbor B: vizualizácia súboru dát v tabuľke 11, pomocou programu Microsoft Excel. ....	70
Graf 3.Br: U3 (prízvukový takt vs. slabika) – súbor B: vizualizácia súboru dát v tabuľke 11, pomocou programu R.....	70

## **Edičná poznámka**

V práci je použitá oficiálna čínska transkripcia s názvom pinyin, rovnako ako oficiálna česká transkripcia vytvorená prof. Švarným. V prípade použitia čínskych znakov používam ich zjednodušenú formu a príslušný textový znakový font.

Bibliografické citácie boli vytvárané pomocou citačnej normy ČSN ISO: 690 a v celej práci používam bibliografický systém poznámok pod čiarou.

## Úvod

Kvantitatívna lingvistika je jedným z odvetví jazykovedy, ktoré kombinuje ako lingvistické, tak aj matematické metódy, aby sa pomocou nich dopátrala k novým poznatkom, alebo overila tie staré. Aj v tejto práci bude kvantitatívna lingvistika zohrávať úlohu sprostredkovateľa a pomocou jednej z jej metód bude skúmaná súčasná hovorená čínština. Touto metódou je jeden z dôležitých lingvistických zákonov a to konkrétne Menzerath-Altmanov zákon, ktorý objasňuje vzťahy medzi jednotlivými jednotami v jazyku. V tejto diplomovej práci, ktorá vznikla na základe projektu *Segmentace pro testování Menzerath-Altmanova zákona a hypotéz s ním souvisejících FF\_2012\_035*, popisujem výskum, ktorý sa zaoberal aplikáciou Menzerath-Altmanovho zákona na hovorenú čínštinu. Jedným z výstupov tohto projektu je článok „An Application of the Menzerath-Altman Law to Contemporary Spoken Chinese“<sup>1</sup>, v ktorom sú publikované určité časti tejto diplomovej práce. Aj keď moja práca vychádza z tohto projektu, je konkrétnejšie zameraná hlavne na vzťah medzi jazykovou úrovňou výpovede meranej v prízvukových taktoch, ktoré sú merané v priemernom počte ich slabík.

Cieľom práce je overiť hypotézu, ktorá stanovuje, že hovorená čínština bude podliehať zákonitostiam popísaným a definovaným Menzerath-Altmanovým zákonom na všetkých skúmaných jazykových úrovniach. Na mnou bližšie skúmanom vzťahu výpovede a prízvukového taktu meraného v slabikách budem navyše overovať hypotézu, ktorá tvrdí, že umelo vytvorený rytmus hudobného žánru akým je rap dokáže ovplyvniť prirodzenú rytmiku reči interpreta, v tomto prípade čínštinu.. Prvá hypotéza bude overovaná na dvoch výberových súboroch na troch úrovniach, druhá hypotéza bude overovaná na troch výberových súboroch na jednej konkrétnej úrovni.

Samotný výskum sa skladá z dvoch častí, pričom prvá fáza bola súčasťou vyššie spomínaného projektu a k druhej fáze som pristúpila z dôvodu porovnania výsledkov z prvej fázy s tretím kontrastným výberovým súborom. Súborny skúmané v prvej fáze výskumu pochádzajú z projektu Korpus hovorové čínštiny FF\_2010\_042, ktorý sa

---

<sup>1</sup> SCHUSTEROVÁ, Denisa. ŠČIGULINSKÁ, Jana et al. **An Application of the Menzerath-Altman Law to Contemporary Spoken Chinese**. *Czech and Slovak Linguistic Review*, 2013, č. 2. (v tisku).

zaoberal transkripčným prepisom nahrávok čínštiny do čínskej fonetickej abecedy. Tretí kontrastný výberový súbor pochádza z diela vynikajúceho českého sinológa prof. Švarného a konkrétne ide o výňatok z materiálu určeného pre študijné účely s názvom Hovorová čínština v príkladoch.

Práca má klasické členenie ku ktorému patrí úvod, jadro a záver, za ktorým nasledujú zoznamy použitých primárnych a sekundárnych zdrojov, a na konci práce sa nachádzajú príslušné prílohy. Samotné jadro práce je členené na niekoľko navzájom súvisiacich kapitol, z ktorých prvé štyri sa zaoberajú prevažne teoretickou a metodologickou časťou výskumu. V prvých kapitolách oboznamujem čitateľa s kvantitatívnou lingvistikou a jej dôležitosťou v rámci lingvistických skúmaní, zaraďujem ju do konkrétnej skupiny v rámci všeobecnej jazykovedy, poukazujem na jej prínos v overovaní lingvistických hypotéz a taktiež na potrebu spolupráce prírodných a humanitných vied. Ďalej predstavujem jeden z nástrojov kvantitatívnej lingvistiky a to konkrétne Menzerath-Altmanov zákon. Vzhľadom na jeho dôležitosť, v príslušnej kapitole stručne približujem oboch vedcov, ktorí prispeli k formulácii tohto zákona a taktiež sa venujem samotnému zákonu, jeho formulácii a definíciám vyjadreným slovne aj matematicky.

V nasledujúcich kapitolách sa už zaoberám metodologickým popisom výskumu a to počínajúc od výberu súborov, stručného popisu čínskej fonetiky až po výber vhodnej transkripcie a definovanie jednotiek, na základe ktorých potom bolo možné pristúpiť k samotnej segmentácii textu a získaniu príslušných dát. Všetky tieto aspekty značne ovplyvňovali samotný výskum, nakoľko definície jednotlivých jednotiek sa pomerne výrazne líšia a bolo nutné pristúpiť k selektívnemu výberu definície vhodnej pre tento výskum. V príslušnej kapitole popisujem rôzne pohľady na túto problematiku, avšak najväčší priestor prenechávam definovaniu úrovne výpovede, nakoľko vzťah konštruktú a konštituentu v spojení s touto jazykovou jednotkou je ten, ktorý ma v práci zaujíma najviac. V závere kapitoly sa snažím o zhrnutie definícií jednotlivých úrovní a o jednoznačné formulovanie konečnej charakteristiky tej ktorej úrovne, na základe ktorých sa k týmto úrovniam pristupovalo.

Posledná a najrozsiahlejšia kapitola sa zaoberá samotnou prezentáciou konečných získaných výsledkov pomocou tabuliek a grafov. Tabuľky a grafy sú sprevádzané príslušnými komentármi a analýzami, ktoré slúžia na interpretáciu a analýzu

jednotlivých výsledkov vizualizovaných v príslušných tabuľkách a grafoch. V tejto kapitole tým pádom rovnako ako k interpretácii výsledkov pristupujem aj k diskusii, v ktorej hodnotím jednotlivé výsledky a vyvodzujem závery. Taktiež tu dochádza aj k posúdeniu hypotéz, nakoľko a či sa tieto hypotézy potvrdili, alebo naopak vyvrátili. Celý výskum, jeho výsledky a možné smery ďalšieho výskumu popisujem a zhŕňam v závere práce.

V práci sa odkazujem na rozličnú českú a zahraničnú literatúru, s ktorou som sa v priebehu výskumu oboznámila a ktorá sa stala oporným bodom teoretickej časti mojej záverečnej práce. Poznatky o kvantitatívnej lingvistike som čerpala hlavne z diel popredných českých autorov ako je Luděk Hřebíček, Marie Těšitelová a ďalší. Pri definovaní jednotiek som sa opierala o rôzne práce a autorov pôsobiacich v českej lingvistike ako napríklad Franišek Daneš, Olga Müllerová, Marie Krčmová a mnohých ďalších. V prípade čínštiny a poznatkov o nej som sa uchýľovala predovšetkým k dielam českého sinológa Oldřicha Švarného a Hany Třískovej. Okrem zdrojov dostupných v češtine som spolupracovala a konzultovala aj so zahraničnou literatúrou, ktorú spolu s ostatnými zdrojmi uvádzam v zozname bibliografie.

Čínština je z hľadiska kvantitatívnej lingvistiky pomerne neprebádanou oblasťou, preto si myslím, že je dôležité zaoberať sa týmto starým a stále rozšírenejším jazykom aj z tejto stránky. Jedným z cieľov tejto práce je preto vrhnúť o trochu viac svetla na problematiku aplikácie Menzerath-Altmanovho zákona na súčasnú hovorenú čínštinu a na základe analyzovaných výsledkov potom určiť a prípadne rozšíriť smer ďalšieho možného bádania.

# 1. Lingvistika

Čo znamená slovo lingvistika je všeobecne známe. Tento pojem označuje vedu, ktorá sa zaoberá skúmaním jazyka, skúmaním jazykového systému, jeho používania, vývine, vzťahov s mimojazykovými skutočnosťami a pod. V rámci lingvistiky rozlišujeme niekoľko jej odvetví: formálna lingvistika, korpusová lingvistika, matematická lingvistika, pod ktorú sa radí napr. kvantitatívna, algebraická a počítačová lingvistika.<sup>2</sup> Jazykoveda ako taká ďalej pod sebou zahŕňa množstvo ďalších pod oblastí, ktoré sa venujú konkrétnym čiastkovým odvetviám lingvistiky ako napríklad: fonetika (náuka o zvukoch v jazyku), morfológia (náuka o stavbe a skladaní slov), sémantika (tá sa zaoberá konkrétnym významom jednotlivých slov), prípadne syntax (ktorý sa zaoberá stavbou a skladbou vety). Čím konkrétne sa ale zaoberá kvantitatívna lingvistika?

Hřebíček vo svojej práci ilustratívne používa citát slávneho fyzika Ernesta Rutherforda (1871-1937), ktorý v preklade znie: „Kvalitatívni není nic jiného než uboze kvantitatívni“<sup>3</sup> Čo to ale znamená v spojení s vedou o jazyku? Slovo kvantitatívny vytvára v mysli predpoklad, že sa budeme zaoberať predovšetkým počtami a číslami. Podľa Šimkovej *Výberového slovníka termínov z korpusovej lingvistiky* tento pojem označuje: „odbor matematickej lingvistiky, [ktorý] skúma jazyk pomocou štatistických metód so zameraním na frekvenciu jazykových jednotiek a javov.“<sup>4</sup>

## 1.2 Matematická lingvistika

K matematickej lingvistike, ktorá vznikla na začiatku šesťdesiatych rokov dvadsiateho storočia radí Černý dva základné prúdy, a to kvantitatívnu lingvistiku a algebraickú lingvistiku, ktorá pracuje hlavne s formálnou logikou. K týmto dvom prúdom Černý priraduje aj strojovú lingvistiku, ktorá využíva poznatky oboch vyššie zmienených odvetví matematickej lingvistiky, aplikuje tieto poznatky v praxi, k čomu používa počítaacie stroje a matematickú informatiku.<sup>5</sup> Podľa Černého je matematická lingvistika

---

<sup>2</sup> ŠIMKOVÁ, Mária. *Výberový slovník termínov z korpusovej lingvistiky*. [2013-03-04] Dostupné z: <http://korpus.juls.savba.sk/attachments/what/2006-simkova-vyberovy%20slovník%20terminov.pdf>

<sup>3</sup> HŘEBÍČEK, Luděk. *Vyprávění o lingvistických experimentech s textem*. Praha: Academia, 2002. ISBN 80-200-0973-6. s. 52.

<sup>4</sup> ŠIMKOVÁ.

<sup>5</sup> ČERNÝ, Jiří. *Stručné dějiny lingvistiky III*. Olomouc: Univerzita Palackého, 1981. s. 5.

tzv. „pomezní disciplína“, která rovnakým podílem využívá matematiku aj lingvistiku, s čím, ako sám autor vraví, nesúhlasí jedna z najväčších znalkýň kvantitatívnej lingvistiky Marie Těšitelová. Tá sa prikláňa ako aj mnohí ďalší k teórii, že kvantitatívna lingvistika je „jen lingvistická disciplína používajúci matematických metod.“<sup>6</sup>

Používanie matematických metód v lingvistike nebol ojedinelý jav ani pred vznikom matematickej lingvistiky v roku 1957. Tieto metódy boli prevažne kvantitatívne alebo štatistické. Podľa Černého sa na túto skutočnosť môžeme pozerat' spôsobom, že kvantitatívna lingvistika je zo všetkých odvetví matematickej lingvistiky jediným odvetvím, ktoré má určitú tradíciu.<sup>7</sup>

### 1.2.1 Kvantitatívna lingvistika

Okrem názvu kvantitatívna lingvistika sa v literatúre objavuje aj pomenovanie štatistická lingvistika. Hřebíček však upozorňuje, že to rozhodne neznamená, že čísla budú tou hlavnou stránkou výskumu a naše myslenie bude obmedzené len na ne.<sup>8</sup> Ako sám tvrdí: „Být kvantitativní znamená uvažovat v matematických veličinách...Nějaké konkrétní číslo, numerická hodnota, by měla být ve vědeckém myšlení jen zástupcem veličiny pro daný případ vyhovující určitým daným okolnostem.“<sup>9</sup>

Prvé vážnejšie pokusy zaviesť kvantitatívne hľadisko do jazykovedy začali naberať na sile na prelome 19. a 20. storočia. V tomto období sa matematika začala uplatňovať aj v rámci iných vedných odborov a tým pádom sa dostala aj do jazykovedy. Ako píše Černý, jedným z priekopníkov používania matematiky v jazykovede bol ruský matematik V. J. Buňakovskij. Tejto tematike sa venoval aj Herman Paul, ktorého záujem bol zameraný na štatistiku, ďalej Augustin Seydler, ktorý sa zaoberal pravdepodobnosťou a prípadne poľský vedec Jan Baudouin de Courtenay, podľa ktorého je možné v lingvistike používať základnú, ale aj vyššiu matematiku.<sup>10</sup> Tieto osobnosti však boli prevažne vedci prírodných vied a ich záujem sa koncentroval predovšetkým na frekvenciu. Prvým lingvistom, ktorý sa zaoberal kvantitatívnou

---

<sup>6</sup> ČERNÝ, *Stručné*, s. 6-7.

<sup>7</sup> Tamtiež s. 7.

<sup>8</sup> HŘEBÍČEK, *Vyprávění*, s. 52.

<sup>9</sup> Tamtiež, s. 52.

<sup>10</sup> ČERNÝ, Jiří. *Dějiny lingvistiky*. Olomouc: Votobia, 1996. ISBN 80-85885-96-4. s. 249.



lingvistikou bol Američan W. D. Whitney. Objektom jeho záujmu bola ale taktiež frekvencia, pozorovaná na anglických hláskach.<sup>11</sup>

Kvantitatívna lingvistika tak ako každá iná veda stavia na hypotézach a ich potvrdení, resp. vyvrátení. Čísla a matematika sú pri tejto úlohe pomocnou silou, vďaka ktorej sa k danému cieľu môžeme dostať. Neznamená to však, že kvantitatívny lingvista musí zákonite byť geniálnym matematikom. Matematika len sprostredkováva cestu k cieľu, je takzvanou „barličkou“, ktorá nám umožňuje dopátrať sa resp. priblížiť sa k potvrdeniu alebo vyvráteniu určitej hypotézy. Ako hovorí Hřebíček v kvantitatívnej lingvistike ide:

„o snahu dát výrokom o jazyce charakter vědecké teorie, která je testovatelná a je formulována s veškerým možným rizikem: buď platí, což je stvrzeno testem, anebo neplatí, a tedy nemá smysl. Cíl kvantitativní lingvistiky je v podstatě dvojitý:

A. Docílit, aby hypotéza byla formulována ‚na ostro‘.

B. Umožnit testování hypotézy.“<sup>12</sup>

Kvantitatívna lingvistika je dôležitým prostriedkom na prepojenie humanitných a exaktných prírodovedných vied. Ako tvrdí Wimmer: „Veda, ktorej chýbajú hypotézy, je protoveda a veda, ktorej hypotézy sú netestovateľné, je pseudoveda.“<sup>13</sup> Vznikom kvantitatívnej lingvistiky však dochádza k tomu, že pre potvrdenie hypotézy formulovanej na kvalitatívnej úrovni sú nutné kvantitatívne metódy a postupy. Hřebíček podporuje tvrdenie, že vzniknutie požiadavky aplikovať tieto metódy, menovite napríklad matematická štatistika, alebo teória pravdepodobnosti, stavia lingvistom do cesty prekážky.<sup>14</sup> Podľa Wimmera je to ale nutné, nakoľko poznanie určitej veci je vytvárané prevažne využitím kvalitatívnych pojmov, avšak pre presnejšie vyjadrenie je nevyhnutné prejsť k pojmom kvantitatívnym. Wimmer zároveň tvrdí, že vedy, ktoré používajú kvantitatívne metódy sa vyvíjajú rýchlejšie.<sup>15</sup> Pri kvantitatívnej lingvistike je dôležité uvedomiť si, že tento odbor kombinuje humanitné a exaktné vedy, preto je nutné, aby príslušníci týchto odborov navzájom spolupracovali. Wimmer opisuje

---

<sup>11</sup> ČERNÝ, *Stručné*, s. 8-9.

<sup>12</sup> HŘEBÍČEK, *Vyprávění*, s. 19.

<sup>13</sup> WIMMER, Gejza. *Úvod do analýzy textov*. Bratislava: Veda, 2003. ISBN 80-224-0756-9. s. 13.

<sup>14</sup> HŘEBÍČEK, Luděk. **Filologie versus lingvistika**. *Vesmír* [online] 2008, roč. 87, č. 7. [cit. 2012-04-15]. Dostupné z: <<http://www.vesmir.cz/clanek/filologie-versus-lingvistika>>.

<sup>15</sup> WIMMER, *Úvod*, s. 14.

jednotlivé fázy výskumu v šiestich krokoch, ktoré vyžadujú vzájomnú spoluprácu medzi lingvistom a matematikom, kde lingvista formuluje kvalitatívnu hypotézu, matematik na jej základe stanoví matematickú formuláciu hypotézy, následne dochádza k výberu dát, testovaniu a záver pozostáva zo štatistickej a filologickej interpretácie.<sup>16</sup> Podľa môjho názoru kombináciou a vzájomnou spolupracou dvoch odlišných odborov je možné v rámci kvantitatívnej lingvistiky dosiahnuť lepšie a presnejšie výsledky, nakoľko humanitný odbor poskytne testovateľnú hypotézu a matematika zas ponúkne vhodné metódy na jej overenie. Z tohto dôvodu je aj výskum v tejto práci realizovaný formou kombinácie lingvistickej hypotézy overovanej pomocou kvantitatívnej metódy.

### 1.2.2 Algebraická lingvistika

Ako už bolo spomínané vyššie, kvantitatívna lingvistika spadá pod oblasť matematickej lingvistiky. Spolu s ňou do tejto oblasti spadá aj algebraická lingvistika. Hřebíček tieto dve oblasti od seba odlišuje a definuje nasledovne: kvantitatívna lingvistika je podľa jeho definície empirický obor, ktorý používa testovateľné teórie a pomocou nich posúva vpred poznanie jazyka. Pri tejto úlohe kvantitatívni lingvisti často používajú jazykovú štatistiku.<sup>17</sup>

Na druhú stranu cieľom algebraickej lingvistiky je: „analýza nebo vytváření řetězců nějakých symbolů, jež mají nebo mohou mít vztah k řetězcům, které lze pozorovat také v textech.“<sup>18</sup> Algebraická lingvistika sa rozvíjala predovšetkým v období 19. storočia spolu s formálnou logikou. Na základe spolupráce týchto dvoch odvetví, dosiahla matematika „vysoký stupeň abstrakcie“ a tým pádom sa rozšírilo jej využitie v iných vedných odvetviach, pričom jedným z nich bola aj jazykoveda. Tieto nové metódy a postupy v lingvistike už ale nemajú kvantitatívny charakter.<sup>19</sup> Podľa Hřebíčkovho názoru taktiež generatívna lingvistika je súčasťou algebraického oboru, pri ktorom je nutné definovať určité jednotky, prípadne kategórie, a taktiež musí byť splnená druhá podmienka a to, definovanie substitučných vzťahov medzi jednotkami, prípadne kategóriami.<sup>20</sup> Hlavným impulzom pre vznik algebraickej lingvistiky v 50. a 60. rokoch 20. storočia bola práca na strojových prekladoch so samočinnými počítačmi, pri ktorej

---

<sup>16</sup> WIMMER, s. 15-16.

<sup>17</sup> HŘEBÍČEK *Vyprávění*, s. 20.

<sup>18</sup> Tamtiež, s. 22.

<sup>19</sup> ČERNÝ, *Stručné*, s. 52-53.

<sup>20</sup> HŘEBÍČEK, *Vyprávění*, s. 22.

sa prišlo na nevyhnutnosť odstránenia nepresností, ku ktorým prišla vtedajšia jazykoveda. Tieto nedostatky, ako píše Černý, bolo možné odstrániť pomocou prísnej formalizácie lingvistických teórií, zavedenie symbolov a operácií, ktoré umožňujú presnú prácu s týmito symbolmi. Z tohto dôvodu sa lingvistika obrátila k matematike, ktorá umožňovala tieto podmienky splniť.<sup>21</sup>

Z algebraickej lingvistiky ťažil zakladateľ generatívnej resp. transformačnej gramatiky Noam Chomsky, ktorý svojho času kvantitatívnu lingvistiku kritizoval. Chomského gramatika je jedným z najtypickejších príkladov generatívnej gramatiky. Medzi ďalšie príklady môžu patriť „funkčný generatívni popis jazyka“ podľa P. Sgalla a jeho tímu a sovietský „aplikačne-generatívni model jazyka.“<sup>22</sup>

Medzi výrazne kladný rys matematickej jazykovedy patrí to, že v niektorých prípadoch boli pomocou matematickej lingvistiky nastolené úplne nové otázky, ktoré tradičná jazykoveda prehliadla. Ďalšou výhodou je, že matematická lingvistika je schopná exaktnými metódami potvrdiť výsledky, ktoré jazykoveda dosiahla bez použitia matematických metód. Podľa Černého vzájomná znalosť, rešpektovanie a spolupráca matematických lingvistov s tými, ktorí pri svojej práci matematiku nevyužívajú by mohla jazykovede priniesť mnoho pozitívnych výsledkov.<sup>23</sup>

Táto práca sa zameriava predovšetkým na fonologickú štatistiku, ktorá je jednou z oblastí kvantitatívnej lingvistiky. Ako píše Tešitelová, fonologická štatistika zisťuje frekvenciu, distribúciu, vzťahy zvukovej jednotky (na najnižšej úrovni je to najčastejšie foném), prípadne vyšších jednotiek, ktoré vznikli spojením v texte. Následne fonologická štatistika tieto jednotky kvantifikuje a modeluje ich vzájomné vzťahy.<sup>24</sup> Pri fonologickej štatistike je nutné rovnako ako pri všetkých ostatných oblastiach dbať na metodológiu a s ňou spojené problémy. Základom je podľa Tešitelovej definovanie konkrétnych jednotiek skúmaného súboru a ďalším krokom je správny a vhodný výber a rozsah výberového súboru.<sup>25</sup>

---

<sup>21</sup> ČERNÝ, *Stručné*, s. 82-83.

<sup>22</sup> Tamtiež, s. 53.

<sup>23</sup> Tamtiež, s. 83.

<sup>24</sup> TEŠITELOVÁ, Marie. *Kvantitatívni lingvistika*. Praha: Univerzita Karlova, 1987, s. 116.

<sup>25</sup> Tamtiež, s. 116.

## 2. Menzerath-Almannov zákon

Lingvistika už dlhé roky považovala jazyk za systém, ktorý je zložený z jednotiek na rôznych jazykových úrovniach. Po celú dobu sa ale tieto jazykové úrovne skúmali spôsobom, ktorý bol v lingvistike dovtedy bežnou praxou. Jazykové úrovne boli od seba oddelené a ich jednotky boli vyňaté, kvôli možnosti preskúmania vzťahov medzi nimi v rámci viet. Problém nastal, keď sa na svetle sveta zjavila textová lingvistika, ktorej úlohou bolo taktiež pokúsiť sa skúmať nadvetné jazykové úrovne. V tejto situácii, ako píše Benešová sa Menzerath-Altmanov zákon zjavil ako „svetlo v temnote.“<sup>26</sup> Novátorským prístupom, ktorý MAL priniesol, je uchopiť jazykové úrovne vo vzájomných vzťahoch v rámci celého systému.

### 2.1 Vývoj Menzerath-Altmanovho zákona

Menzerath-Altmanov zákon je jedným zo základných zákonov v kvantitatívnej lingvistike. Ako je však zrejmé z názvu tohto zákona, ten nie je dielom len jedného individuálneho jazykovedca. Ako prvý položil základy tohto zákona nemecký jazykovedec Paul Menzerath v roku 1928. Hřebíček vo svojej práci *Vyprávění o lingvistických experimentech s textem* tvrdí, že bez usilovnej a dlhodobej práce jazykovedca a matematika Gabriela Altmanna, ktorý prácu nemeckého jazykovedca rozvíjal, by dnes nebolo možné využívať tento zákon v rozsahu v akom je používaný v kvantitatívnej lingvistike dnes. Rovnako ani výsledky, ktoré počas tejto doby vznikli by bez Altmanna neboli k dispozícii.<sup>27</sup> Altmann rozšíril dovtedy známe poznatky a metódy v kvantitatívnej lingvistike a položil základný kameň kvantitatívnej lingvistiky pre budúce smerovanie jazykovedných skúmaní rôznych jazykov a textov.

Vedci a bádatelia overovali a stále overujú Menzerath-Altmanov zákon na rôznych jazykových úrovniach, v rámci rozličných lingvistických disciplín ako je napríklad fonológia, morfológia, syntax, sémantika atď., na výberových súboroch líšiacich sa z hľadiska jazykovej typológie a štylistiky. Ale rovnako sa skúma vplyv a zhoda s MAL

---

<sup>26</sup> BENEŠOVÁ, Martina. *Kvantitatívni analýza textu se zvláštním zřetelém k analýze fraktální*. Olomouc, 2011. disertační práce (Ph.D.). UNIVERZITA PALACKÉHO V OLOMOUCI. Filozofická fakulta. s. 23.

<sup>27</sup> HŘEBÍČEK, *Vyprávění*, s. 11.

aj v iných vedách ako sú napríklad evolučná biológia alebo genetika.<sup>28</sup> Výskum popísaný v tejto práci sa ale zameriava na skúmanie vzájomných vzťahov jednotlivých jednotiek v rámci fonetiky na čínskej reči.

Samozrejme, kvantitatívna lingvistika nevznikla práve Altmannovým doplnením Menzerath-Altmanovéhó zákona, existovala už dlhú dobu pred ním, ale ako hovorí Hřebíček: „byla většinou užita deskriptivně, zatímco Altmann pochopil, že úkolem této metodologie je důsledně směřovat k *vědecké explanaci*.“<sup>29</sup> Na základe Altmannovho doplnenia daného lingvistického zákona došlo k prehĺbeniu a k prepojeniu lingvistiky a matematiky, vďaka čomu sa otvorili nové dovtedy netušené možnosti jazykovedného skúmania rôznych jazykových súborov a textov.

Vtedajšia lingvistika a lingvisti sa zaoberali hlavne relatívnou samostatnosťou jazykových podsystémov, zatiaľ čo Paul Menzerath si všimol zvláštne skutočnosti. Na najnižšej úrovni je táto skutočnosť popisovaná tým, že hláska/foném je tým kratší, čím dlhší je celok v ktorom sa vyskytuje (čo predstavuje zákon kvantity) a rovnako platí, že čím sa v slove nachádza viac hlások, tým menšia je ich relatívna dĺžka.<sup>30</sup> Z týchto čiastkových pozorovaných skutočností sa neskôr vykryštalizovala definícia, ktorá stále pracuje s konkrétnymi hladinami a jej znenie je: „čím víc je ve slově slabik, tím víc se slabika jaksi smršťuje.“<sup>31</sup> Na tento Menzerathov poznatok neskôr nadviazal Gabriel Altmann, ktorý v roku 1980<sup>32</sup> zaviedol pojmy konštrukt a konštituent a dopomohol k zovšeobecneniu Menzerathovej teórie. Medzi zvoleným konštruktom (čo je jednotka na vyššej jazykovej úrovni) a konštituentom (čo znamená jednotka na najbližšej nižšej jazykovej úrovni) existuje nasledujúci vzťah:

„Čím delší je v jazyce nějaký konstrukt, tím kratší jsou v průměru jeho konstituenty.“<sup>33</sup>

Ako ale sám Hřebíček ďalej poznamenáva, nie je možné, aby sa takýto vzťah bral doslova, nakoľko je nevyhnutné vnímať ho len ako vyjadrenie určitej tendencie v jazyku. Podľa Menzerathovej hypotézy je medzi konštruktom a konštituentom vzťah nepriamej úmernosti. To znamená, že čím je väčší konštrukt, tým menší je konštituent.

---

<sup>28</sup> KUŁACKA, Agnieszka. **The Coefficients in the Formula for the Menzerath-Altman Law.** *Journal of Quantitative Linguistics*, [online]. 2010. roč. 17, č. 4, s. 257-268. [cit. 2012-03-27]. Dostupné z: <<http://dx.doi.org/10.1080/09296174.2010.512160>> s. 257.

<sup>29</sup> HŘEBÍČEK, *Vyprávění*, s. 11.

<sup>30</sup> ALTMANN, Gabriel. **Prolegomena to Menzerath's law.** *Glottometrika*, č.2, 1980. s. 1–10. s. 1.

<sup>31</sup> HŘEBÍČEK, *Vyprávění*, s. 53.

<sup>32</sup> KUŁACKA, s. 258.

<sup>33</sup> HŘEBÍČEK, *Vyprávění*, s. 53.

Dĺžka konštruktu je meraná v konštituentoch, preto, lebo ako vysvetľuje Hřebíček, ich počet bude celé číslo. Ďalej pokračuje vo vysvetlení, že dĺžka konštituentov bude ale rôzna, pretože je vyjadrovaná priemerom jednotlivých zistených dĺžok. To znamená, že priemer bude v drvivej väčšine prípadov desatinné číslo.<sup>34</sup>

## 2.2 *Matematické a slovné vyjadrenie MAL*

Altmann Menzerathovu hypotézu doplnil o predpoklad priamej úmernosti, ktorý hovorí: „čím väčší je konštituent, tým väčší je jeho prírústek.“<sup>35</sup> Tento zákon a tieto dva vzťahy sú vyjadrené matematickou formuláciou Menzerath-Altmanovho zákona ktorá má nasledujúce znenie:

$$y = A \cdot x^{-b}$$

$x$  predstavuje dĺžku konštruktu, ktorý je meraný v jeho konštituentoch

$y$  predstavuje dĺžku konštituentu, ktorý je meraný v jednotkách najbližšej nižšej úrovne (t.j. v jeho konštituentoch)

$A, b$  predstavujú reálne parametre.<sup>36</sup>

Tento vzorec je používaný ako jednoduchá formulácia MAL, existuje však aj zložitejšia, kompletná algebraická formulácia tohto zákona:

$$y = A \cdot x^{-b} \cdot e^{cx}$$

$x$  predstavuje dĺžku konštruktu, ktorý je meraný v jeho konštituentoch

$y$  predstavuje dĺžku konštituentu, ktorý je meraný v jednotkách najbližšej nižšej úrovne

$A, b, c$  predstavujú reálne parametre.<sup>37</sup>

V roku 1984 si Köhler všimol, že hodnota koeficientu  $A$  sa rovná jazykovému konštruktú s empiricky získanou priemernou dĺžkou jedného konštituentu. Rovnako

---

<sup>34</sup> HŘEBÍČEK, Vyprávění, s. 54.

<sup>35</sup> Tamtiež, s. 54.

<sup>36</sup> ANDRES, Jan. BENEŠOVÁ, Martina et al. **Methodological Note on the Fractal Analysis of Texts.** *Journal of Quantitative Linguistics*, [online]. 2012, roč. 19, č. 1, s. 1-31. [cit. 2012-03-21]. Dostupné z: <<http://dx.doi.org/10.1080/09296174.2011.608604>>. s. 2.

<sup>37</sup> Tamtiež, s. 2.

tvrdil, že koeficient  $b$  je zodpovedný za vykazovanie skracujúcej sa tendencie konštruktového konštituentu, za predpokladu, že konštrukt sa predlžuje. Hodnoty oboch koeficientov sú závislé na jazykovej úrovni.<sup>38</sup> Kulacka vo svojej práci porovnáva výsledky a poznatky dvojice Hammerl a Sambor s prácou Cramerovej, ktorá si podobne ako dvaja vyššie zmienení vedci všimla negatívnu koreláciu medzi koeficientmi. Dá sa predpokladať, že hodnota jedného koeficientu určuje hodnotu toho druhého. Cramerová taktiež upozornila na to, že v rámci rovnakej jazykovej úrovne sú hodnoty koeficientov veľmi podobné, na základe čoho stanovila predpoklad, že v budúcnosti môže byť objavená celá škála hodnôt pre jednotlivé jazykové úrovne.<sup>39</sup>

Ako píše Andres a kol. v článku *Optimization of parameters in the Menzerath-Altmann law* teoreticky by sa mohlo očakávať, že výsledky dosiahnuté pomocou kompletnej algebraickej formulácie budú presnejšie a optimálnejšie ako tie, ktoré boli dosiahnuté zjednodušenou algebraickou formuláciou. Avšak na základe svojho experimentu sám tvrdí, že situácia je nie je taká zrejmalá a jednoduchá a predošlý výrok je značne chúlостivý.<sup>40</sup>

Jazyk je štruktúrovaný celok, ktorý sa skladá z množstva na sebe naviazaných úrovní. Hřebíček pojednáva o tom, že v lingvistike sa prejavuje tendencia skúmať a presadzovať len tie jazykové úrovne, ktoré už boli rozpoznané. Avšak nie je správne tvrdiť, že počet jazykových úrovní je konečný na základe tradičnej filológie. Samozrejme je ale nevyhnutné, najskôr preskúmať známe jazykové úrovne a na základe poznatkov sa dopátrať a objaviť tie ďalšie.<sup>41</sup> Empirické overenie platnosti pozostáva z testovania tejto hypotézy na množstve jazykov. Tieto koncepty nazývané konštrukt a konštituent môžu byť interpretované rôznymi spôsobmi a môžeme ich nahradiť s empirickejšími konceptmi ako sú hláska, slabika, morfém, slovo, skupina slov, slovné spojenie, fráza, klauza, veta atď.<sup>42</sup>

---

<sup>38</sup> CRAMER, Irene. **The Parameters of the Altmann-Menzerath Law.** *Journal of Quantitative Linguistics*, [online] 2005, roč. 12, č. 1. s. 41-52. [cit. 2012-03-28]. Dostupné z: <<http://www.tandfonline.com/doi/pdf/10.1080/09296170500055301>>.s. 45.

<sup>39</sup> KUŁACKA, s. 260.

<sup>40</sup> ANDRES, Jan et al. **Optimization of parameters in the Menzerath-Altmann law.** *Mathematica*, 2012, roč. 51, č. 1, s. 5-27. s. 7.

<sup>41</sup> HŘEBÍČEK, *Vyprávění*, s. 59.

<sup>42</sup> ALTMANN, s. 4.

Keďže text, tak ako aj jazyk je skladaný z množstva úrovní, nie je možné izolovať len dve úrovne nezávisle na ostatných a podrobiť ich skúmaniu v izolovanom prostredí. Tento postup nie je možný preto, lebo každá z úrovní (s výnimkou začiatocnej a konečnej úrovne) vstupuje zároveň ako nadradená nižšej ale podriadená vyššej úrovni. To v lingvistických termínoch znamená že: „každá jazyková entita vŕči vŕsem vyŕším jazykovým úrovním je konstituentem a vŕči vŕsem nižŕším úrovním je konstruktem.“<sup>43</sup> Z tohto teda vyplýva nasledujúci reŕazec, ktorý je v texte realizovaný:

„... ↔ KONSTRUKT / KONSTITUENT ↔ KONSTRUKT / KONSTITUENT ↔ KONSTRUKT / KONSTITUENT ↔ KONSTRUKT / KONSTITUENT ↔ ...“<sup>44</sup>

Neŕ sa odvodila a overila táto postupnosť, existovala definícia, ktorá hovorila, ŕe dĺŕka komponentu *y* zobrazuje monotonicky klesajúcu funkciu, ktorá zobrazuje dĺŕku konŕstrukt *x*. Tu vŕsak nastáva otázka v tom, ŕo a ako je treba považovať za konŕstituent. Na jednej strane je za konŕstituent považovaná nižŕšia jednotka, napríklad klauza ako konŕstituent vety, slabika ako najbliŕŕŕi (fonetický) konŕstituent slova atď. Na druhej strane sa za konŕstituenty považovali aj ostatné jednotky napr. slovo bolo konŕstituentom vety, ale foném ako konŕstituent slova atď.<sup>45</sup> Z tejto dilemy nakoniec vznikol vyŕŕšie zmienený reŕazec, v ktorom každá jednotka, s výnimkou tých koncových, funguje ako konŕstrukt najbliŕŕŕej nižŕšej jazykovej jednotky a zároveň je konŕstituentom najbliŕŕŕej vyŕŕšej jazykovej jednotky.

Ako stanovuje Altmann vo svojej práci, úlohou výskumu by mali byť ŕtyri nasledujúce základné povinnosti. Prvou úlohou je testovanie hypotéz na základe ŕo možno najväčŕšieho počtu dát, ďalej by to malo byť skúmanie rozsahu Menzerathovho zákona na základe zostavovania a overovania stále nových a nových hypotéz. Ďalŕou z úloh je taktieŕ ŕpecifikovať krivky na prísluŕných údajoch a koeficienty dávať do súvislostí s ďalŕšími fenoménmi vo vzŕahu k danému jazyku a v neposlednom rade integrovať Menzerathov zákon do systému jazykových zákonov, alebo rozvíjať princípy, ktorých sa drŕí.<sup>46</sup>

---

<sup>43</sup> HŘEBÍČEK, *Vypravění*, s. 59.

<sup>44</sup> Tamtieŕ, s. 59.

<sup>45</sup> ALTMANN, s. 124.

<sup>46</sup> Tamtieŕ, s. 129.



Okrem nástroja štatistickej lingvistiky, ktorou je Menzerath-Altmannov zákon je nutné sa venovať aj ďalšej metodologickej problematike, do ktorej Těšitelová zahrňa definovanie jednotiek skúmaného súboru a taktiež poukazuje na dôraz výberu a rozsahu súboru.<sup>47</sup> Vymedzenie jednotiek je veľmi problematické, nakoľko rôzni jazykovedci k tejto úlohou pristupujú rôzne a tým pádom sa definície a vymedzenie jednotiek navzájom od seba líšia. Čo sa týka výberu súboru, Těšitelová poukazuje na závislosť frekventovanosti jednotky a dĺžky súboru, kde všeobecne platí, že čím je jednotka v súbore frekventovanejšia, tým menší výberový súbor je nutný k jej preskúmaniu a naopak.<sup>48</sup> V tejto kapitole sme sa zaoberali prevažne predstavením a zaradením kvantitatívnej lingvistiky v rámci všeobecnej jazykovedy a zoznámením sa so základnými črtami Manzerath-Altmannovho zákona. V nasledujúcich kapitolách sa vak budeme venovať metodologicky dôležitému vymedzeniu skúmaných jednotiek a taktiež si priblížime čínštinu ako jazyk, hlavne jej fonetický a rytmický systém, ktoré sú predmetom daného výskumu.

---

<sup>47</sup> TĚŠITELOVÁ, s. 116.

<sup>48</sup> Tamtiež, s. 117.

### 3. Segmentácia hovoreného textu

Slovo text primárne evokuje písomnú formu odovzdania rôznych typov informácií. Prostriedkom na zaznamenávanie a tým pádom na odovzdávanie informácií je však jazyk, ktorý je hierarchicky usporiadaný systém. Všetky jeho zložky majú úlohu na dorozumievaní, a to je zároveň jeho hlavnou úlohou. Daneš, ktorý viac menej prijíma Horáلكove ponímanie jazykového systému, definuje jazykový systém ako prostriedok k vytváraniu prehovorov, jeho jazykové zložky sa nachádzajú na rôznych stupňoch a plnia túto funkciu len nepriamo.<sup>49</sup> Jazyk teda vytvára pomocou svojich prostriedkov štruktúru, na základe ktorej sme schopný vytvárať písomné texty a jazykové prejavy.

Avšak kvantitatívna lingvistická analýza hovorenej reči môže narážať na úskalia, nakoľko hovorená reč plynie v čase a nie je možné ju analyzovať okamžite. Preto je nutné vytvoriť záznam reči. Müllerová vo svojej práci odporúča vtedajšiu metódu analýzy textu, získaného na základe prepisu magnetofónovej nahrávky.<sup>50</sup> V dnešných podmienkach však lingvisti už nie sú odkázaný len na nie úplne kvalitné nahrávky z magnetofónu, ale majú možnosť získať vysoko kvalitné zvukové alebo obrazové materiály pre ďalšiu analýzu, ktorých prepis je tým pádom omnoho jednoduchší a presnejší. Podľa definície Jaklovej sú hovorené texty menej prehľadné ako písané nakoľko „jednotky psaného textu bývajú navzájem zreteľne oddelené, tempo vnímaní textu není vázáno na tempo jeho vytvárení...“<sup>51</sup> Aj keď Jaklová súhlasí s Müllerovou ohľadom možnosti vytvorenia zvukového záznamu, stále však tvrdí, že hovorený prejav nedosahuje prehľadnosti písaného prejavu a rovnako nevykazuje možnosť prispôsobenia vnímania textu príjemcovým potrebám.<sup>52</sup>

Pri analýze hovoreného textu sa však nemôžeme spoliehať na syntaktické jednotky, ktoré sa používajú pri analýze písomného textu. Pri procese tvorby hovoreného textu dochádza k problému, že rečník sa sústreďuje skôr na obsahové oznámenie, tým pádom

---

<sup>49</sup> DANEŠ, František. *Jazyk a text I : výbor z lingvistického díla Františka Daneše. Část 1*. Praha : Univerzita Karlova, 1999. ISBN 8085899701. s. 58.

<sup>50</sup> MÜLLEROVÁ, Olga. *Mluvený text a jeho syntaktická výstavba*. Praha: Academia, 1994. ISBN 80-200-0489-0. s. 23.

<sup>51</sup> JAKLOVÁ, Alena. *Úvod do textové lingvistiky*. České Budějovice: Pedagogická fakulta, 1991. ISBN 80-7040-035-8. s. 104.

<sup>52</sup> Tamtéž, s. 104.

ako píše Müllerová: „nemá čas na jeho formální a logické uspořádání,...často není jasné, kde by vlastně byla hranice mezi větami.“<sup>53</sup> Veta v ponímaní Müllerovej je definovaná ako základná syntaktická jednotka písomného textu. Predtým než sa ale pustíme do definovania jednotlivých jazykových úrovní, ktoré boli použité pri ďalej popisovanom výskume, pozrime sa najprv na základné vlastnosti a rozdiely v hovorených textoch.

### **3.1 Hovorený text**

Hovorené prejavy, na základe ktorých je možné vytvoriť textový prepis, môžeme rozdeliť do niekoľkých skupín na dopredu pripravené/nepripravené (spontánne), čítané/nečítané, monologické/dialogické/špeciálny typ, ktorým je dialogizovaný monológ a posledným rozdielom je stupeň verejnosti.<sup>54</sup> V mnohých prípadoch tieto typy nie sú osamotené, ale dajú sa rôzne miešať a kombinovať. Pre kvantitatívnu analýzu prvej fázy tohto výskumu bolo nevyhnutné vybrať výberový súbor prevažne monologický, nečítaný, podľa možnosti čo najprirodzenejší, a nepripravený t.j. text ktorý „se realizuje jako bezprostřední reakce na podnět.“<sup>55</sup> Ako Jaklová ďalej uvádza, nepripravený text sa od toho pripraveného výrazne líši. Medzi základné znaky patrí hlavne jednoduchšia skladba a stavba vety, je v nej určitá miera uvoľnenosti a objavujú sa v nej časté korekcie a stakatovosť.<sup>56</sup>

Texty produkované za rozdielnych podmienok a patriace do rozdielnych vyššie spomínaných kategórií majú medzi sebou nesporné rozdiely. Tieto rozdiely v Danešovom ponímaní spôsobujú rozdielnu voľbu jazykových prostriedkov a prejavujú sa aj v spôsobe reči. Nastávajú tu rozdiely v gramatickej a lexikálnej výstavbe textu, ďalej v zvukovej realizácii a pod. Spolu so štylistikou môžeme taktiež rozoznať rozdiely v jazykových prostriedkoch, ktoré sú spisovné, ale podľa miery formality/neformality sa do rečového prejavu môžu vmiesiť aj prostriedky slangové, nárečové, profesionalizmy a iné.<sup>57</sup> S Danešom súhlasí aj Müllerová, podľa ktorej je „syntaktická strukturace mluvených projevů ovlivňována jednak činiteli interakčními, jednak různými a navíc

---

<sup>53</sup> MÜLLEROVÁ, s. 23.

<sup>54</sup> DANEŠ, František. *Jazyk a text I : výbor z lingvistického díla Františka Daneše. Část 2.* Praha : Univerzita Karlova, 1999. s. 249-471. ISBN 8085899701. s. 268.

<sup>55</sup> JAKLOVÁ, s. 105.

<sup>56</sup> Tamtéž, 105.

<sup>57</sup> DANEŠ. *Jazyk a text I : výbor z lingvistického díla Františka Daneše. Část 2.* s. 268.

proměňujícími se dispozicemi mluvčích v různých komunikačních situacích.“<sup>58</sup> Rovnako pozoruje rozdiel medzi dramatickým textom, ktorý je síce hovorený, ale na základe písomného textu, a spontánnym dialógom. Ten sa v niektorých rysoch výstavby od toho dramatického líši hlavne v tých, ktoré podliehajú podmienkam produkcie textu v prirodzenej komunikácii. Medzi tieto rysy patria napríklad „některé druhy oprav, nefunkční opakování, tzv. syntaktické nepravidelnosti atd.“<sup>59</sup> Tieto prejavy sa vyskytujú aj v skúmaných výberových súboroch A a B, čo potvrdzuje, že ich rečový prejav nebol vopred pripravený a tým pádom dochádza ako k opravám, tak aj k hezitačným zvukom na rozdiel od kontrastného výberového súboru C, ktorý bol pre porovnanie vybraný v druhej fáze výskumu. Pri oboch súboroch A a B taktiež dochádza k neukončeným výpovediam, keďže v procese tvorenia myšlienok a uceleného rečového prejavu často dochádza buď k odmlkám, kedy si človek premýšľa ďalšiu myšlienku a jej nadviazanie na tú predošlú, alebo dochádza k prehodeniu začiatku prejavu, človek začne novú myšlienku a tá predošlá ostane neukončená. Na rozdiel od písomného prejavu, v ktorom sú zachované už vopred pripravené, premyslené a jasne formulované myšlienky, v spontánnom nepripravenom rečovom prejave k takémuto javu dochádza pomerne často. Samozrejme autor si je toho vedomý a vychádza zo znalosti gramatiky a syntaktickej skladby vety, avšak ako píše Müllerová, hlavným cieľom autora je pomocou týchto pravidiel primárne znázorniť a vyjadriť obsah svojich myšlienok a tým pádom veta ako syntaktická štruktúra nie je pre neho taká dôležitá.<sup>60</sup> K tejto tematike sa vrátim pri popise a interpretácii výsledkov výskumu, v ktorom sa prejavili takmer všetky vyššie spomínané úskalia analýzy hovoreného textu ako sú hezitácia a hezitálne zvuky, spontánnosť, opravy, nedokončené výpovede a pod. Momentálne je však nutné zaoberať sa rôznymi prístupmi charakteristiky textových jednotiek a ich definovanie v hovorenom texte.

### **3.2 Textové jednotky**

Pre tento výskum bolo stanovených päť jazykových jednotiek, a to menovite hláska – slabika – prízvukový takt – výpoveď – replika. Vzhľadom na to, že táto práca je zameraná hlavne na úroveň výpovede, ostatným úrovniám sa budem venovať len okrajovo.

---

<sup>58</sup> MÜLLEROVÁ, s. 24.

<sup>59</sup> Tamtéž, s. 28.

<sup>60</sup> Tamtéž, s. 23.

### 3.2.1 Hláska

Hláska ako jazyková kategória je podľa starších definícií Hálu charakterizovaná ako najmenšia ďalej nedeliteľná zvuková jednotka, ktorú je možné vysloviť artikulačnými orgánmi a predstavuje základný zvukový prvok konkrétneho jazyka. Krčmová však s týmto tvrdením nesúhlasí a za minimálnu artikulačnú jednotku považuje až slabiku. Hláska môže byť samostatnou slabikou, za predpokladu, že je tvorená samohláskou a má funkciu spojky. V slovenčine a češtine sú to spojky napr. a, i. Krčmová ďalej tvrdí, že samostatné spoluhlásky ako také nemôžu byť slabikou, pretože ak vyslovíme akúkoľvek spoluhlásku napr. pri hláskovaní určitého slova, nutne a prirodzene túto spoluhlásku doplnujeme o neurčitú samohlásku.<sup>61</sup> Tým pádom vlastne z hlásky pridaním ďalšej hlásky vytvoríme slabiku. Preto nie je možné, aby hlásky mohli stáť samostatne, s výnimkou vyššie zmieňovaných samohlások.

### 3.2.2 Slabika

Ako je všeobecne známe, slabiky sú tvorené nižšími jednotkami nazývanými hlásky. Určité zoskupenie hlások vytvára slabiku, avšak ako je to zmienené vyššie, aj určité samostatné hlásky dokážu vytvoriť slabiku. Krčmová definuje slabiku, inak nazývanú ako sylaba, ako najtesnejšiu a najjednoduchšiu artikulačnú jednotku, ktorá vyhovuje dorozumievaniu.<sup>62</sup> Rovnako ako Krčmová, aj Daneš považuje slabiku za základnú jednotku hovorenej reči a definuje ju ako: „koordinovaný (souhlasně a jednotně provedený) popud k takové činnosti mluvidel, jíž se uvolní cesta výdechovému proudu za současného silového crescenda.“<sup>63</sup> Ako Krčmová, tak aj Daneš sa zhodujú v tom, že slabika je pomerne inštinktívny jav, uvedomujú si ho aj deti od raného detstva a ktorý sa v nich umocňuje pomocou riekaniak a rytmických básničiek. Je samozrejmé, že všetky slabiky nevyslovujeme rovnako silným hlasom. V češtine a slovenčine je slovný prízvuk prevažne na prvej slabike, ale v iných jazykoch tento prízvuk môže variovať. Preto Daneš rozlišuje dva typy slabík, ktoré delí na prízvučné, to sú tie ktoré nesú prízvuk a na neprízvuchné.<sup>64</sup> V čínštine je distribúcia prízvukov podriadená pomerne výrazne prepracovanej hierarchii na siedmich stupňoch. Čínskej slabike sa budem

---

<sup>61</sup> KRČMOVÁ, Marie: *Fonetika* [online] Brno: Filozofická fakulta MU, 2007. [cit. 2013-04-10] Dostupné z: <is.muni.cz/elportal/estud/ff/js07/fonetika/materialy/ch04.html>.

<sup>62</sup> Tamtéž.

<sup>63</sup> DANEŠ, František a spol. *O mluveném slově*. Praha: Státní pedagogické nakladatelství, 1954. s. 37.

<sup>64</sup> DANEŠ, *O mluveném*, s. 87.

konkrétne venovať po všeobecnom predstavení jazykových jednotiek používaných v tomto výskume.

### 3.2.3 Prízvukový takt

Pri analýze hovoreného textu dochádza k problémom hlavne v rámci určenia základnej jednotky. V niektorých prípadoch sa názory jazykovedcov zhodujú, v iných sa čiastočne prekrývajú, prípadne úplne líšia. Možností ako stanovovať jednotky hovoreného textu je niekoľko. Podľa Müllerovej je možné vychádzať zo zvukovej realizácie textu, ďalšou možnosťou je zameranie na syntaktické vzťahy medzi jednotkami lexikálnymi a tretím spôsobom je dôraz na obsahovo-pragmatické usporiadanie.<sup>65</sup> Alexová nadväzuje na Müllerovú a porovnáva jej koncept základnej jednotky hovoreného textu s definíciou od Hrbáčka.<sup>66</sup> Müllerová považuje za najspoľahlivejší spôsob členenia hovoreného textu pomocou zvukovej segmentácie, ktorá člení text na zvukové jednotky. Názvy a definície týchto jednotiek sa líšia u väčšiny jazykovedcov. V tomto prípade však Müllerová prirovnáva tónové jednotky, ktoré používa ona, k „promluvovým taktům“ u Daneša a k „mluvným taktům“ u Mathesia a charakterizuje ich ako úseky textu oddelené pauzou alebo intonačným pohybom.<sup>67</sup> V tomto výskume je na danej úrovni používaný termín prízvukový takt. Krčmová definuje takt ako skupinu slabík, ktoré patria k jednému slovnému prízvuku a túto jednotku priraduje k slovu, aj keď zároveň tvrdí, že prízvukový takt môže byť aj vyššou jednotkou ako je slovo.<sup>68</sup> Rovnaký názor má aj Daneš, ktorý tvrdí, že slovo je síce gramatickým základom reči, avšak hovorená reč sa deje prevažne v taktoch, kde k prízvučnej slabike priradzujeme niekoľko neprízvučných a vytvára sa tak jeden zvukový celok.<sup>69</sup> Navyše v čínštine, ktorá je predmetom skúmania tohto výskumu je kategória slova veľmi zložitá a nie je úplne definovaná. Tento výskum sa zaoberá prevažne fonetickou stránkou reči, preto je preň slovo nevyhovujúca jednotka a namiesto neho bola zvolená fonetická jednotka prízvukový takt, ktorý ako píše Krčmová obsahuje slabiku charakterizovanú prízvukom a k nej sú priradené slabiky bez prízvuku.<sup>70</sup> Nakoľko čínština je tónová reč a má jasnú

---

<sup>65</sup> MÜLLEROVÁ, s. 40.

<sup>66</sup> *Jazykověda = Linguistica*. Ostrava: Ostravská univerzita, 2004. Sborník prací Filozofické fakulty Ostravské univerzity. ISBN 80-7042-664-0. s. 239.

<sup>67</sup> MÜLLEROVÁ, s. 41.

<sup>68</sup> KRČMOVÁ.

<sup>69</sup> DANEŠ, *O mluveném*, s. 37.

<sup>70</sup> KRČMOVÁ.

hierarchiu prominencie (prízvučnosti) jednotlivých slabík, táto jednotka bola zvolené ako vyhovujúca pre účely výskumu.

### 3.2.4 Výpoveď

Následná vyššia zvuková jednotka sa všeobecne označuje ako výpoveď. Krčmová charakterizuje výpoveď nasledovne:

„výpoveď je obvykle menší než promluva, může však být i minimální promluvou. Z fonetického hlediska je vyznačena větnou intonací a relativními pauzami, mnohdy jen potenciálními...Je zvukově členěna větným přízvukem a intonací a je tedy nositelkou intonačních prostředků modulace řeči.“<sup>71</sup>

Táto definícia umožňuje používať názov výpoveď ako jednotku hovoreného textu, aj keď Alexová považuje výpoveď skôr za jednotku „písomného komunikátu“ a prirovnáva ju k Müllerovej „obsahově-pragmatickej“ jednotke.<sup>72</sup> Alexová sa opiera o Hrbáčkovú definíciu výpovedi, v ktorej je táto je charakterizovaná ako základná jazyková jednotka, ktorá v sebe nesie určitý obsah a má výpovednú funkciu.<sup>73</sup> Zatiaľ čo Krčmovej definícia sa zameriava hlavne na fonetickú stránku výpovede, ostatné definície sa zameriavajú prevažne na jej syntaktickú a sémantickú skladbu.

Vo vyššie zmienených prípadoch sa výpoveď viac-menej považuje za základnú jednotku hovoreného textu s rozdielmi v náhľade na ňu, avšak v literatúre narazíme aj na výrazné rozdiely v ponímaní výpovede. Jeden z prúdov považuje výpoveď za ekvivalent vety, ktorého predstaviteľmi sú napr. Daneš, či Mathesius. Daneš rozlišuje vetu a výpoveď tým, že považuje termín veta za definíciu „veta ako jednotku jazykového systému“ a výpoveď označuje ako „veta ako súčasť jazykového systému“.<sup>74</sup> Mathesius zas považuje výpoveď ako vetu, ktorá v sebe nesie vyjadrenie aktuálneho

---

<sup>71</sup> HRBÁČEK, Josef. **Promluva. Naše řeč**, [online] 2003, roč. 86, č. 1. [cit. 2012-03-28]. Dostupné z: <<http://nase-rec.ujc.cas.cz/archiv.php?art=7705>>.

<sup>72</sup> *Jazykověda*, s. 239.

<sup>73</sup> Tamtéž, s. 239.

<sup>74</sup> Daneš, *Věta a text*, 8.

postoja hovorcú.<sup>75</sup> Daneš samozrejme dodáva, že v hovorenom texte sú časté nevetné výpovede, keďže v prípade hovoreného textu je reč spontánna a dochádza k rôznym vplyvom, ktoré nemajú gramatickú štruktúru vety, napriek tomu ale výpoveďou sú, keďže nesú určitú výpovednú hodnotu a sémantický obsah.<sup>76</sup> Daneš charakterizuje jazyk ako hierarchický systém, ktorý je zložený zo systematicky usporiadaných prostriedkov na rôznych stupňoch. Podľa jeho názoru najnižšie jednotky plnia funkciu jazyka len sprostredkovane, pretože funkciou jazyka je dorozumievanie a tá je splnená až na úrovni vetných jednotiek, kedy sa veta stáva skutočnou výpoveďou a tým pádom spĺňa funkciu jazyka.<sup>77</sup> Tu je zrejme, že Daneš ako ekvivalent vety používa výpoveď alebo „promluvu“. Rovnako aj Jaklová vo svojej práci dáva medzi výpoveď a vetu znamienko rovná sa, nakoľko pri komunikácii výpovede za sebou nasledujú v určitej následnosti a ona tieto výpovede taktiež nazýva ako „vety za sebou nasledujúci v určenom poradí.“<sup>78</sup>

Na druhú stranu Barthes vo svojom článku o štrukturálnej analýze príbehu poznamenáva, že lingvistika sa často sústreďuje na vetu ako na poslednú jednotku, ktorá môže byť skúmaná, pretože všetky vyššie jednotky sú z nej vystavané. Tým pádom je veta považovaná za pôvodnú jednotku, a výpoveď je charakterizovaná ako vyššia jednotka, ktorá sa skladá z viet, teda z niekoľkých nižších jednotiek.<sup>79</sup> Barthes však pracuje skôr s textom písaným a ako je vyššie zmienené, pri hovorenom texte sa významy vety a výpovede často prelínajú. Väčšinou sa však veta považuje za jednotku textu písaného a výpoveď je jej ekvivalent v texte hovorenom.

Aj keď sa výpoveď javí ako jednotka, ktorú je možné v texte rozlíšiť pomocou fonetických pravidiel popisovaných Krčmovou, Daneš poukazuje na nepopierateľný vzťah medzi zvukovou (príp. intonačnou) realizáciou a obsahom resp. významom výpovede. Podľa neho je možné vo výpovedi rozlišovať dôraznú a nedôraznú časť, s čím súvisí aj intonačná kadencia. Na rozdiel od dôraznej (prízvukovej) časti však

---

<sup>75</sup> DANEŠ, *Věta*, s. 8.

<sup>76</sup> Tamtéž, s. 8.

<sup>77</sup> DANEŠ, *O mluveném*, s. 90.

<sup>78</sup> JAKLOVÁ, s. 95.

<sup>79</sup> BARTHES, Roland. *Úvod do strukturální analýzy vyprávění. Znak, struktura, vyprávění*. Brno: Host, 2002, s. 9-43. s. 12.



nedôrazná môže byť vynechaná.<sup>80</sup> Tejto problematike sa venuje aj Müllerová, ktorá zdôrazňuje a poukazuje na rozdiely medzi písaným a hovoreným textom. V tom hovorenom je omnoho častejší výskyt neukončených výpovedí a pri zvukovej segmentácii nie je možné spoľahlivo sa riadiť intonačnou signalizáciou pri rozhodovaní, či ide o koncový alebo nekoncový úsek prejavu.<sup>81</sup> V rámci reči potom dochádza k tomu, že rečník pokračuje v nadväzovaní myšlienky, alebo ju mení v rámci svojho prejavu bez toho, aby túto zmenu jasne intonačne naznačil a tým dochádza k problému segmentácie daného prejavu. V tomto prípade už nie je možné spoliehať sa len na fonetickú stránku, ale je nutné do segmentácie včleniť stránku syntaktickú a sémantickú, aby sme od seba dokázali oddeliť dve nezávisle výpovede, ktoré boli foneticky spojené. Müllerová túto situáciu popisuje tak, že: „i poměrně dlouhé úseky textu, zejména monologického, se někdy člení jen na výpovědní úseky s nekoncovou intonační kadencí, přestože jejich významová a syntaktická výstavba poskytuje možnost uplatnění koncové intonace a tím rozčlenění textu do jednotlivých výpovědních celků.“<sup>82</sup> Tento postup bol uplatňovaný aj pri segmentácii textov potrebných na výskum popisovaný v príslušnej časti tejto práce, v rámci ktorého sme sa rozhodli aplikovať definíciu výpovedí na základe fonetického hľadiska, ktoré bolo síce prvotné, ale s výrazným doplnením o hľadisko syntaktické a sémantické. Týmto spôsobom sme dlhšie časti textu rozdelili na menšie jednotky, ktoré sme nazývali výpoveďou s prihliadnutím na obsah danej výpovede.

### 3.2.5 Replika

Posledná jazyková úroveň ktorou sa vo výskume zaoberáme sa nazýva replika. Táto úroveň je spojená hlavne ucelenou významovosťou, kde jednotlivé výpovede na seba nadväzujú a sú súčasťou vyššieho významového celku. Replika, v českej literatúre inak nazývaná aj promluva je charakterizovaná ako rečový celok, ktorý je oddelený dvoma absolútnymi pauzami, zjednocuje ju obsahová súdržnosť a jediný rečník.<sup>83</sup> Termín promluva ako prvý použil a definoval v tridsiatych rokoch 20. storočia V. Skalička.<sup>84</sup> Hrbáček vo svojom zborníkovom príspevku uvádza stručný prehľad rôznych definícií repliky/promluvy, avšak tvrdí, že ani jedna z nich nie je úplná, každá z nich sa totiž

---

<sup>80</sup> DANEŠ, František. **Příspěvek k rozboru významové výstavby výpovědi.** *Studie a práce lingvistické.* I. 1. vyd. Praha : Nakladatelství Československé akademie věd, 1954. s. 67.

<sup>81</sup> MÜLLEROVÁ, s. 41.

<sup>82</sup> Tamtéž, s. 41-42.

<sup>83</sup> KRČMOVÁ.

<sup>84</sup> HRBÁČEK.

zameriava na časť celkovej definície. Sám preto zhŕňa tieto poznatky a definuje repliku ako jednotkou reči, ktorá je tvorená jednotou autorského subjektu, obsahom a dominantnou funkciou.<sup>85</sup> V rámci popisovaného výskumu sme za repliky označili jednotlivé vstupy medzi prerušovanými videoklipmi, nakoľko v každom vstupe sa interpret venoval určitému tematicky ucelenému prejavu. Vzhľadom na to, že výpoveď ako taká nepresahuje útvary vety, výpovede sémantický spojené do jedného celku spoločne vytvárajú prejav, ktorý je možné nazvať ako promluva alebo v našom prípade replika. Charakteristickým znakom každej repliky, je tematická nadväznosť jednej výpovede na druhú, tieto sa vzájomne dopĺňujú a tvoria ucelený celok. V tomto prípade bolo toto rozčlenenie uľahčené už spomínanými predelmi pomocou videoklipov, ktoré tým pádom rozčlenili prejav daného interpreta do niekoľkých obsahovo ucelených častí. V ďalšej časti sa bližšie pozrieme na charakteristiku jednotlivých jazykových úrovní v rámci čínštiny, nakoľko je nutné vyššie spomínané všeobecné jazykové poznatky aplikovať na skúmaný jazyk, samozrejme s prihliadnutím na charakteristické odlišnosti daného jazyka. Taktiež je nutné porovnať a pojednať o odlišnostiach názvoslovia a ustáliť konkrétne definície jednotlivých úrovní, ktoré sú pre daný experiment nevyhnutné.

### ***3.3 Jazykové úrovne v rámci čínštiny***

Aj keď s vyššie zmienenými definíciami pracujú lingvisti hlavne v súvislosti s európskymi jazykmi, myslím, že tieto poznatky sa dajú aplikovať aj na ázijské jazyky, konkrétne na čínštinu, ktorá bola predmetom tohto skúmania. V čínštine, tak ako v každom inom jazyku, je možné rozdeliť tento jazyk na menšie jednotky. Problémom je ale nejednoznačnosť a relatívna variabilita týchto jednotiek, nakoľko koncept slova je v čínštine veľmi diskutabilný a preto je stanovovanie jednotiek pomerne zložitá. Jedným z jazykovedcov, ktorý sa v hlbokej miere zaoberal čínštinou, jej fonetikou a štruktúrou bol aj profesor Švarný. Pri svojich experimentoch a skúmaniach čínštiny rozdeľoval vety a súvetia na menšie jednotky, tzv. kóla (v jednotnom čísle nazývaný kólón), ďalej to boli segmenty, bináry, slabiky a nakoniec ako najmenšiu jednotku používal hlásky.<sup>86</sup> Švarný charakterizuje kolón ako jednotku, ktorá je zložená z niekoľkých segmentov, má ucelenú významovú a rytmickú štruktúru a charakterizuje

---

<sup>85</sup> HRBÁČEK.

<sup>86</sup> ŠVARNÝ, Oldřich a kol. *Gramatika hovorovej čínštiny v príkladoch*. Bratislava: Vydavateľstvo Univerzity Komenského, 1993. s. 23-25.

ich taktiež aj neukončujúca modifikácia poslednej tonickej slabiky. Tempo reči podľa Švarného nemá vplyv na dĺžku kolónu, avšak v rámci dlhších viet bývajú kóla dlhšie.<sup>87</sup>

### 3.3.1 Segment

Nižšou jazykovou jednotkou, používanou v Švarného terminológii je segment, ktorý vytvára vyššie zmienenu jednotku – kólon. Segmenty pozostávajú zo skupiny slabík, z ktorých minimálne jedna musí obsahovať rytmický prízvuk (iktus). Rytmus reči je závislý práve na spájaní a vyslovovaní segmentov, a tie sú ako píše Švarný variabilné a ich dĺžka závisí na tempe reči.<sup>88</sup> Nakoľko čínština je tónový jazyk, význam je rozlišovaný tónom, táto tendencia sa nevyhnutne musí preukázať aj v reči ako takej. Avšak na rozdiel od samostatne vyslovenej slabiky, vo vete resp. pri reči nie je možné realizovať tóny na všetkých slabikách. V reči preto Švarný rozlišuje niekoľko typov slabík: slabiky so zvýraznenou prominenciou, ďalej sú to plne tonické, oslabene tonické a atonické slabiky.<sup>89</sup>

### 3.3.2 Slabika

Čínska slabika je považovaná za základnú fonologickú jednotku čínštiny a skladá sa z troch častí. Týmito časťami sú iniciála, finála a tón.<sup>90</sup> Čínsky jazyk vytvára slabiky kombinovaním 21 resp. 23 iniciál a 36 finál. Ako píše Trísková, čínština nerozlišuje medzi znelými a neznelými spoluhláskami, ale delí ich na aspirované a neaspirované. Finálu tvorí zvyšok slabiky a môže byť tvorená samostatnou samohláskou, prípadne kombináciou viacerých samohlások alebo kombináciou samohlásky a nazály.<sup>91</sup> Vo všeobecnej lingvistike je slabika alebo sylaba je považovaná za „nejjednoduchšiu a nejtěsnější možnou artikulační jednotou funkčních prvků řeči, která vyhovuje dorozumívání.“<sup>92</sup> Ako Krčmová stanovuje, slabiky delíme na otvorené, to sú také, ktoré majú na konci samohlásku, a zatvorené, to sú tie, ktoré končia spoluhláskou. V Čínštine prevládajú hlavne otvorené slabiky, jedinými zatvorenými slabikami sú tie, na ktorých konci sa nachádzajú nazály n a ng. Čínština toleruje na konci slabiky jedine tieto dve spoluhlásky, ostatné spoluhlásky môžu vystupovať len v pozícii iniciály.

---

<sup>87</sup> ŠVARNÝ, *Gramatika*, s. 24.

<sup>88</sup> Tamtiež, s. 24.

<sup>89</sup> Tamtiež, s. 25.

<sup>90</sup> TRÍSKOVÁ, Hana. *Transkripce čínštiny: Tabulky a návody*. Praha: Česko-čínská společnost, 1999. ISBN 80-902515-3-6. s. 8.

<sup>91</sup> Tamtiež, s. 9-11.

<sup>92</sup> KRČMOVÁ.

### 3.3.3 Rozdiely v názvosloví

Slabiky a hlásky sú úrovne, ktoré sa za všeobecne za sporné nepovažujú a sú viac menej jasne definované a charakterizované. Avšak na ďalších úrovniach sa názvoslovie môže líšiť. Jiangfen Cao z jazykového inštitútu vo svojej práci uvádza ako vyššie jazykové jednotky prozodické slovo (prosodic word) a prozodickú frázu (prosodic phrase). Rovnako ako Švarný, aj on uvádza, že slabiky v čínštine majú tendenciu sa zoskupovať do väčších dvoj-trojslabičných celkov. Tieto celky, ktoré Švarný nazýva segmentami, podľa Caoa sú tieto časti centrom prozodického slova a slúžia ako základný kameň výstavby vety. Prozodické slovo je tvorené prevažne dvoma až tromi slabikami a veľmi ojedinele môžu obsahovať monosylabické slovo.<sup>93</sup> V prípade, že sa ale zoskupí niekoľko takýchto dvoj-trojslabičných častí, vytvárajú prozodickú frázu.<sup>94</sup> Prozodická fráza sa môže označovať aj ako fonemická klauza alebo fonologická fráza a je pomerne náročné ju definovať, aj keď je považovaná za najčastejšiu a najdôležitejšiu rytmickú jednotku v čínskej reči.<sup>95</sup> Cao ďalej pokračuje, že na každej prozodickej úrovni sa v rámci rytmickej jednotky objavujú určité črty a hranice medzi jednotlivými jednotkami.

### 3.4 Rytmus reči

Hovorená čínština je výrazne ovplyvňovaná rytmom, ktorý sa podieľa na celkovej skladbe vety a prehovoru. Rytmické členenie viet je veľmi dôležité, nakoľko prízvuk a rytmus spolupracujú a tým pádom vytvárajú jednotlivé prehovory. V čínštine je rytmus veľmi variabilný a môže závisieť od niekoľkých faktorov. Ako píše Cao, to, ako sa prejavuje trvanie určitej rytmickej jednotky, výrazne závisí na kontexte a je obzvlášť citlivé na zmeny v para lingvistických informáciách, ako je napríklad zdôraznenie určitého aspektu, alebo rečníkom pridaný dôraz v danej vete.<sup>96</sup> Rytmus a tempo reči výrazne ovplyvňujú samostatnú realizáciu konkrétneho prejavu. Švarný taktiež tvrdí, že priemerná dĺžka segmentu je veľmi variabilná a je výrazne ovplyvňovaná tempom reči. Reč a jej rytmus vzniká uceleným vyslovením a spojením segmentov do kól a tie sa následne spájajú do viet. Dĺžka segmentu však môže oscilovať podľa rýchlosti, v akej

---

<sup>93</sup> CAO, Jianfen. **Rhythm of Spoken Chinese**. [cit. 2012-04-06]. Dostupné z: <[http://ling.cass.cn/yuyin/report/files/2000\\_4.pdf](http://ling.cass.cn/yuyin/report/files/2000_4.pdf)>. s. 9.

<sup>94</sup> CAO, Jianfen. **The Rhythm of Mandarin Chinese**. [cit. 2012-04-08]. Dostupné z: <[http://www.researchgate.net/publication/228722757\\_The\\_rhythm\\_of\\_Mandarin\\_Chinese?citationList=incoming](http://www.researchgate.net/publication/228722757_The_rhythm_of_Mandarin_Chinese?citationList=incoming)>. s. 24.

<sup>95</sup> CAO, **Rhythm**, s. 10.

<sup>96</sup> CAO, **The Rhythm**, s. 24.

daný rečník rozpráva. Namiesto termínu dôraz Švarný používa termín iktus, to znamená že iktová slabika je tá, ktorá je prominentnejšia a tón na nej je realizovaný, neiktová slabika je tá, u ktorej sa tón nerealizuje v jeho plnej podobe a je oslabený. Na základe pozorovaní a skúseností potom Švarný prišiel k záveru, že v pomalšej reči prevládajú segmenty s jednou iktovou slabikou (t.j. slabika, na ktorej sa nachádza rytmický dôraz). V rýchlejšej reči sa ale segmenty na základe určitých gramatických pravidiel bližšie spájajú a vytvárajú tak rytmický celok, v ktorom sa nachádza niekoľko iktových slabík.<sup>97</sup> Dĺžka segmentu tým pádom taktiež závisí na tempe reči a môžeme z toho vyvodit' závislosť priamej úmernosti, čo znamená, že čím je tempo reči rýchlejšie, tým dlhšie sú jednotlivé segmenty. Švarný odhaduje dĺžku segmentu na priemer 2,5 - 4,5 slabiky. Na rozdiel od neho ale Cao za najdôležitejšiu stavebnú jednotku, ktorá ovplyvňuje produkciu a percepciu reči, považuje prozodickú frázu. Jej dĺžka variuje rovnako ako dĺžka segmentu v závislosti od tempa reči, avšak tieto variácie sa zdajú byť obmedzená v priemere 7 slabík s toleranciou 2 slabiky.<sup>98</sup> Bez ohľadu na rozdiely Švarného a Caoa v názvosloví a v priemere dĺžky základnej rytmickej stavebnej jednotky v čínštine sa obaja zhodujú v tom, že rýchlosť reči rovnako ako jej štýl majú tú schopnosť a ovplyvňujú dĺžku jednotlivých jednotiek. Rytmus odráža dočasnú distribúciu lingvistických informácií a silu vzťahu medzi jednotlivými rečovými jednotkami. Vo všeobecnosti platí, že vzťah je tým tesnejší, čím viac tieto dve alebo viaceré jednotky spolu súvisia a patria do jednej a tej istej rytmickej skupiny. A naopak tento vzťah medzi jednotkami je tým voľnejší, čím viac sú jednotky oddelené do iných rytmických skupín.<sup>99</sup>

### **3.5 Definovanie jednotiek**

V rámci experimentu bola ako stredná a základná jazyková jednotka zvolený prízvukový takt, ktorý spĺňa definíciu Caovej prozodickej frázy. Aj keď Švarný pripodobňuje ním používaný segment prízvukovému taktu, obe jednotky sa stopercentne nezhodujú. Prízvukový takt sme preto určili ako jednotku, ktorá stojí medzi dvoma jednotkami prof. Švarného a to konkrétne medzi segmentom a kólonom. Prízvukový takt bol definovaný ako rytmická jednotka, ktorá je ovplyvnená tempom reči, je tvorená slabikami zoskupenými do menších celkov a obsahuje minimálne jeden

---

<sup>97</sup> ŠVARNÝ, *Gramatika*, s. 24.

<sup>98</sup> CAO, *The Rhythm*, s. 25.

<sup>99</sup> CAO, *Rhythm*, s. 10.

lexikálny prízvuk, pričom skupina slabík s minimálne jedným lexikálnym prízvukom je od ostatných slabík oddelená relatívnymi pauzami. Táto práca sa zameriava hlavne na vyššiu jazykovú jednotku t.j. výpoveď, ktorá je ale meraná v počte prízvukových taktov, ktoré sú merané v priemernej dĺžke jeho slabík. Výpoveď bola stanovená za základe vyššie zmienených definícií, ktoré sa opierajú o českú lingvistiku, avšak tieto pravidlá sa dajú aplikovať aj na čínsky jazyk. Napriek tomu, že experiment sa mal zameriavať na jazyk predovšetkým z fonetického hľadiska, túto úroveň nebolo možné vyčleniť z textu bez pomoci hľadiska syntaktického a sémantického, čo vyplýva z vyššie zmienených definícií. Preto v danom výskume chápeme výpoveď ako jazykovú úroveň, ktorá je tvorená jedným alebo viacerými prízvukovými taktami, je oddelená pauzou v reči, má ucelenú syntaktickú a štruktúru a jej súčasti navzájom spolu sémanticky súvisia.

Replika, v literatúre nazývaná aj ako prehovor, je charakterizovaná ako celok reči nachádzajúci sa medzi dvoma absolútnymi pauzami. Zjednocuje ju obsahová súdržnosť a jeden rečník. Dĺžka repliky je potenciálne neobmedzená.<sup>100</sup> Pri segmentácii výberového súboru A a B v prvej fáze výskumu absolútne pauzy vytvárali hudobné klipy vkladané medzi jednotlivé vstupy interpreta a tým pádom oddeľovali obsahovo ucelené repliky. Úroveň repliky bola meraná v dĺžke výpovedí, ktoré boli merané v priemernej dĺžke ich prízvukových taktov. Vzťah repliky a výpovede, rovnako ako vzťah prízvukového taktu a slabiky nebol skúmaný na výberovom súbore C, nakoľko v druhej fáze výskumu som sa zamerala na vzťah výpovede a prízvukového taktu. Tento vzťah som skúmala na výberových súboroch A a B z prvej fázy výskumu a na kontrastnom výberovom súbore C pridanom v druhej fáze výskumu.

---

<sup>100</sup> KRČMOVÁ.

## 4. Problematika vhodnej transkripcie

Po zvážení a vybratí výberových súborov experiment pokračoval prepisom hovorenej reči do textovej podoby a následný prepis do transkripcie, ktorá bola pre spracovanie dát najvýhodnejšia. Dostupné boli: oficiálna čínska transkripcia pinyin, anglická transkripcia Wade-Giles, česká transkripcia prof. Švarného a fonetická transkripcia IPA (The International Phonetic Alphabet). Vzhľadom na fakt, že predmetom skúmania bol počet hlások to znamená kvantita a nie kvalita jednotlivých hlások, fonetická abeceda IPA bola pre túto úlohu nevhodná. IPA je štandardizovaná medzinárodná fonetická abeceda, ktorá sa používa na zaznamenanie rečových zvukov. Jej konkrétnu podobu a prehľad jednotlivých symbolov je možné nájsť napr. v príručke Martina Dokoupila.<sup>101</sup> Táto transkripcia totiž zaznamenáva aj kvalitu hlások a používa na to špeciálne symboly, ktoré by v tomto prípade narušovali kontinuitu a efektivitu práce a spôsobovali pomerne výrazne problémy pri spracovávaní dát. Ďalšou možnosťou bola anglická transkripcia hovorenej čínštiny známa ako Wade-Gilesova transkripcia. Táto možnosť sa ale po bližšom preskúmaní taktiež ukázala ako neefektívna, nakoľko v rámci kvantifikácie dát bolo nutné čo najpresnejšie spočítať jednotlivé vyslovené hlásky a táto transkripcia by sa pri tejto úlohe neosvedčila. Rovnako ako fonetická abeceda IPA by proces prepisu a následnej kvantifikácie spomaľovala, nehovoriac o riziku častejších nepresností a chýb, nakoľko počet hlások v slabike podľa tejto transkripcie vo väčšine prípadov nezodpovedá skutočnému počtu realizovaných hlások. Táto vlastnosť by pri procese spracovávaní skôr narušovala proces a napomáhala väčšiemu počtu chýb.

Jednou z možností bolo taktiež prepísať hovorený text do čínskej fonetickej abecedy, ktorá sa nazýva Hanyu Pinyin Fang'an 汉语拼音方案<sup>102</sup>. Skrátené sa tejto abecede vraví pinyin. Táto abeceda bola vyvinutá ČĽanmi špeciálne na prepis čínskych znakov do latinky. Pinyin síce zachytáva výslovnosť daného znaku, avšak pre laika je veľmi náročné a bez osvojenia si základných pravidiel pinyinu aj nemožné byť schopný text napísaný v tejto abecede správne prečítať. Aj keď transkripcia používa latinku, v mnohých prípadoch sa hláska alebo kombinácia hlások inak píše a inak vyslovuje. Tým pádom je veľké riziko vzniku chýb pri kvantitatívnom spracovávaní dát.

---

<sup>101</sup> DOKOUPIL, Martin. *Jazykověda*. Vrbátky: M. Dokoupil, 2012. ISBN 978-80-260-2129-2. s. 135.

<sup>102</sup> TRÍSKOVÁ, s. 4.

Na vyriešení problému s výberom transkripcie sa vo veľkej miere podieľa profesor Švarný a jeho štandardná česká transkripcia. Už pred ním boli pokusy vypracovať systém prevodu výslovnosti čínštiny do češtiny a ako píše Třísková, prvý ucelený systém vypracoval v 30. rokoch 20. storočia profesor Jaroslav Průšek, ktorého prepis sa používal až do 50. rokoch 20. storočia. V tejto dobe profesor Oldřich Švarný vypracoval nový návrh českej transkripcie a použil ho v *Učebnici hovorové čínštiny* z roku 1951.<sup>103</sup> Tento prepis čínskych znakov je založený hlavne na ich výslovnosti a jeho cieľom je čo možno najbližšie priblížiť čitateľovi správnu čínsku výslovnosť daného znaku či slova. Problematickosť pinyinu spočíva v niekoľkých spoluhláskach, ktoré sa síce píšu jedným písmenom, ale môžu reprezentovať niekoľko hlások, alebo naopak, kde daná hláska je zapisovaná viacerými písmenami. Samozrejme nie je to výnimka len spoluhlások, v čínštine nájdeme aj samohlásky, ktoré podliehajú tomuto fenoménu. Preto sa prof. Švarný rozhodol zostaviť čo možno najpresnejšiu transkripciu a na základe podrobnej úvahy bola táto transkripcia vybraná ako vhodný nástroj na prepis hovoreného čínskeho textu na ďalšiu kvantitatívnu analýzu.

#### ***4.1 Fonetické problémy a porovnanie pinyinu a českej transkripcie***

Ako som už zmienila vyššie, česká transkripcia sa javí efektívnejšia pri kvantitatívnom spracovávaní čínskeho jazyka. V nasledujúcej časti priblížim niekoľko základných problémov, ktoré sa objavili pri výbere transkripcie a ktorých riešenie nakoniec vyústilo do konečného výberu transkripcie pre ďalšiu prácu.

##### **4.1.1 Finálna posícia i vs. e**

Nakoľko čínština obsahuje prevažne otvorené slabiky, tieto slabiky končia buď jednou alebo viacerými samohláskami. V čínštine sa však nie všetky samohlásky realizujú v tej podobe, ako sú zapísané pinyinom. V prípade pinyinových hlások zapisovaných písmenami zh, ch, a sh nastáva problém v spojení s finálami i a e. Obe tieto finály sa nerealizujú na základe ich grafickej podoby, ale poukazujú len na kvalitu neurčitej samohlásky, ktorá za týmito spoluhláskami nasleduje. Ak po nich nasleduje i, táto slabika sa realizuje ako spoluhláska s neprízvučnou neurčitou samohláskou v IPA

---

<sup>103</sup> TŘÍSKOVÁ, s. 18.



označovanou ako /ə/. Ak však po nich nasleduje samohláska e, táto samohláska naznačuje zmenu kvality realizovanej neurčitej samohlásky, ktorá je v tomto prípade prízvučná a označuje sa /ɜ/. Česká transkripčia tieto prípady rozlišuje tým, že v prvom prípade za spoluhlásku pripojí apostrof a v druhom prípade ponechá písmeno e. Na názornom príklade dvoch čínskych slabík prepísaných do českej transkripcie si ukážeme tento rozdiel. V čínskej abecede napísané slabiky zhi a zhe sa v českej transkripcii budú vyskytovať ako č' a če. Vzhľadom na povahu výskumu, ktorý sa zameriaval hlavne na kvantitatívnu a nie kvalitatívnu zložku jazyka sa česká transkripčia javí ako efektívnejšia. Apostrof totiž považujem za grafém, ktorý indikuje neurčitú samohlásku. Tým pádom sú pri oboch slabikách v pinyin realizované dve hlásky, ale zaznamenané sú tri, na rozdiel od českej transkripcie, kde je zaznamenaný rovnaký počet hlások ako tých realizovaných a teda dve.

#### **4.1.2 Iniciálne aspirované spoluhlásky**

Čínština nerozlišuje na rozdiel od slovenčiny alebo češtiny znelé a neznelé spoluhlásky. Avšak namiesto toho používa aspirované a neaspirované spoluhlásky. V pinyin používame písmená b, d, g na zaznamenanie zvukov p, t, k. Písmená p, t, k v pinyin sú v českej transkripcii zaznamenané ako pch, tch, kch.<sup>104</sup> Pri spracovávaní dát na základe pinyinu tým pádom môže veľmi ľahko dôjsť k systémovej chybe, pri ktorej by v pinyin slabika pan mala započítané 3 hlásky, avšak jej skutočná výslovnosť je pchan a sú tu realizované hlásky štyri. Nakoľko hláska ch sa síce skladá z dvoch písmen, ale máme zaužívané považovať ju za jednu hlásku, postupujeme na základe toho aj v tomto experimente. Znova sa ukazuje, že česká transkripčia je efektívnejšia, nakoľko výrazne odlišuje aspirované a neaspirované slabiky a tým pádom napomáha lepšej orientácii v texte pri procese kvantifikácie.

#### **4.1.3 Iniciálne palatálne hlásky**

Iniciálne palatálne hlásky v pinyin sú zaznamenané ako zh, ch, sh avšak ich prepis v českej transkripcii je č, čch, š. Aj keď v pinyin tieto spoluhlásky pozostávajú z dvoch písmen, ich fonetická realizácia pri zh a sh pozostáva len z jednej hlásky. Palatálna hláska zaznamenaná v pinyin ako ch je výnimkou, nakoľko táto konzonanta sa

---

<sup>104</sup> TRÍSKOVÁ, s. 24.

realizuje ako čch a keďže písmeno ch chápeme ako jednu jednotku tak táto spoluhláska tým pádom pozostáva z jednej písmenovej jednotky, ale realizuje sa dvoma hláskami. V oboch prípadoch však česká transkripcia profesora Švarného presnejšie zachytáva skutočnú fonetickú realizáciu daných hlások a tým pádom aj slabík. Napríklad slabiky zhu a shu pozostávajú v pinyinine z troch písmen, ale realizujú sa pri oboch len dve hlásky t.j. ču a šu, čo presne zaznamenáva česká transkripcia. Preto aj v tomto prípade je pinyin neefektívny v rámci procesu kvantifikácie a môže znovu spôsobiť problémy a systémové chyby, ktoré by sa mohli odraziť na výsledku. Aj tu sa teda česká transkripcia ukázala ako vhodnejší nástroj pre potreby kvantifikácie.

#### **4.1.4 Iniciálna spoluhláska q**

Pinyin používa písmeno q na označenie zvuku, ktorý je veľmi podobný zvuku písmena č. Takisto patrí do podskupiny afrikatív a v našom jazyku sa nevyskytuje. Toto pinyinové písmeno však v žiadnom prípade nezastupuje písmeno q v našom ponímaní s výslovnosťou /kv/. V českej transkripcii sa ale používa na prepis tejto hlásky kombinácia písmen čch. Táto kombinácia sa používa aj pri prepise pinyinového písmena ch, avšak vzhľadom na nasledujúcu finálu je možné zreteľne a bez akýchkoľvek pochybností určiť, či ide o pinyinové písmeno ch alebo q. Čo sa ale týka efektivity pri kvantitatívnom spracovaní textu, dochádza tu k zaujímavému fenoménu. Pinyinové písmeno q reprezentuje jednu hlásku, avšak v českom prepise sa používanú hlásku dve. Zobrazené na príklade to môžeme porovnať nasledovne: v pinyinine používaná slabika qu obsahuje dve písmená a rovnako sa pri nej realizujú dve hlásky. Naproti tomu tá istá slabika v českej transkripcii sa prepisuje ako čchü a tým pádom je zapisovaná tromi písmenovými jednotkami, ale realizujú sa tu len dve hlásky. Znovu podotýkam, že české písmeno ch považujem za samostatný celok a počítá sa za jedno písmeno. V tomto prípade je zrejme, že čo sa týka tejto spoluhlásky, je pinyin na rozdiel od českej transkripcie pri kvantitatívnom spracovaní efektívnejší.

#### **4.1.5 Trojhlásky /iou/, /uei/**

Ako posledný fonetický problém vyvstal problém pinyinových finál –iu a –ui. Obe tieto finály sa v pinyinine zapisujú pomocou dvoch písmen, avšak pri realizácii je medzi ne vložená ďalšia samohláska -o alebo -e. To v praxi znamená, že pri finále –iu je medzi prvú a druhú hlásku vložená samohláska -o a v prípade finály –ui je medzi prvú a druhú

hlásku vložená samohláska -e. Tým pádom pinyinová slabika liu pozostáva z troch písmen, rovnako ako pinyinová slabika dui, avšak v oboch prípadoch ich fonetická realizácia vyžaduje hlásky štyri. Česká transkripčia prepisuje tieto slabiky ako liou a duej, čo znovu presnejšie vystihuje skutočnú realizáciu daných slabík, keďže používa štyri písmená na zobrazenie výslovnosti štyroch hlások. Pre kvantitatívny výskum je teda aj v tomto prípade efektívnejšia česká transkripčia, nakoľko znovu vernejšie zobrazuje skutočnú realizovanú výslovnosť a tým pádom zabraňuje vzniku systémovej chyby pri spracovávaní dát.

Nižšie umiestnená tabuľka 1 sumarizuje vyššie popísané hlavné problematické aspekty pinyinu a českej transkripcie, ktoré boli uvažované pri výbere vhodného nástroja na prepis. Táto tabuľka slúži na zhrnutie názorných rozdielov v zaznamenávaní hlások pri pracovaní s pinyinom a českou transkripciou.

Tabuľka 1: Zhrnutie foneticky problematických aspektov<sup>105</sup>

Problematický aspekt	Pinyin (počet písmen/počet fonémov)	Česká transkripčia (počet písmen/počet fonémov)
Finálna pozícia i vs. E	<i>zhi</i> (3/2) <i>zhe</i> (3/2)	<i>č'</i> (2/2) <i>če</i> (2/2)
Iniciálne aspirované spoluhlásky	<i>ban</i> (3/3) <i>pan</i> (3/4)	<i>pan</i> (3/3) <i>pchan</i> (4/4)
Iniciálne spoluhlásky	<i>zhu</i> (3/2)	<i>ču</i> (2/2)
Iniciálna konzonanta q	<i>qu</i> (2/2)	<i>čchü</i> (3/2)
Trojhlásky /iou/, /uei/	<i>liu</i> (3/4) <i>dui</i> (3/4)	<i>liou</i> (4/4) <i>tuej</i> (4/4)

Z tabuľky a z popisu vyššie je zrejmé, že v prípade kvantitatívneho spracovania dát je výhodnejšie používať českú transkripciu, nakoľko lepšie kopíruje skutočnú realizáciu slabík a tým pádom zaručuje efektívnejšie a presnejšie spracovanie dát. Samozrejme aj pri tejto transkripcii sa vyskytujú výnimky, kde počet realizovaných hlások nezodpovedá počtu písmen, avšak v porovnaní s ostatnými transkripciami je výskyt

<sup>105</sup> SCHUSTEROVÁ, Denisa. ŠČIGULINSKÁ Jana et al. **An Application of the Menzerath-Altmann Law to Contemporary Spoken Chinese.** *Czech and Slovak Linguistic Review*, 2013, č. 2, (v tlači). s. 4.

týchto výnimiek relatívne marginálny. Je ale potrebné byť si vedomý týchto výnimiek a pri spracovaní sa na ne zamerať a neprehliadať ich, aby nedochádzalo k nepresnostiam. Pre účely tohto skúmania a kvantifikácie však česká transkripcia bola vyhodnotená ako najvhodnejší nástroj pre prípravu súboru na ďalšie skúmanie.

Po konečnom definovaní jednotiek, výbere transkripcie a segmentácii výberového súboru následne môžeme pristúpiť k prezentácii, analýze a interpretácii získaných výsledkov. V nasledujúcej kapitole sa venujem praktickej časti výskumu, predkladám získané dáta vo forme tabuliek a grafov a pristupujem k diskusii, v ktorej na základe jednotlivých výsledkov formulujem príslušné závery.

## 5. Výsledky a diskusia

Ako je spomenuté v predchádzajúcich častiach, celý výskum sa odvíjal od vybratia vhodného skúmaného súboru, stanovenia jednotiek a ich definícií, segmentácie a následnej aplikácie Menzerath-Altmanovho zákona. So všetkými týmito súčasťami, ich úskaliami a definíciami sme sa mohli oboznámiť v predošlej teoretickej časti tejto práce. Nasledujúca časť bude zameraná na prezentáciu výsledkov, ich vysvetlenie, interpretáciu, ako aj na návrhy ďalších postupov a smerov výskumu. Aj keď hlavnou súčasťou tejto práce je úroveň zameraná na vzťah výpovede a prízvukového taktu, meraného v priemernom počte jeho slabík, v práci sa okrajovo zmienim aj o ďalších dvoch úrovniach, nakoľko v rámci Menzerath-Altmanovho zákona sú jednotlivé úrovne previazané a konštrukt na jednej úrovni je zároveň konštituentom na tej druhej. Znovu spomínam jednotky, ktoré boli stanovené a s ktorými prebiehala prvá fáza výskumného procesu. Sú to hláska – slabika – prízvukový takt – výpoveď – replika. V druhej fáze výskumu bol pre lepšie porovnanie k dvom súborom z prvej fázy pridaný odlišný tretí výberový súbor, ktorý bol ukážkou segmentácie prof. Švarného. V tomto súbore sú jednotlivé jednotky nazývané: hláska – slabika – binára – segment – kolón – veta – text. Pripomínam, že rozdiely medzi jednotlivými jednotkami sú taktiež opísané v príslušnej kapitole teoretickej časti.

Štruktúrne je nasledujúca časť usporiadaná nasledovne: príslušné dáta na každej úrovni sú v rámci prehľadnosti usporiadané do tabuliek s popisom, ktorý označuje ku ktorej úrovni a výberovému súboru sa dané dáta vzťahujú. Pod tabuľkou nasleduje bodový graf vygenerovaný na základe príslušných dát v programe Microsoft Excel, na ktorom je možné vidieť tendenciu usporiadania bodov. Nakoniec nasleduje graf vygenerovaný programom R, ktorý na rozdiel od excelového grafu zobrazuje okrem bodov aj tendenciu MALu pomocou krivky, vytvorenej na základe vstupných dát. Táto krivka sa snaží čo najpresnejšie ukázať zhodnú prípadne nezhodnú tendenciu Menzerath-Altmanovho zákona s daným výberovým súborom a je ohýbaná a upravovaná na základe bodov, ktoré nekopíruje, ale snaží sa nájsť čo možno najpresnejší prienik medzi bodmi, na základe Menzerath-Altmanovho zákona. Po predstavení grafov následne pristúpim k interpretácií daných výsledkov, ich vzájomnému porovnaniu a na záver z nich vyvodím zhodnosť prípadne nezhodnosť s vopred stanovenými hypotézami.

Konkrétny výskum prebiehal v rámci troch úrovní, ktoré sú nazvané podľa najvyššieho konštruktú, a ktoré zahŕňajú okrem tejto jednotky ešte dve nižšie jednotky. Tieto úrovne nazývam U1, U2 a U3, kde:

U1 = Replika meraná vo výpovediach, ktoré sú merané v priemernom počte ich prízvukových taktov

U2 = Výpoveď meraná v prízvukových taktoch, ktoré sú merané v priemernom počte ich slabík

U3 = Prízvukový takt meraný v slabikách, ktoré sú merané v priemernom počte ich hlások.

Aj keď tieto tri úrovne boli v rámci výskumu rovnocenné a bol im venovaný rovnaký priestor a pozornosť, táto práca sa zameriava hlavne na U2, preto tejto úrovni venujem najväčší priestor a taktiež sa budem zaoberať interpretáciou výsledkov najmä na tejto úrovni. Z rovnakého dôvodu budem porovnávať výsledky z pôvodných výberových súborov spolu s vybranou časťou práce prof. Švarného len na úrovni U2.

Jednotlivé súbory sú rozdelené na výberový súbor A, ktorý pozostáva z reči popovej speváčky, výberový súbor B, ktorý má na úrovni U1 a U2 dve verzie a pozostáva z reči čínskeho rapera a nakoniec je to výberový súbor C, ktorý má taktiež dve verzie a je vybraný z diela prof. Švarného. Tento súbor, ktorý pozostáva z vybraných častí učebného materiálu Hovorová čínština v *příkladech*, a výsledky získané na základe jeho segmentácie a následného skúmania, používam len na U2.

### ***5.1 U1: Replika meraná vo výpovediach, ktoré sú merané v priemernom počte ich prízvukových taktov***

Na rozdiel od nasledujúcich úrovní, konštrukt tejto konkrétnej úrovne stojí na kraji vopred zvolených jednotiek. V takýchto prípadoch sa výskumníci väčšinou stretávajú s tým istým problémom, ktorým je možná nedostatočná frekvencia konštituentu v rámci daného konštruktú. V tomto výskume boli jednotlivé repliky navzájom oddeľované pomocou hudobných vsuviek originálneho programu, nakoľko medzi danými

prerušeniami bola zrejماً významová príbuznosť jednotlivých výpovedí. Preto bolo rozhodnuté za repliku považovať celok tematicky príbuzný, oddelený hudobným klipom daného interpreta. V experimente boli na tejto úrovni v rámci výberového súboru B vytvorené dve verzie, keďže na U2 boli počas segmentácie taktiež vytvorené dve verzie. Tým pádom boli hranice a dĺžky konštruktov U2, ktoré zároveň vystupujú ako konštituenty U1 pozmenené, a preto bolo nutné vytvoriť dve príslušné verzie aj na U1. Dôvodom rozhodnutia pre vytvorenie dvoch verzií na U2 sa budem venovať v príslušnej časti, na tomto mieste je však potrebné zmieniť, že práve kvôli týmto dvom verziám v rámci výberového súboru B na U2 bolo nutné ponechať aj dve verzie na U1.

Tabuľka 2: U1: konštrukt  $x_1$  – dĺžka repliky (v počte jej výpovedí),  $z_1$  – frekvencia konštruktov, konštituenty  $y_1$  – priemerná dĺžka výpovede (v počte prízvukových taktov), výberový súbor A.

$x_1$	Súbor A	
	$z_1$	$y_1$
2	2	3,0000
4	1	6,0000
7	1	5,2857
10	1	2,3000
12	1	4,3333
15	1	5,3333
19	1	4,3158
20	1	2,9000
21	1	3,7143
39	1	2,8462

Tabuľka 3: U1: konštrukt  $x_1$  – dĺžka repliky (v počte jej výpovedí),  $z_1$  – frekvencia konštruktov, konštituenty  $y_1$  – priemerná dĺžka výpovede (v počte prízvukových taktov), výberový súbor B, verzia 1 a 2.

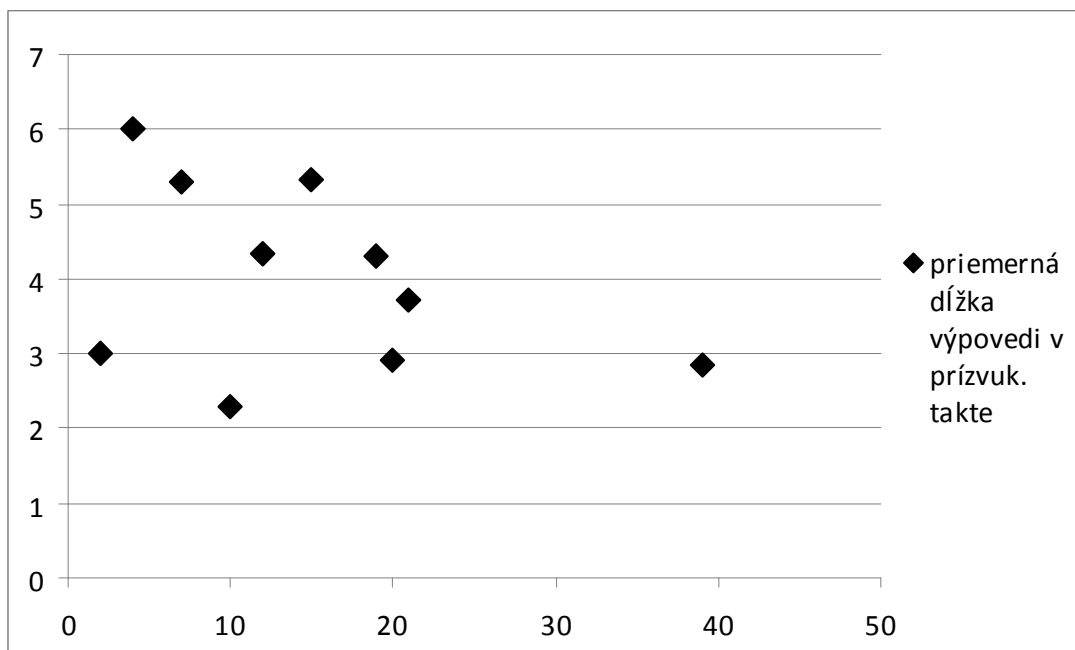
$x_1$	Súbor B, verzia 1		Súbor B, verzia 2	
	$z_1$	$y_1$	$z_1$	$y_1$
1	3	3,0000	1	5,0000
2	1	4,5000	3	2,1667
3	4	3,0000	1	3,3333
4	1	5,2500	3	2,1667
5	1	5,4000	2	4,8000
7	1	6,0000	–	–

8	1	4,6250	1	5,2500
11	–	–	1	3,3636
12	2	4,1250	–	–
13	1	5,1538	–	–
15	–	–	1	3,2667
27	–	–	1	1,8889
29	–	–	1	2,3103

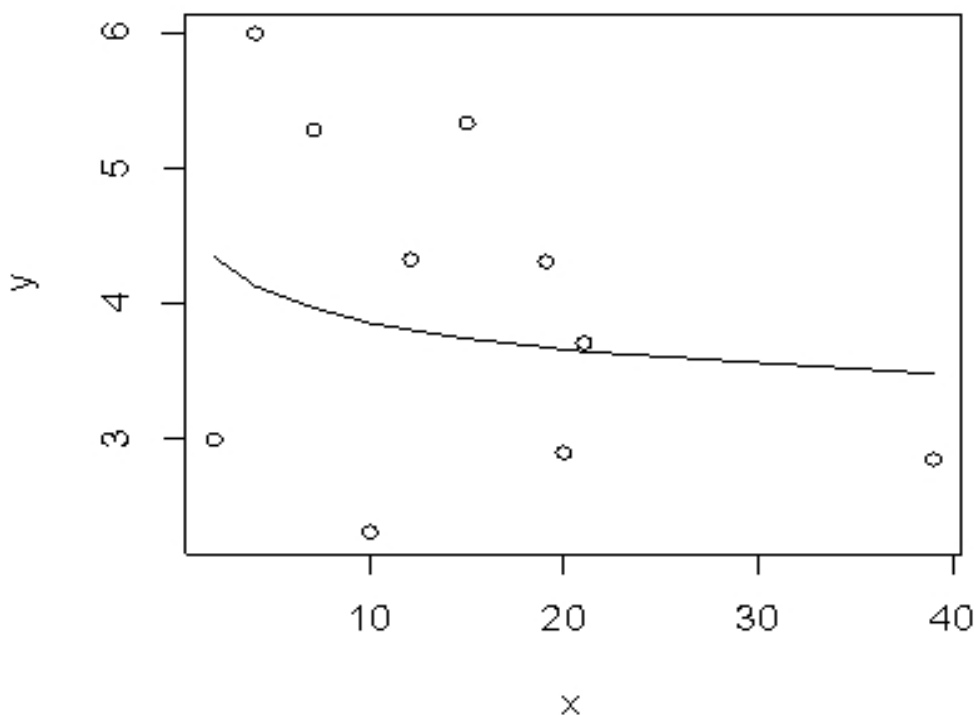
Na základe tabuliek 2 a 3 je možné vidieť vyššie spomínaný problém frekvencie jednotlivých dĺžok konštruktu. V tomto prípade dĺžka konštruktu t.j. repliky variuje v rozmedzí 1 – 39 konštituentov t.j. výpovedí v jednom konštrukte. Taktiež je možné z tabuliek vidieť, že frekvencia týchto konštruktov nepresahuje číslo tri, čo je pre výskum veľmi nedostatočná frekvencia. Ako som už spomínala vyššie, s týmto problémom sa stretáva veľké množstvo výskumníkov, nakoľko koncové úrovne predstavujú z hľadiska frekvencie výrazný problém, keďže vo väčšine prípadov je daný výberový súbor dostatočne rozsiahly na skúmanie nižších hladín, ale neobsahuje dostatočný počet informácií na presné a relevantné výsledky na hladine vyššej.

Nasledujúce grafy zobrazujú dáta z tabuľky 2 a 3, sú vytvorené v programe Microsoft Excel a pomocou programu R. Keďže to, čo nás zaujíma, sú jednotlivé body, ktoré reprezentujú jednotlivé pozorovania, rozhodla som sa pre formu bodových grafov, na ktorých je možné vidieť určitú tendenciu a pre naše účely je daná forma najvyhovujúcejšia. Pre lepšie porovnanie zoradujem za seba vždy graf z programu Microsoft Excel a vzápätí za ním príslušný graf z R-ka, v ktorom je vizualizovaná krivka MAL. Táto krivka reprezentuje najbližšiu reprezentáciu MALu na základe daných pozorovaní a daný program okrem vytvorenia príslušnej krivky vyčísluje aj percentuálnu zhodu s Menzerath-Altmanovým zákonom.





Graf 1.A: U1 (replika vs. výpoveď) – súbor A: vizualizácia súboru dát v tabuľke 2, pomocou programu Microsoft Excel.



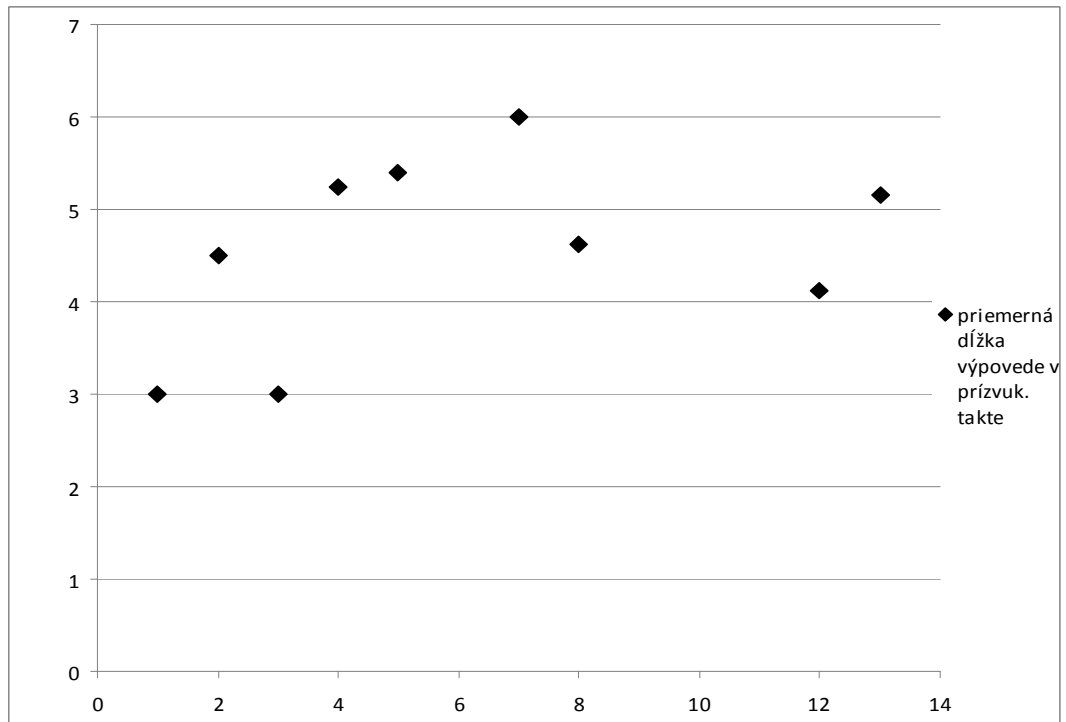
Graf 1.Ar: U1 (replika vs. výpoveď)<sup>106</sup> – súbor A: vizualizácia súboru dát v tabuľke 2, pomocou programu R.

<sup>106</sup> SCHUSTEROVÁ, s. 8.

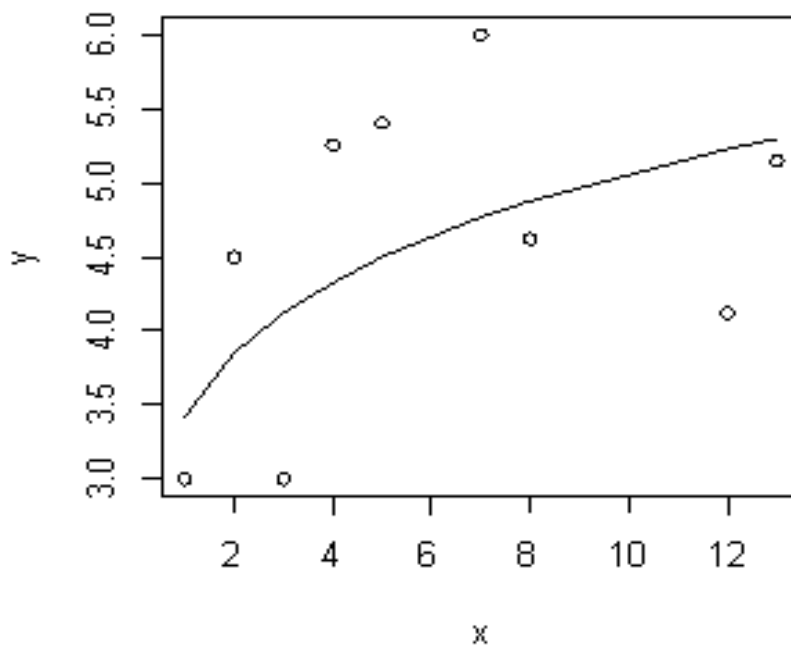
Ako je možné vidieť na grafe 1.A, jednotlivé pozorovania nevykazujú určitú tendenciu, nakoľko sú viac menej roztrúsené a sú medzi nimi pomerne výrazné intervaly. Na grafe 1.Ar je už možné vidieť aj preloženie daných bodov konkrétnou krivkou, ktorá reprezentuje MAL. Táto zhoda je číselne vyčíslená ako  $R^2 = 4.16 \%$  čo je veľmi nízka percentuálna zhoda s vykonštruovaným matematickým modelom, ktorého vizualizáciou je daná krivka. Aj napriek tomu, že krivka vykazuje klesajúcu tendenciu, percentuálna zhoda poukazuje na to, že na danej úrovni a na danom výberovom súbore sa tendencia MALu nepotvrdila. Možnými príčinami tejto malej zhody, resp. výraznej nezhody s MALom sa budem venovať neskôr, čo sa však týka výberového súboru A, z príslušných grafov a číselných údajov je táto malá zhoda očividná.

Pri segmentovaní výberového súboru B som sa rozhodla pristúpiť k dvom verziám segmentácie, nakoľko v reči daného interpreta sa vyskytovalo veľké množstvo vulgarizmov a sémanticky prázdnych slov, ktoré vo svojom prejave používal veľmi často. Táto okolnosť mala potenciál výrazne ovplyvniť konečný počet konštruktov a konštituentov, preto som pristúpila k dvojitej segmentácii. V prípade verzie č.1 boli dané vulgarizmy resp. sémanticky prázdne slová či skupina slov systematicky pričleňované k výpovedi stojacej pred nimi, zatiaľ čo v prípade verzie č. 2 boli vulgarizmy a sémanticky prázdne slová či skupiny slov systematicky vyčleňované ako samostatne stojace výpovede. Podotýkam, že toto vyčlenenie sa týkalo len vulgarizmov, ktoré nemali rozvíjajúcu funkciu a teda nestáli v pozícii adjektíva.

Z nasledujúcich grafov je zrejmé, že toto odlišné definovanie daných problematických častí a následná segmentácia podľa vopred určených kritérií pre verziu č. 1 a verziu č. 2 mali na konečné výsledky výrazný vplyv.

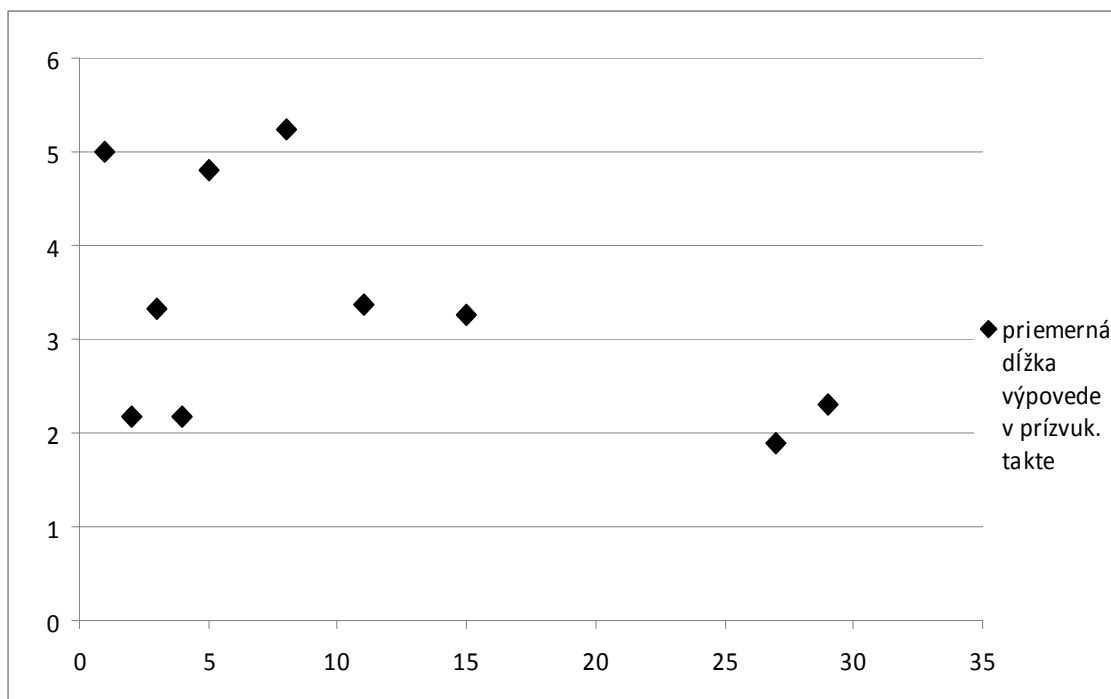


Graf 1.B1: U1 (replika vs. výpoveď) – súbor B verzia 1: vizualizácia súboru dát v tabuľke 3, pomocou programu Microsoft Excel.

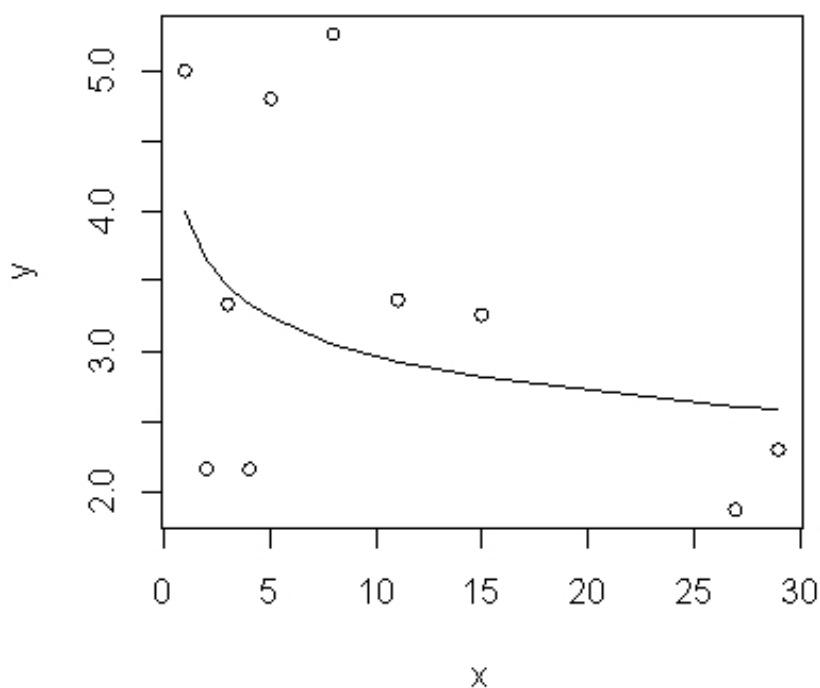


Graf 1.B1r: U1 (replika vs. výpoveď)<sup>107</sup> – súbor B verzia 1: vizualizácia súboru dát v tabuľke 3, pomocou programu R.

<sup>107</sup> SCHUSTEROVÁ, s. 9.



Graf 1.B2: U1 (replika vs. výpoveď) – súbor B verzia 2: vizualizácia súboru dát v tabuľke 3 , pomocou programu Microsoft Excel.



Graf 1.B2r: U1 (replika vs. výpoveď)<sup>108</sup> – súbor B verzia 2: vizualizácia súboru dát v tabuľke 3, pomocou programu R.

Pre názornejší prehľad číselných údajov parametrov a percentuálnej zhody jednotlivých súborov a verzií, vkladám na toto miesto prehľadnú tabuľku s jednotlivými údajmi.

<sup>108</sup> SCHUSTEROVÁ, s. 9.

Tabuľka 4: Prehľad parametrov <sup>b</sup> MAL a koeficientov determinácie R<sup>2</sup> na U1<sup>109</sup>

		Parameter <sup>b</sup>	Koeficient determinácie R <sup>2</sup> (%)
Súbor A		0.0748	4.16
Súbor B	Verzia 1	-0.1717	34.1
	Verzia 2	0.1279	13.97

Z grafov 1.B1 a 1.B2 je zrejmé, že rovnako ako v prípade grafu 1.A sa v jednotlivých pozorovaniach neprejavuje výrazná klesajúca tendencia. Vo všetkých prípadoch sú konkrétne bodové reprezentácie pozorovaní rozmiestnené pomerne nerovnomerne, bez ustálenej tendencie. Avšak z grafov 1.B1r a 1.B2r je okamžite možné vidieť výrazný rozdiel, na ktorý poukazuje krivka reprezentujúca MAL vytvorená na základe vstupných dát. Na grafe 1.B1r má táto krivka stúpajúcu tendenciu, na rozdiel od grafu 1.B2r, kde má tendenciu klesajúcu. Čo sa týka percentuálnej zhody s MAL na tejto úrovni, dáta výberového súboru B v 1. verzii vykazujú zhodu  $R^2 = 34.1\%$  a dáta rovnakého súboru v 2. verzii vykazujú percentuálnu zhodu  $R^2 = 13.97\%$ . Aj keď je v prípade súboru B verzie 1 v porovnaní s ostatnými pomerne vysoká zhoda s MAL, krivka vykazuje stúpajúcu tendenciu, čo má na svedomí parameter b, ktorý je v tomto prípade záporný a tým pádom spôsobuje túto stúpajúcu krivku.

Na všetkých zobrazených grafoch 1.A, 1.B1 a 1.B2 sa nevyskytuje žiadna podobná tendencia. Na každom grafe sú body reprezentujúce jednotlivé pozorovania usporiadané veľmi nesúrodne, avšak z týchto grafov nie je možné vyčítať možnú zhodu s MAL, nakoľko na zistenie tejto zhody je nutná krivka, ktorá daný zákon reprezentuje a je aplikovaná na dané pozorovania. V prípade grafov 1.Ar, 1.B1r a 1.B2r na ktorých sa vizualizácia matematického modelu pomocou krivky nachádza, je možné vidieť klesajúcu tendenciu na v prípade grafov 1.Ar a 1.B2r a stúpajúcu tendenciu na grafe 1.B1r. Avšak vo všetkých troch prípadoch bola zhoda s MAL minimálna a zanedbateľná, vzhľadom na to, že v lingvistických kruhoch sa zhoda nižšia ako 75-80% nepovažuje za dostatočne preukázateľnú.

Príčin tejto pomerne nízkej zhody môže byť niekoľko. V prvom rade, ako som to už spomínala vyššie, jedným z dôvodov je krátkosť výberových súborov a tým pádom nízka frekvencia jednotlivých konštruktov. Ďalším dôvodom môže byť prípadná

<sup>109</sup> SCHUSTEROVÁ, s. 10.

medzera medzi jednotlivými hladinami, čo znamená, že medzi jednotkami replika a výpoveď sa môže nachádzať ďalšia jazyková hladina. Výsledky mohli byť taktiež ovplyvnené tým, že pri segmentácii na tejto úrovni bolo nutné používať aj syntaktické a sémantické hľadisko, ktoré boli použité ako doplnkové kritériá ku fonetickému hľadisku. Ak by sme sa zamerali na mimojazykové dôvody nízkej zhody, veľký vplyv mohol mať umelý zásah do výberových súborov, ktorý sa udial na výberových súbor pre účely zostrihania daného hudobného programu. Na základe získaných výsledkov sa tendencia Menzerath-Altmanovho zákona na danej úrovni na hovorenej čínštine neprejavuje, čo vyvracia pôvodnú hypotézu, ktorá predpokladala, že hovorená čínština bude spĺňať podmienky a tendencie Menzerath-Altmanovho zákona. Avšak pre potvrdenie tejto novej hypotézy, podľa ktorej táto tendencia neplatí je nutné prihliadnuť k možným príčinám nízkej zhody popísaným vyššie a rozhodne je potrebné daný vzťah overiť na ďalších súboroch.

Vzťah medzi jednotkami na tejto úrovni poskytuje množstvo ďalších možností pre nasledujúci výskum, nakoľko výrazné rozdiely medzi jednotlivými grafmi vytvárajú priestor pre nové otázky a hypotézy. V tejto chvíli nie je možné zo stopercentnou istotou tvrdiť, čo spôsobilo nezhodu s MAL. Vzhľadom na to, že jednotlivé grafy sa od seba diametrálne líšia, je nutné vzťahy na tejto úrovni hlbšie preskúmať. To znamená, že je potrebné prehodnotiť stanovenie jednotiek a prípadne pridať ďalšie jazykové úrovne, rovnako ako je nevyhnutné zväčšiť rozsah výberového súboru. Po konečnom predefinovaní a navýšení jazykových jednotiek a zväčšení výberového súboru je potom možné znovu prísť k preskúmaniu vzťahov medzi jednotkami na tejto úrovni.

## **5.2 U2: Výpoved' meraná v prízvukových taktoch, ktoré sú merané v priemernom počte ich slabík.**

Na tejto úrovni bol meraný vzťah konštruktú, to znamená výpovede, a konštituentu, čo bol v tomto prípade prízvukový takt, ktorý zároveň vystupoval ako konštrukt ku svojmu konštituentu t.j. k slabike a bol meraný v priemernom počte slabík. Vzhľadom k tomu, že táto úroveň je práve tá, ktorou som sa zaoberala najviac, rozhodla som sa pri nej pristúpiť k segmentácii viacerých výberových súborov. Okrem dvoch vyššie spomínaných súborov A a B z prvej fáze výskumu, to znamená súbor s popovou speváčkou a súbor s raperom, pri ktorom som pracovala s dvoma verziami, som sa rozhodla pridať v druhej fáze výskumu ďalšie dve verzie súboru vybraného z diela prof. Švarného Hovorová čínština v príkladoch. Tento súbor pre lepší prehľad nazývam súbor C a jeho dve verzie označujem príslušným číslom.

Segmentácia súboru A bola viac menej jednoznačná, preto nebol dôvod pristúpiť k vytvoreniu viacerých verzií. Avšak pri súboroch B a C sa objavili určité komplikácie, ktoré mohli ovplyvniť výsledky experimentu, preto som sa rozhodla pre segmentáciu v dvoch verziách segmentácie a na základe výsledkov z oboch verzií pristúpiť k ich interpretácii. Rapový spevák, hlavný rečník vo vzorke B, okrem občasných anglických výrazov, ktoré čiastočne mohli spôsobiť odchýlku v U3 (tejto odchýlke sa venujem v príslušnej časti), prekladá svoj prejav častými spontánnymi vulgarizmami prípade sémanticky prázdnyimi slovami. Tieto, či už vulgárne, alebo sémanticky prázdne slová a slovné spojenia používa ako vatu, buď na vyplnenie priestoru, alebo ich môže používať ako spôsob pre získanie času, aby si mohol premyslieť ďalšiu svoju odpoveď. Jednou z možností je ale aj predpoklad, že vulgarizmy k rapovému štýlu neodmysliteľne patria, ako súčasť imidžu rapera, a preto by sa im nemala venovať prílišná pozornosť, nakoľko je to len anomália rečového prejavu daného jedinca.

Moje dve verzie počítajú s predpokladom, že vyčlenenie vulgarizmov, ako samostatných výpovedí ovplyvní celkovú zhodu daného súboru s MAL. Pôvodná hypotéza totiž tvrdí, že na danej úrovni by sa zhoda s MAL mala prejavovať, avšak jednou zo skúmaných hypotéz je aj tvrdenie, že rytmicky agresívny hudobný štýl akým rap nepochybne je, môže ovplyvniť a možno dokonca narušiť prirodzenú rytmiku reči,

v našom prípade je to rytmika čínštiny. Vo verzii 1 výberového súboru B pričleňujem vulgarizmy a sémanticky prázdne slovné spojenia k výpovedi pred nimi, a spájam ich do jedného celku. Vo verzii 2 toho istého súboru som sa ale rozhodla vyskúšať experiment a vulgárne slovné spojenia, ako aj sémanticky prázdne výpovede vyčleňujem samostatne a počítam ich ako samostatne stojacu výpoveď. K tomuto kroku avšak pristupujem len v tom prípade, že daný vulgarizmus neslúži ako rozvíjajúci člen, to znamená, že nevystupuje v pozícii zhodného alebo nezhodného prívlastku. Ako bude možné vidieť na príslušných grafoch, táto rozdielna dvojitá segmentácia má rozhodne vplyv na zhodu s MAL.

Čo sa týka výberového súboru C aj v tomto prípade som sa rozhodla uskutočniť dve rozdielne segmentácie, avšak u daného súboru boli kritériá a dôvody pre dvojitú segmentáciu a tým pádom aj dve verzie odlišné. Pri súbore C som sa v prvej verzii držala segmentácie, ktorú sme používali pri ostatných dvoch výberových súboroch. To znamená, že na U2 som súbor C vo verzii 1 rozčlenila na výpovede, merané v prízvukových taktoch, ktoré sú merané v priemernom počte ich slabík. Vzhľadom na to, že prof. Švarný používa o niečo odlišnú segmentáciu aj názvoslovie, rozhodla som sa druhú verziu spracovať podľa jeho segmentácie a teda kolón, meriam v segmentoch, ktoré meriam v priemernom počte jeho slabík. Aj keď Švarný pripodobňuje segment k prízvukovému taktu<sup>110</sup> a tým pádom by sa výpoveď mala rovnať Švarného kolónu, v prvej fáze výskumu je prízvukový takt číselne vyjadrený medzi segmentom a kólonom. Keďže súbor A aj B sú produkované pomerne rýchlo rečou, prízvukový takt sa naťahuje a tým sa pádom akoby zmazáva rozdiel medzi ním a kólonom a miestami sa takmer prekrývajú. Súbor C pochádza zo súboru viet, ktoré boli vytvorené pre študijné účely, preto je samozrejmé, že tempo reči je prispôsobené tomuto cieľu. Na základe toho je potom zrejmé, že segmentácia prof. Švarného bude prispôsobená týmto okolnostiam. Avšak porovnanie spomínaných dvoch segmentácií na tom istom súbore C považujem za zaujímavý experiment. Ako som už spomínala na začiatku, pôvodná hypotéza stanovuje, že na danej úrovni by sa popísaný vzťah MALu mal na hovorenej čínštine prejavovať.

---

<sup>110</sup> ŠVARNÝ, *Gramatika*, s. 57.



Formálne zoskupenie nasledujúcich tabuliek a grafov zachováva vopred určenú štruktúru. Ako prvé sú zaradené tabuľky s označením, aby bolo zrejmé, ktoré hodnoty súvisia s ktorým súborom a verziou. Za tabuľkami nasledujú grafy usporiadané od súboru A cez dve verzie súboru B až po dve verzie súboru C. Grafy uvádzam v dvojiciach, to znamená, že prvý stojí graf vytvorený programom Microsoft Excel a následne za ním umiestňujem príslušný graf z programu R, v ktorom je možné na základe daných pozorovaní sledovať krivku, ktorá predstavuje vizualizáciu daných údajov a s jej pomocou je možné daný graf presnejšie interpretovať vo vzťahu k MAL.

Tabuľka 5: U2: konštrukt  $x_2$  – dĺžka výpovede (v počte jej prízvukových taktov),  $z_2$  – frekvencia konštruktov, konštituenty  $y_2$  – priemerná dĺžka prízvukového taktu (v počte slabík), výberový súbor A.

$x_2$	Súbor A	
	$z_2$	$y_2$
1	20	9,3000
2	39	6,7308
3	26	5,7949
4	21	6,5238
5	14	5,2571
6	13	4,7051
7	7	4,8980
8	4	4,8125
9	2	5,7778
10	1	3,5000
11	2	3,6364
13	2	4,9231

Tabuľka 6: U2: konštrukt  $x_2$  – dĺžka výpovede (v počte jej prízvukových taktov),  $z_2$  – frekvencia konštruktov, konštituenty  $y_2$  – priemerná dĺžka prízvukového taktu (v počte slabík), výberový súbor B, verzia 1 a 2.

$x_2$	Súbor B, verzia 1		Súbor B, verzia 2	
	$z_2$	$y_2$	$z_2$	$y_2$
1	2	4,5000	37	4,2162
2	15	4,8667	30	4,3833
3	15	4,5556	18	4,4259
4	17	4,2059	17	4,5735

5	5	4,6400	5	4,6400
6	10	4,0500	4	4,0000
7	3	4,1905	4	3,8929
8	1	4,6250	–	–
9	6	3,9815	3	4,3704
10	–	–	1	3,7000
11	1	3,5455	–	–
12	1	5,3333	–	–
15	1	3,2667	1	3,3333

Tabuľka 7: U2: konštrukt  $x_2$  – dĺžka výpovede (v počte jej prízvukových taktov),  $f_{x_2}$  – frekvencia konštruktov, konštituenty  $y_2$  – priemerná dĺžka prízvukového taktu (v počte slabík), výberový súbor C, verzia 1.

	Súbor C, verzia 1	
$x_2$	$z_2$	$y_2$
1	48	5,6667
2	120	5,3875
3	86	5,1395
4	18	4,5278
5	4	4,8500

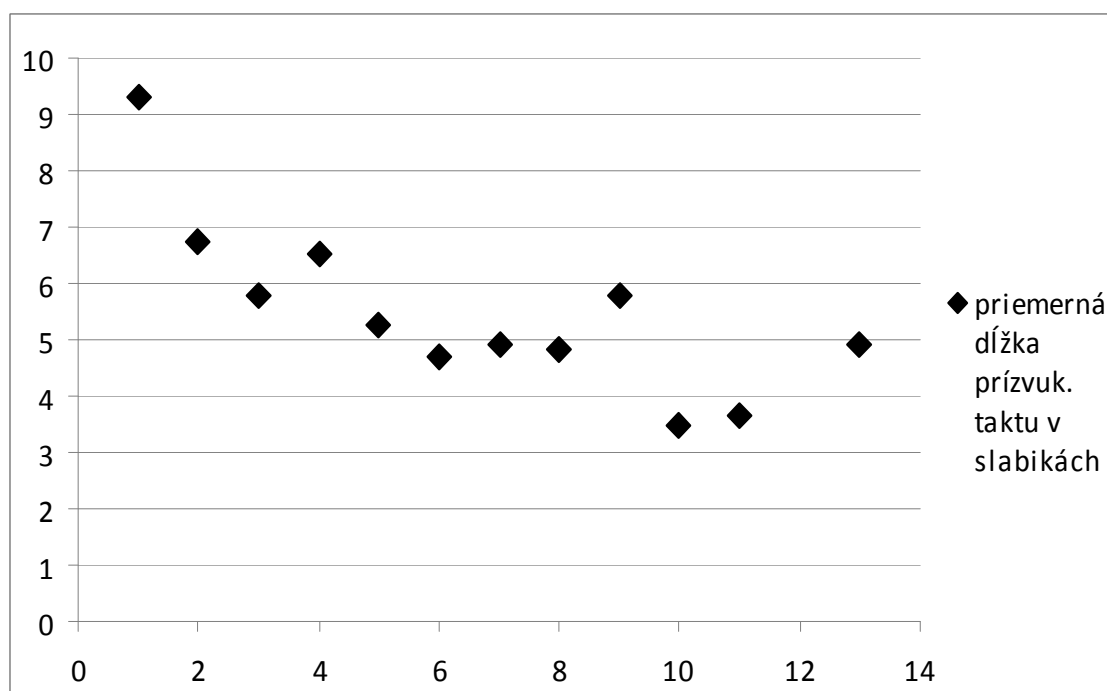
Tabuľka 8: U2: konštrukt  $x_2$  – dĺžka kolónu (v počte jeho segmentov),  $f_{x_2}$  – frekvencia konštruktov, konštituenty  $y_2$  – priemerná dĺžka segmentu (v počte slabík), výberový súbor C, verzia 2.

	Súbor C, verzia 2	
$x_2$	$z_2$	$y_2$
1	171	3,4221
2	290	2,8638
3	105	2,7302
4	23	2,5870
5	1	2,6000

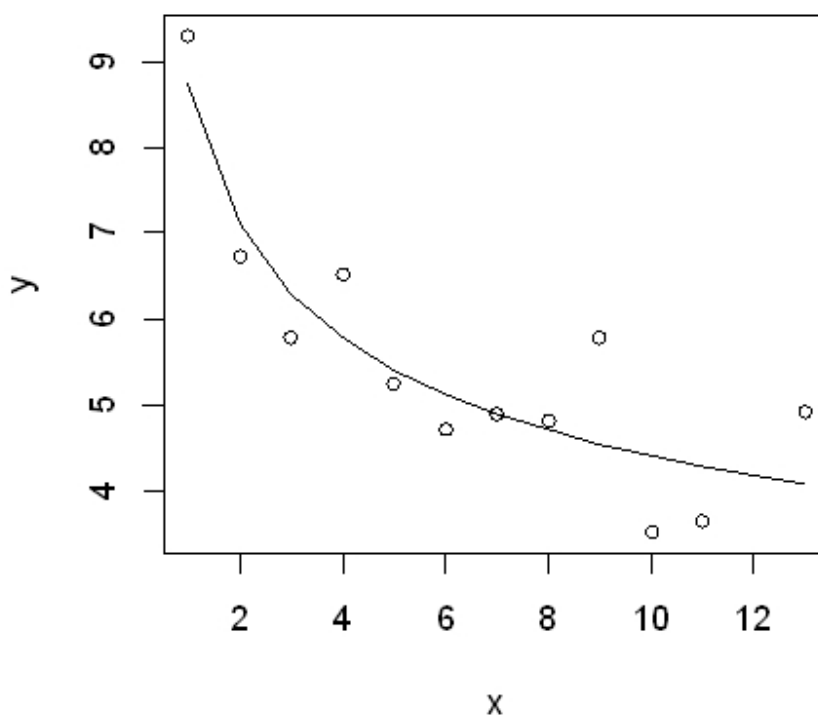
Na tejto úrovni dĺžka konštrukt (v tomto prípade je to výpoveď) variuje pri súbore A a B v intervale od 1-15 prízvukových taktov. V prípade súboru C je to v oboch verziách len maximálne 5 prízvukových taktov resp. segmentov v rámci konštrukt. Tento mierny nepomer je spôsobený tým, že pôvodné súbory A a B pochádzajú zo

spontánnej reči, zatiaľ čo výberový súbor C bol pre účely kontrastu v experimente vybraný zo súboru viet, ktoré na seba navzájom nenadväzujú a sú pomerne jednoduché. Tým pádom sa v rámci kratšieho konštruktú nachádza menší počet konštituentov.

Pri pohľade na vyššie uvedené tabuľky je ale možné vidieť, že čím je dĺžka konštruktú väčšia, tým nižšia je jej frekvencia. Tento fenomén sa objavuje ako pri súboroch A a B, tak aj pri kontrastnom súbore C. Tieto nižšie hodnoty môžu mať vplyv na konkrétny výsledok, preto vo svojom ďalšom štúdiu plánujem podrobiť tieto výsledky a s nimi súvisiace hypotézy ďalšiemu rozsiahlejšiemu výskumu. V nasledujúcej časti sa znovu zameriame na konkrétne grafické reprezentácie údajov z tabuliek 5-8 vytvorené pomocou dvoch programov.



Graf 2.A: U2 (výpoveď vs. prízvukový takt) – súbor A: vizualizácia súboru dát v tabuľke 5, pomocou programu Microsoft Excel.

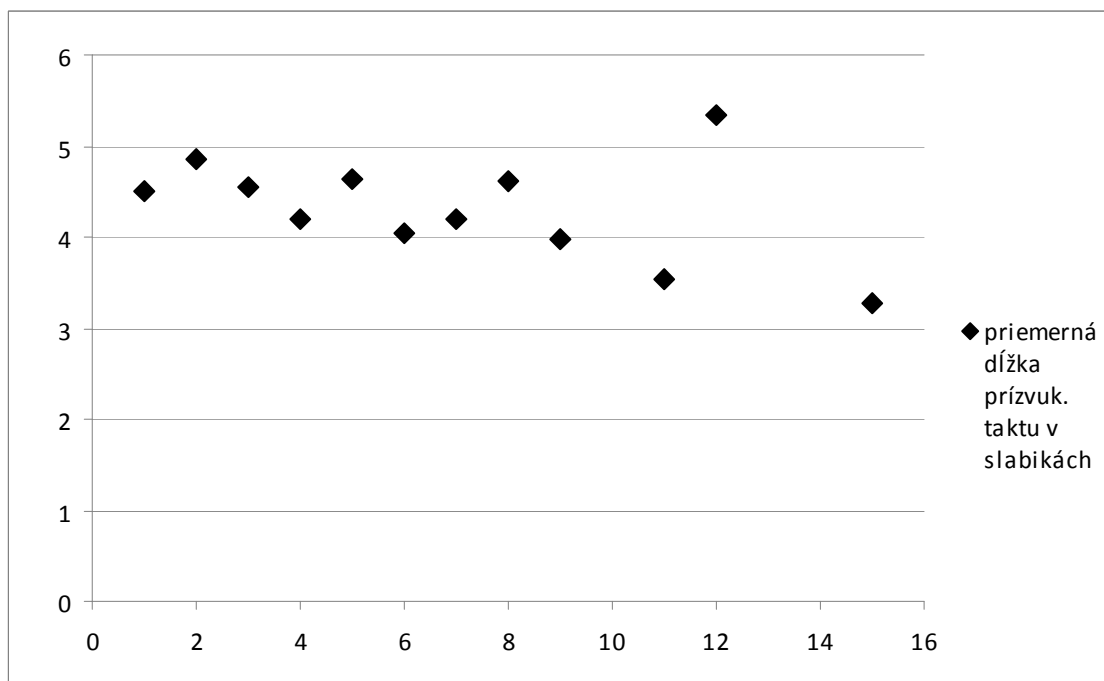


Graf 2.Ar: U2 (výpoveď vs. prízvukový takt)<sup>111</sup> – súbor A: vizualizácia súboru dát v tabuľke 2, pomocou programu R.

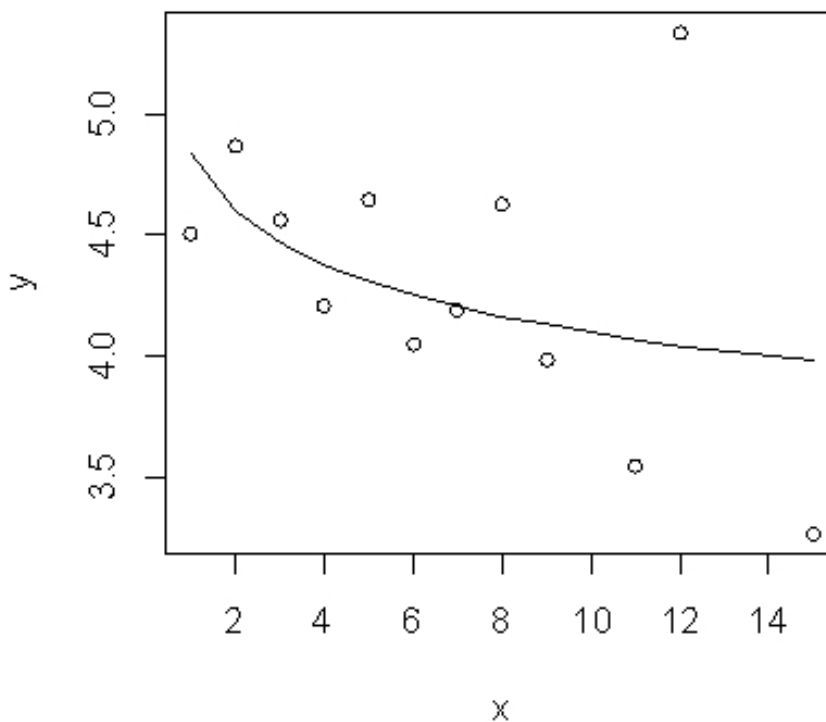
V prípade výberového súboru A je z oboch grafov jasne viditeľná výrazná klesajúca tendencia. To znamená, že vzťah medzi konštruktom a konštituentom na danej úrovni je nepriamo úmerný, čo zodpovedá predpokladu, ktorý je v lingvistike stanovený MALom. Krivka pomerne výrazne kopíruje ideálnu krivku a percentuálna zhoda s matematicky vypočítaným modelom je  $R^2 = 73.11\%$ . Ak by sme odstránili body s nízkou frekvenciou, čo sa pri podobných experimentoch považuje v lingvistike za bežnú prax, dosiahli by sme ešte výraznejšiu zhodu vyše 90%. Avšak vzhľadom na čo možno najkorektnejšie a najpresnejšie výsledky som sa rozhodla žiadne pozorovania, ani tie menej frekventované, nevypúšťať.

Na základe grafov 2.A a 2.Ar, rovnako ako na základe percentuálnej zhody je možné potvrdiť, že prirodzená hovorená číňština na tejto úrovni a v rámci tohto súboru pomerne dobre spĺňa pravidlá a definície stanovené MALom. Autorom tohto rečového úseku bola popová interpretka, autorom nasledujúcich dvoch verzií súboru B bol rapper. Jednou z otázok je, či a nakoľko ovplyvní tento hudobný žáner jeho prirodzenú reč.

<sup>111</sup> SCHUSTEROVÁ, s. 10.

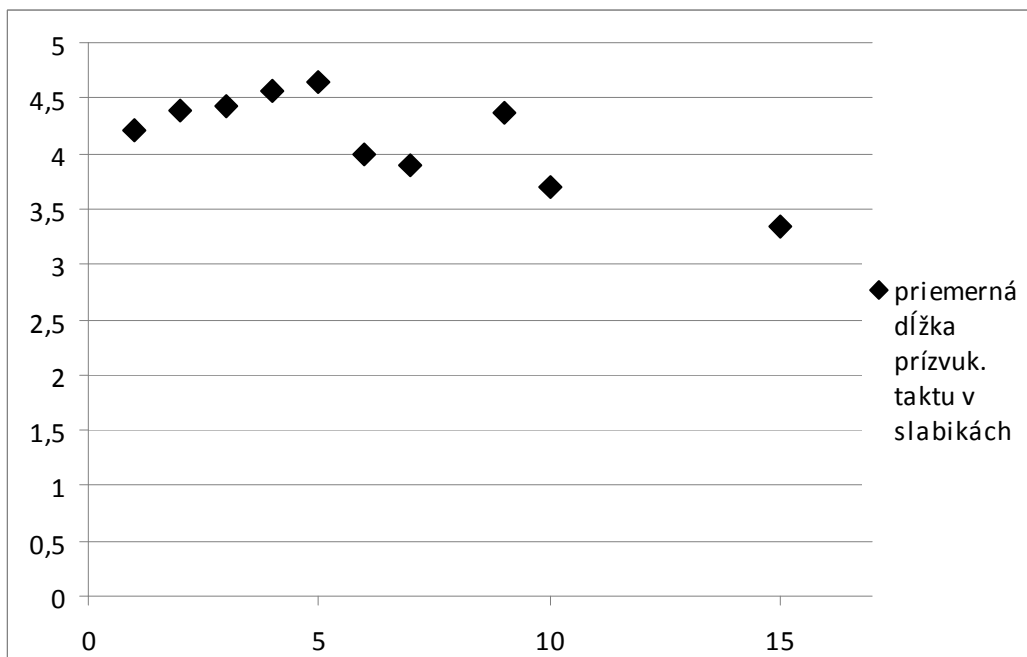


Graf 2.B1: U2 (výpoveď vs. prízukový takt) – súbor B verzia 1: vizualizácia súboru dát v tabuľke 6, pomocou programu Microsoft Excel.

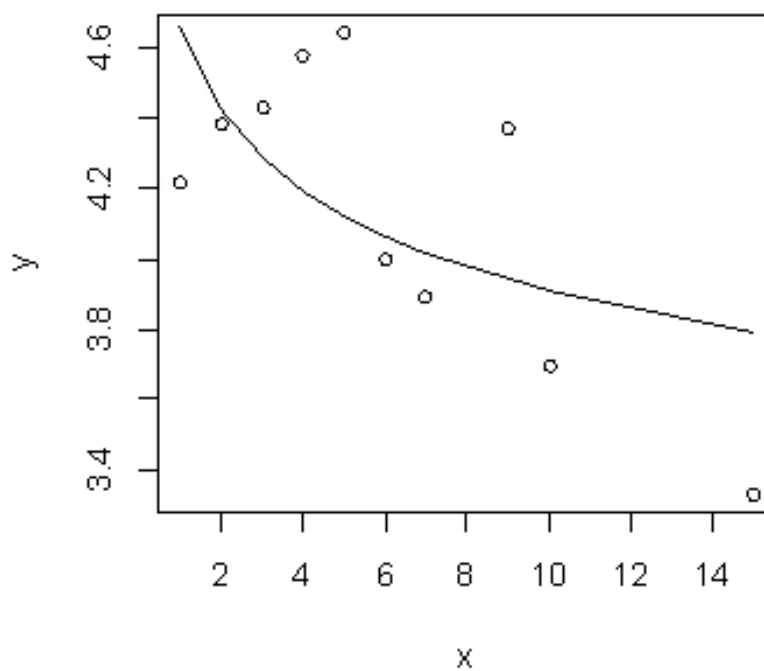


Graf 2.B1r: U2 (výpoveď vs. prízukový takt)<sup>112</sup> – súbor B verzia 1: vizualizácia súboru dát v tabuľke 6 pomocou programu R.

<sup>112</sup> SCHUSTEROVÁ, s. 11.



Graf 2.B2: U2 (výpoveď vs. prízvukový takt) – súbor B verzia 2: vizualizácia súboru dát v tabuľke 6, pomocou programu Microsoft Excel.



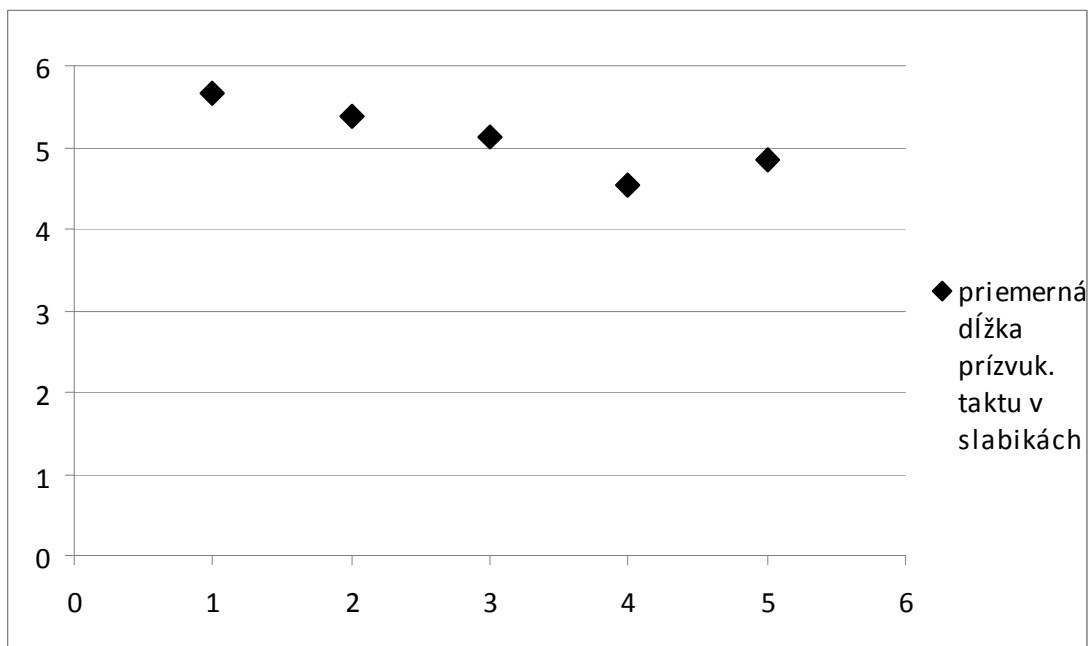
Graf 2.B2r: U2 (výpoveď vs. prízvukový takt)<sup>113</sup> – súbor B verzia 2: vizualizácia súboru dát v tabuľke 6 pomocou programu R.

<sup>113</sup> SCHUSTEROVÁ, s. 11.

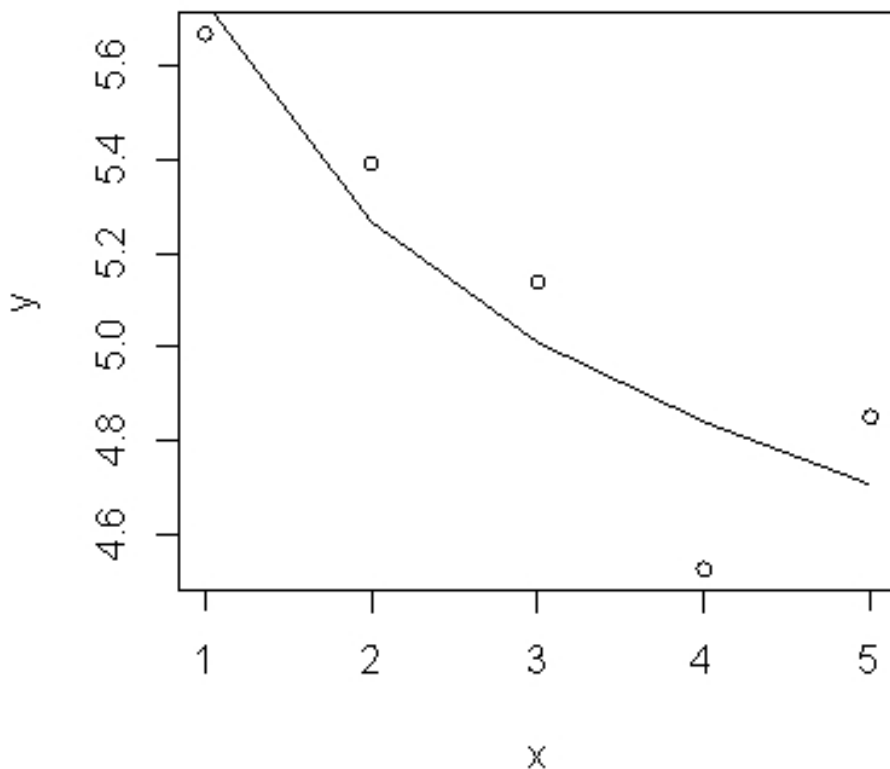
Z uvedených grafov 2.B1 a 2.B2 je klesajúca tendencia výraznejšia a lepšie viditeľnejšia u verzie číslo 1. Aj keď jednotlivé pozorovania pomerne variujú a nenachádzajú sa priamo na jednej klesajúcej priamke, určitý pokles je viditeľný. Na grafe 2.B2 je ale najskôr vidieť tendenciu stúpajúcu, až potom sa mierne stáča nadol. Čiastočne to pôsobí štýlom, že pozorovania s kratšou dĺžkou konštruktu sformovali určitú množinu, ktorá sa vymyká zákonitostiam MALu. Pozoruhodné ale je, že dlhšie konštrukty si už klesajúcu tendenciu pomerne dobre udržiavajú. V oboch prípadoch sú však body pomerne výrazne rozptýlené.

Čo sa týka kriviek na grafoch 2.B1r a 2.B2r, obe krivky vykazujú klesajúcu tendenciu, čo by zodpovedalo matematickému vyjadreniu MALu. Problémom však ostáva pomerne výrazný rozptyl jednotlivých pozorovaní, nakoľko krivka len zobrazuje najideálnejšiu stopu v rámci daných pozorovaní. Na rozdiel od grafu 2.Ar, kde je možné pozorovať veľmi malý rozptyl a vzdialenosť bodov reprezentujúcich jednotlivé pozorovania od krivky, u zvyšných grafov 2. B1r a 2.B2r sú tieto vzdialenosti výrazne dlhšie. Rovnako aj percentuálna zhoda s MALom nedosahuje ani polovicu zhody v prípade súboru A. U súboru B je pri verzii 1 percentuálna zhoda s matematickým modelom  $R^2 = 22.38 \%$ , pri verzii 2 je táto zhoda o čosi vyššia a predstavuje číselnú podobu  $R^2 = 34.78 \%$ . Oba číselné údaje len dokazujú, že daný súbor nevykazuje na tejto úrovni výraznú zhodu s matematickým modelom, a tým pádom nepodlieha vo väčšej miere Menzerath-Altmanovmu zákonu.

Nasledujúce grafy patria ku kontrastnému výberovému súboru C, ktorý pochádza z materiálu prof. Švarného z GHČP. Kontrastný je preto, že tento súbor úplne nespĺňa podmienku prirodzenej spontánnej reči, aj keď je zvukovo realizovaný rodenou Číňankou, ktorá daný text nečíta, ale produkuje ho ako bežný rečový prejav. Táto vzorka bola vybraná hlavne kvôli pomalšiemu tempu a tým pádom o niečo rozdielnejšej segmentácii prof. Švarného. Pre väčší kontrast som sa rozhodla v rámci 1. verzie nasegmentovať túto vzorku pomocou nami určených jazykových jednotiek a segmentáciu 2. verzie som zachovala podľa prof. Švarného. Jedným z dôvodov bolo porovnanie oboch typov súborov, to znamená pomerne spontánnej a prirodzenej reči u vzoriek A a B a čiastočne umelej reči, upravenej pre študijné potreby cudzincov u vzorku C. K dvom verziám pri súbore C som sa rozhodla z dôvodu porovnania potvrdenia resp. vyvrátenia správnosti Švarného segmentácie na danom súbore.

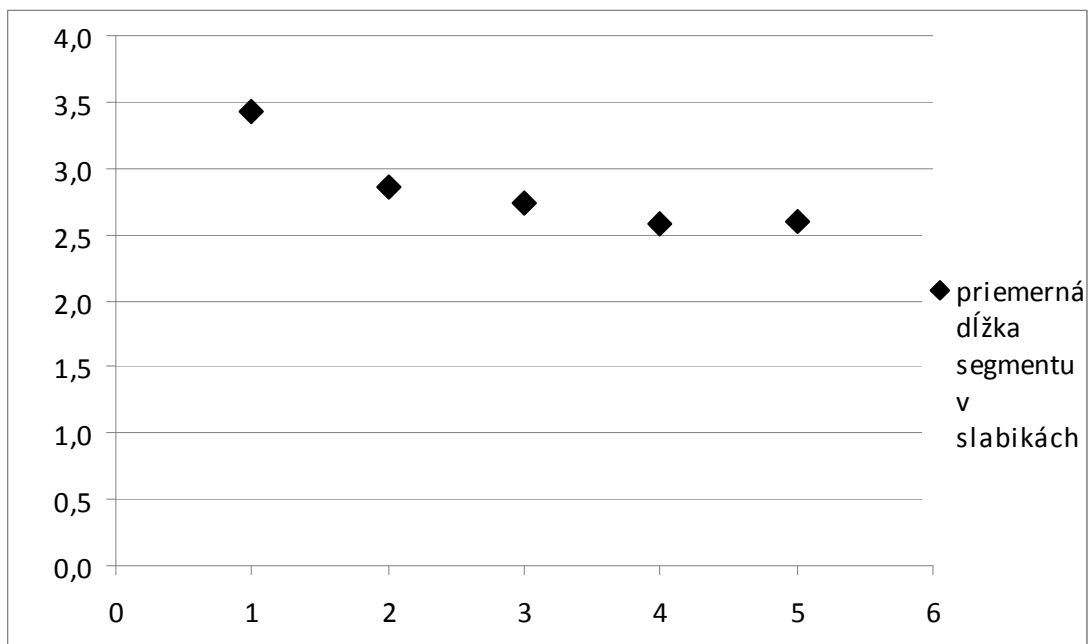


Graf 2.C1: U2 (výpoveď vs. prízvukový takt) – súbor C verzia 1: vizualizácia súboru dát v tabuľke 7, pomocou programu Microsoft Excel.

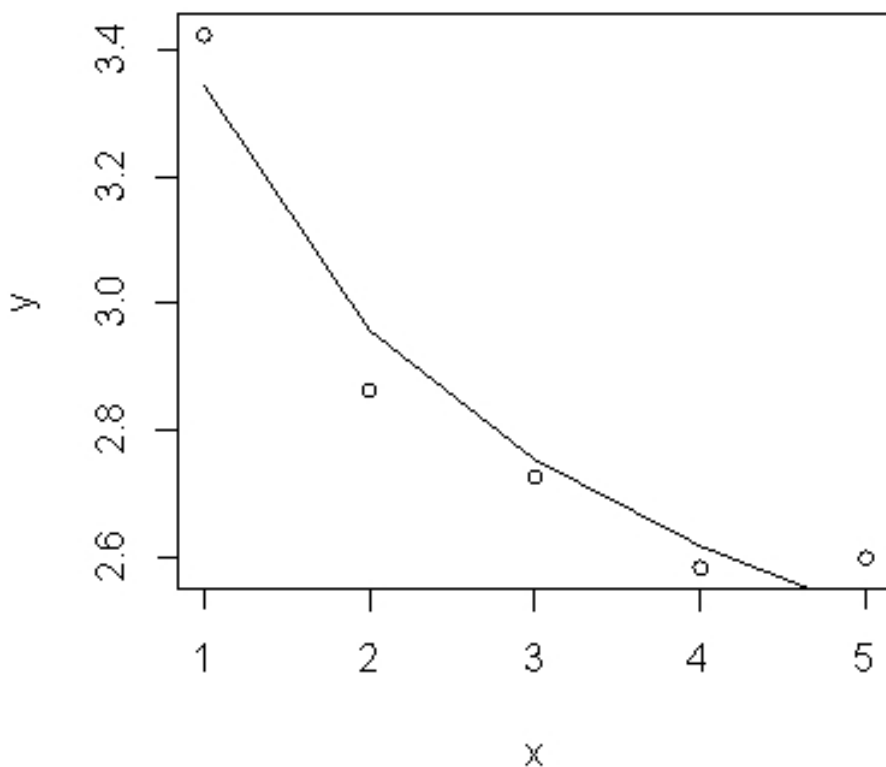


Graf 2.C1r: U2 (výpoveď vs. prízvukový takt) – súbor C verzia 1: vizualizácia súboru dát v tabuľke 7, pomocou programu R.





Graf 2.C2: U2 (výpoveď vs. prízvukový takt) – súbor C verzia 2: vizualizácia súboru dát v tabuľke 8, pomocou programu Microsoft Excel.



Graf 2.C2r: U2 (výpoveď vs. prízvukový takt)– súbor C verzia 2: vizualizácia súboru dát v tabuľke 8 pomocou programu R.

Na grafoch 2.C1 a 2.C2 je možné vidieť mierne klesavú tendenciu, prvým predpokladom by mohlo byť, že zhoda s MAL nebude nijak výrazná. Určitý pokles je síce očividný, ale pomyselná spojnica jednotlivých bodov nevykazuje prudký pokles. Tento predpoklad je však okamžite po prezretí si grafu 2.C1r a 2.C2r vyvrátený. Na príslušných grafoch, kde dochádza k preloženiu konkrétnych pozorovaní reprezentovaných pomocou jednotlivých bodov krivkou, ktorá reprezentuje MAL, je okamžite pozorovateľná prudká klesajúca tendencia. Tento dojem ešte umocňuje číselné vyjadrenie zhody s MAL, pričom v prípade 1. verzie výberového súboru C činí táto zhoda  $R^2 = 78,72 \%$ , čo sa v lingvistických kruhoch považuje za dostatočne vysokú zhodu s matematickým modelom. Avšak v prípade verzie 2 tohto istého súboru je toto číslo ešte vyššie a percentuálne činí zhodu až  $R^2 = 94,57\%$ . To znamená, že v rámci vzťahu konštrukt a konštituentu na U2 sa vzťah daný MAL preukázal vo vysokej percentuálnej zhode. V nasledujúcej tabuľke č. 9 uvádzam pre lepší prehľad číselné hodnoty parametra b, rovnako ako percentuálnu zhodu jednotlivých výberových súborov a ich percentuálnej zhody s MAL.

Tabuľka 9: Prehľad parametrov <sup>b</sup> MAL a koeficientov determinácie  $R^2$  na U2

		Parameter <sup>b</sup>	Koeficient determinácie $R^2$ (%)
Súbor A		0.2982	73.11
Súbor B	Verzia 1	0.2737	22.38
	Verzia 2	0.0761	34.78
Súbor C	Verzia 1	0.1226	78.72
	Verzia 2	0.1755	94.57

V priebehu výskumu sa preukázalo, že v prípade výberového súboru C je segmentácia prof. Švarného presnejšia, nakoľko sa jedná o pomalšiu reč, a prízvukový takt definovaný v prvej časti výskumu bol dĺžkou ustálený na pomedzí segmentu a kolónu. Aj napriek tomu, vykazujú na danej úrovni obe verzie vysokú zhodu s MAL. Percentuálne obe verzie výberového súboru C súhlasia s výberovým súborom A z prvej fáze výskumu, ktorý taktiež na danej úrovni dosiahol vysokú zhodu s MAL. Navyše, táto zhoda sa pri odobratí málo frekventovaných bodov ešte výrazne zvýšila. To by mohlo znamenať, že vzťah konštrukt – konštituent na tejto úrovni podlieha pravidlám

MALu. Avšak výberový súbor B vykazuje v oboch prípadoch priveľmi nízke zhody na to, aby spĺňal podmienky zhody s MALom.

Dôvodov odlišnosti tohto vzorku môže byť niekoľko. Na základe výsledkov uskutočneného výskumu sa javí, že rytmus hudobného žánru, ovplyvňuje prirodzenú rytmiku reči. V prípade popovej speváčky, ako aj hudobne nepríznačného súboru C sa zhoda s MAL výrazne potvrdila. Na rozdiel od týchto výsledkov však vzorka produkovaná raperom zákonitosti MALu tak výrazne nespĺňa. Samozrejme, na úplné potvrdenie tejto hypotézy plánujem realizovať výskum na väčšom množstve vzoriek a takisto analyzovať väčší počet rapových interpretov. Na základe doposiaľ získaných výsledkov je však možné tvrdiť, že pomerne umelá rytmika rapu, dokáže ovplyvniť prirodzenú rytmiku reči u daného interpreta.

Ďalším dôvodom pre takéto rozdiely medzi výberovými súbormi mohla spôsobiť aj pomerne výrazná emocionalita u interpreta výberového súboru B. Na rozdiel od popovej speváčky u súboru A a Číňanky, ktorá vytvorila zvukový záznam vzorky C, muzikant zo vzorky B prejavuje výraznú a často až prehnanú emocionalitu a rozrušenie vo svojom prejave. Tento fakt taktiež mohol narušiť prirodzenú rytmiku reči, nakoľko vo výrazne emocionálne zafarbených pasážach sa tento subjekt pomerne často zasekáva, nepravidelne zrýchľuje a spomaľuje tok svojej reči a tým narúša neutrálnosť a pravidelnosť prejavu.

V priebehu tohto výskumu sa podarilo potvrdiť, že vzťah konštruktú a konštituentu daný MALom v rámci úrovne, ktorá meria vzájomný vzťah výpovede a prízvukového taktu meraného v slabikách, platí a taktiež je možné tvrdiť, že sa potvrdila aj hypotéza, že rytmus hudobného žánru môže ovplyvniť prirodzenú rytmiku reči daného interpreta. Aj napriek tomu, by však bolo vhodné pristúpiť k ďalšiemu výskumu a obe tieto hypotézy podrobiť detailnejšiemu výskumu.

### 5.3 U3: Prízvukový takt meraný v slabikách, ktoré sú merané v priemernom počte ich hlások.

Na tejto úrovni bol skúmaný vzájomný vzťah konštrukt, t.j. prízvukového taktu a konštituentu t.j. slabiky, ktorá vstupovala zároveň ako konštrukt voči svojmu konštituentovi, t.j. hláske. Prízvukový takt je základnou úrovňou celého experimentu, nakoľko sa vyskytuje v skúmaníach na všetkých troch úrovniach, jediné čo sa mení, je jeho vzťah voči ostatným jednotkám na danej úrovni. Na U1 prízvukový takt vystupuje len ako konštituent voči výpovedi, ktorá je konštituentom repliky, na U2 má prízvukový takt dve funkcie a to ako konštrukt meraný v slabikách, tak aj konštituent k vyššiemu konštruktovi a teda k výpovedi. Na U3 prízvukový takt zas vystupuje len ako konštrukt meraný v slabikách, ktoré sú merané v priemernom počte hlások. Ako som spomínala v teoretickej časti, dĺžka prízvukového taktu v čínštine môže rôzne variovať v závislosti od rýchlosti reči interpreta. V nasledujúcej časti sa pozrieme na to, či výberové súbory A a B na tejto úrovni zodpovedajú zákonitostiam a pravidlám MALu a rovnako sa pozrieme na odlišnosti jednotlivých súborov.

Tabuľka 10: U3: konštrukt  $x_3$  – dĺžka prízvukového taktu (v počte jeho slabík),  $z_3$  – frekvencia konštruktov, konštituenty  $y_3$  – priemerná dĺžka slabiky (v počte jej hlások), výberový súbor A.

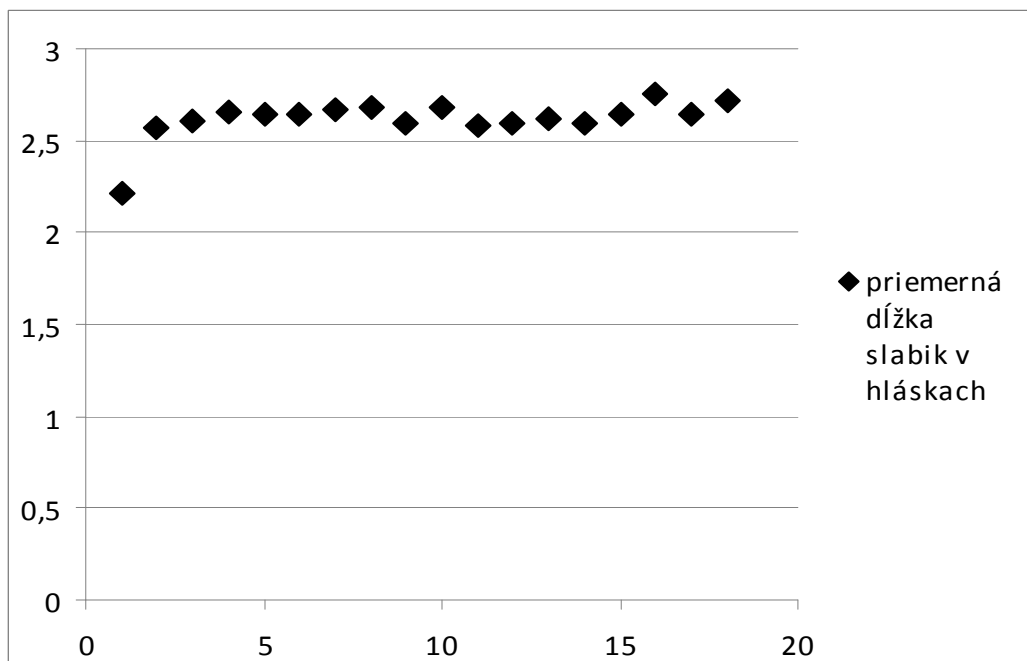
	Súbor A	
$x^3$	$z^3$	$y^3$
1	19	2,2105
2	76	2,5658
3	73	2,6073
4	69	2,6522
5	84	2,6452
6	65	2,6462
7	48	2,6667
8	36	2,6840
9	33	2,5926
10	24	2,6792
11	9	2,5859
12	8	2,5938
13	9	2,6239
14	7	2,5918

15	3	2,6444
16	2	2,7500
17	2	2,6471
18	1	2,7222

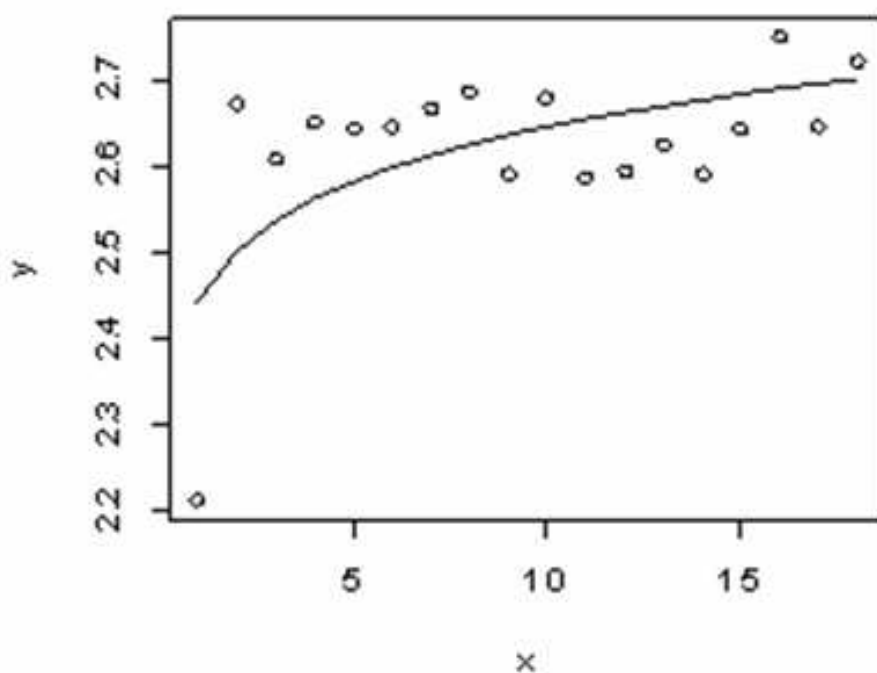
Tabuľka 11: U3: konštrukt  $x_3$  – dĺžka prízvukového taktu (v počte jeho slabík),  $z_3$  – frekvencia konštruktov, konštituenty  $y_3$  – priemerná dĺžka slabiky (v počte jej hlások), výberový súbor B.

$x^3$	Súbor B	
	$z^3$	$y^3$
1	10	3,6000
2	82	2,7500
3	47	2,7163
4	63	2,6032
5	48	2,7250
6	37	2,6036
7	25	2,6400
8	20	2,6313
9	8	2,6528
10	3	2,7000
11	2	2,3182
13	2	2,7308

Z oboch tabuliek 5 a 6 je vidieť, že dĺžka prízvukového taktu výrazne variuje v rozmedzí od 1 až po 18. Samozrejme frekvencia prízvukových taktov s dĺžkou väčšou ako 10 rapídne klesá z desiatok na jednotky, nakoľko takáto dĺžka je v čínštine pomerne zriedkavá. Avšak ako som už spomínala, dĺžka prízvukové taktu resp. segmentu podľa názvoslovia prof. Švarného je závislá na tepme reči, čo znamená že dĺžka prízvukového taktu je v rýchlejšej reči dlhšia na rozdiel od reči pomalšej. Vzhľadom na to, že čo sa týka tejto úrovne, boli výberové súbory dostatočne obsiahle, je možno získať relevantnejšie údaje, ako je to na najvyššej úrovni, na ktorej bola frekvencia jednotlivých konštruktov pomerne nedostatočná. Nasledujúce grafy rovnako zoradujem podľa súboru a typu grafu, kvôli porovnaniu grafu z programu Microsoft Excel a programu R a po predstavení grafov z oboch súborov pristúpim k interpretácii daných výsledkov a ich porovnaniu.

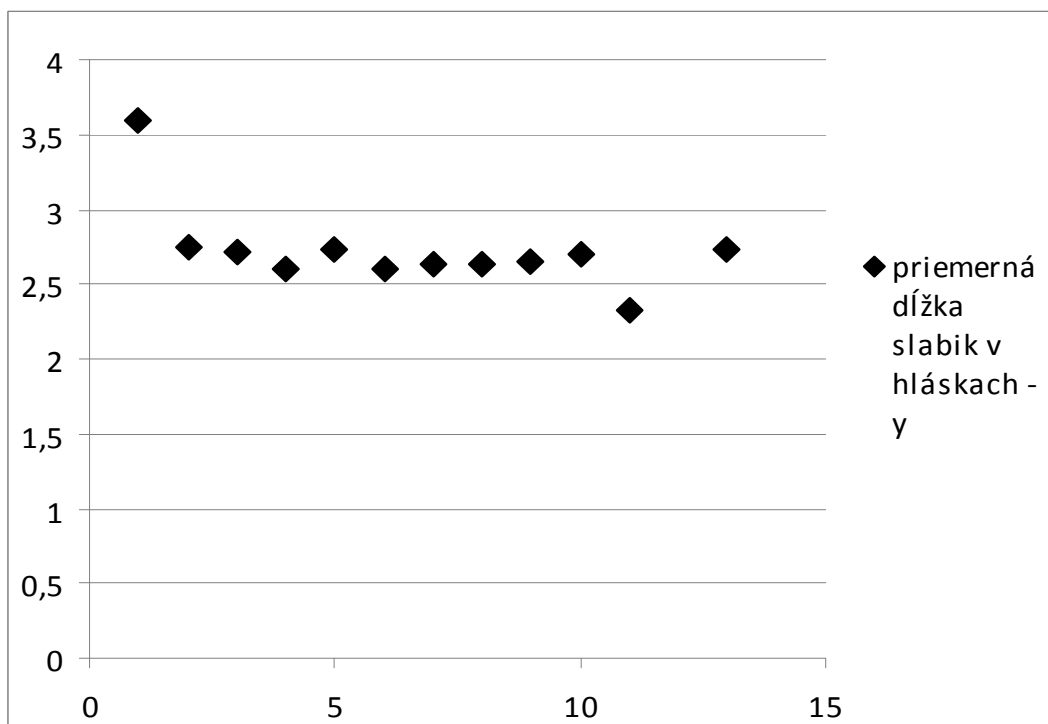


Graf 3.A: U2 (prízvukový takt vs. slabika) – súbor A: vizualizácia súboru dát v tabuľke 10, pomocou programu Microsoft Excel.

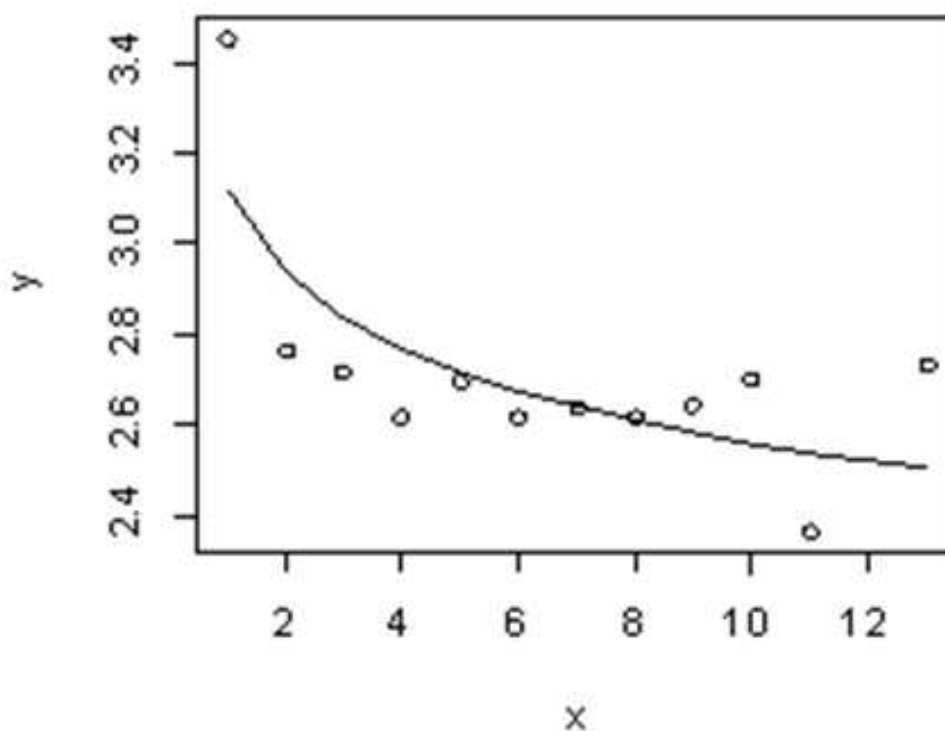


Graf 3.Ar: U3 (prízvukový takt vs. slabika)<sup>114</sup> – súbor A: vizualizácia súboru dát v tabuľke 10, pomocou programu R.

<sup>114</sup> SCHUSTEROVÁ, s. 13.



Graf 3.B: U3 (prízvukový takt vs. slabika) – súbor B: vizualizácia súboru dát v tabuľke 11, pomocou programu Microsoft Excel.



Graf 3.Br: U3 (prízvukový takt vs. slabika)<sup>115</sup> – súbor B: vizualizácia súboru dát v tabuľke 11, pomocou programu R.

<sup>115</sup> SCHUSTEROVÁ, s. 13.

Z vyššie uvedených grafov 3.A a 3.B je výrazne viditeľná pomerne konštantná tendencia jednotlivých bodov, ktoré sa blízko zoskupujú okolo hodnoty 2,5 na osi y. V oboch prípadoch je väčšina bodov začlenená relatívne jednotne na osi y, výrazný rozdiel je možné objaviť v prípade bodov  $x_3 = 1$ . Táto odlišnosť je pravdepodobne spôsobená individualitou jednotlivých rečníkov v daných prípadoch. V prípade súboru A sa popova speváčka pomerne často zasekáva a vytvára jednoslabičné prízvukové takty, ktoré obsahujú nízky počet hlások, prípadne pri vymenúvaní viacnásobných vetných členov sa taktiež objavujú veľmi krátke jednoslabičné takty s minimálnym počtom hlások. Táto anomália spôsobila, že na danej osi bod  $x_3 = 1$  vykazuje nižší priemerný počet hlások, ako je to u ostatných prízvukových taktov s dĺžkou viac ako jedna slabika.

V prípade súboru B sa taktiež objavuje anomália a vybočenie v rámci toho istého bodu  $x_3 = 1$ , avšak v tomto prípade ide o presne opačný fenomén. V tomto prípade daný bod  $x_3 = 1$  vykazuje vyšší priemerný počet hlások, ako je to u ostatných prízvukových taktov s väčším počtom slabík. Táto odlišnosť bola pravdepodobne spôsobená osobitým rečovým prejavom rapového interpreta, ktorý pomerne často používal anglické slová. Nakoľko mal však určité problémy s výslovnosťou a veľmi sa na ňu sústredil, tieto takty sa výrazne skrátili oproti ostatným prízvukovým taktom, to znamená, že tento interpret jednotlivé anglické slová v texte výrazne vyčlenil ako samostatné jednoslabičné prízvukové takty. Anglické slová, často aj tie jednoslabičné, sú zložené z viacerých hlások, ktorých môže byť viac ako v priemernej čínskej slabike, a táto rozdielnosť mohla zapríčiniť výkyv v bode  $x_3 = 1$ . Ak k tomu pridáme výrazné skrátenie jednoslabičného prízvukového taktu v súbore A zistíme, že tým pádom sa táto odlišnosť medzi oboma súbormi stráca.

Ostatné body oscilujú nad hodnotou 2,5 a s výnimkou  $x_3 = 1$  v oboch súboroch a  $x_3 = 11$  v prípade súboru B, kde môže hrať úlohu nedostatočná frekvencia. Všetky ostatné body sa nachádzajú v intervale (2.56; 2.72) v prípade súboru A a v intervale (2.60; 2.75) v prípade súboru B. Je vidieť, že hodnoty intervalov sú veľmi blízke a variujú len v extrémne malom rozmedzí. Táto skutočnosť môže poukazovať na systémovosť čínštiny, vzhľadom na konštantnú tendenciu zobrazenú na grafoch 3.A



a 3.B. Toto zistenie je veľmi dôležité, nakoľko to môže poukazovať na odlišné ponímanie najmenej jazykovej jednotky v čínštine.

Pri pohľade na grafy 3.Ar a 3.Br si môžeme všimnúť, že krivka v prípade súboru A má stúpajúcu tendenciu, ktorá je spôsobená bodom  $x_3 = 1$ . Tento bod je položený nižšie ako tie ostatné a preto výrazne ovplyvňuje priebeh krivky. Na rozdiel od stúpajúcej krivky v grafe 3.Ar krivka v grafe 3.Br má tendenciu klesajúcu, avšak aj táto tendencia je znovu ovplyvnená prvým bodom diagramu. Zhoda s MAL je v prípade súboru A 37.53% a v prípade súboru B 56.65%. Obe krivky však majú najprudší priebeh na začiatku, to znamená v okolí bodu  $x_3 = 1$ , pri ďalšom pokračovaní krivky sa jej priebeh rapídne zmierňuje. Pre lepší prehľad a porovnanie číselných hodnôt najmä parametra  $b$  a percentuálnej zhody znovu poskytujem nižšie uvedenú tabuľku s príslušnými dátami.

Tabuľka 12: Prehľad parametrov <sup>b</sup> MAL a koeficientov determinácie  $R^2$  na U3<sup>116</sup>

	<b>Parameter <sup>b</sup></b>	<b>Koeficient determinácie <math>R^2</math> (%)</b>
Súbor A	- 0.0349	37.53
Súbor B	0.0855	56.65

Aj keď na základe číselných údajov je možné tvrdiť, že v prípade súboru B sa prejavila pomerne vysoká zhoda s MAL, v grafoch 3.A a 3.B., kde sa nevyskytuje krivka, je možné sledovať viac-menej lineárnu tendenciu čínskeho jazyka na úrovni priemerného počtu hlások v rámci slabiky. Táto tendencia sa ukazuje ako dôležitý bod ďalšieho skúmania a na jej potvrdenie je nutné vykonať ďalšie experimenty vo väčšom meradle.<sup>117</sup>

---

<sup>116</sup> SCHUSTEROVÁ, s. 14.

<sup>117</sup>SCHUSTEROVÁ, Denisa.

## Záver

Cieľom tejto práce bolo preskúmať platnosť Menzerath-Altmannovho zákona na hovorenú čínštinu a potvrdiť resp. vyvrátiť hypotézy s ním súvisiace. Stanovená hypotéza predpokladala, že bude možné preukázať platnosť MALu na hovorenej čínštine na všetkých troch úrovniach. Zároveň úlohou druhej hypotézy bolo overiť, nakoľko dokáže umelá rytmika hudobného štýlu ovplyvniť prirodzenú rytmiku reči. V prvej fáze boli tieto hypotézy overované kvantitatívnou metódou na dvoch výberových súboroch A a B, ktoré oba pochádzali z hudobného programu určeného pre zábavu diváka. Na týchto dvoch súboroch sa testovala hypotéza zhody hovorenej čínštiny s MALom v rámci vzájomných vzťahov jazykových jednotiek na troch úrovniach. V druhej fáze výskumu bol neskôr pribraný tretí kontrastný súbor C, ktorý bol pomerne umelo vytvorený pre účely výučby čínštiny pre cudzincov. Tento súbor slúžil ako kontrast k prvým dvom súborom hlavne v rámci vzťahov na druhej úrovni, čo znamená vzťah výpovede a prízvukového taktu meraného v priemernom počte jeho slabík.

Po stanovení podmienok a vybratí výberových súborov sme pristúpili k ďalšej fáze výskumu, v ktorej bolo nutné vybrať vhodnú transkripciu na prepis hovoreného textu pre uľahčenie následného kvantitatívneho spracovania dát. Rovnako dôležité bolo stanoviť a definovať jednotlivé jazykové jednotky nevyhnutné pre ďalší priebeh experimentu. Definovanie týchto textových jednotiek bolo dôležité hlavne kvôli nasledujúcej fáze, v ktorej bolo nutné výberové súbory správne nasegmentovať na nižšie jednotky, aby tieto jednotky mohli byť dané do vzťahu konštrukt – konštituent. Po segmentácii sme pristúpili k samotnému jadru výskumu, ktorým bola aplikácia vzorcov Menzerath-Altmannovho zákona na získané dáta. V tomto momente vstúpil výskum do svojej druhej fázy, v ktorej bol pribraný kontrastný súbor C pre lepšie porovnanie so súbormi A a B a pre overenie hypotézy testujúcej vplyv umelej rytmiky hudobného žánru na prirodzenú rytmiku reči. Posledným krokom výskumu bolo spracovanie a analýza získaných výsledkov a ich následná interpretácia.

Na základe získaných výsledkov sa na najvyššej úrovni, to znamená replika meraná vo výpovedi meranej v priemernom počte jej prízvukových taktov, zhoda s MAL

nepotvrdila. Táto skutočnosť ale mohla byť ovplyvnená viacerými faktormi, medzi ktorými môže figurovať nedostatočná dĺžka vzorku, čo je častým problémom pri skúmaní vyšších jazykových úrovní. Ďalšou príčinou, ktorá zapríčinila nezhodu s pôvodnou hypotézou môže byť chýbajúca medzi-úroveň medzi výpoveďou a samotnou replikou.

Na úrovni, ktorá spája do vzťahu výpoveď a prízvukový takt meraný v priemernom počte jeho slabík a ktorý je hlavným záujmom môjho výskumu, sa na výberovom súbore A a C prejavila vysoko percentná zhoda s MAL, čo zodpovedá predpokladu pôvodnej hypotézy. Výberový súbor B, ktorý sa výsledkami od týchto dvoch súborov pomerne výrazne líši, síce pôvodnú hypotézu zhody s MALom vyvracia, ale zároveň potvrdzuje druhú hypotézu, ktorá testovala vplyv umelej rytmiky na prirodzenú rytmiku reči a stanovovala, že umelý rytmus dokáže ovplyvniť prirodzený rytmus reči. Výberový súbor B, na rozdiel od súborov A a C produkovaných popovou speváčkou a nespeváckym subjektom, bol vyprodukovaný rapovým interpretom. Na základe získaných výsledkov sa javí, že umelý rytmus tohto hudobného štýlu je natoľko zaväzujúci, že dokáže ovplyvniť aj prirodzenú rytmiku reči. Pri tomto výberovom súbore som sa na základe rečového prejavu, v ktorom sa objavovalo veľké množstvo vulgarizmov, rozhodla pristúpiť k dvom odlišným verziám segmentácie. Avšak obe verzie nepotvrdili výraznú zhodu s MAL, preto inklinujem k tvrdeniu, že rapový rytmus dokázal ovplyvniť prirodzenú rytmiku reči daného interpreta. Do úvahy však treba vziať aj miestami prílišnú emocionalitu rečníka, ktorá ovplyvňovala tempo reči a tým pádom sa na výsledkoch taktiež mohla podpísať. Rozhodne je však nutné realizovať ďalšie podrobnejšie skúmanie na viacerých výberových súboroch, aby bolo skutočne možné danú hypotézu definitívne potvrdiť.

Na najnižšej úrovni, na ktorej bol skúmaný prízvukový takt meraný v slabikách meraných v priemernom počte ich hlások sa objavil zaujímavý fenomén. Výsledky ukázali, že v prípade oboch výberových súborov jednotlivé pozorovania oscillovali v rámci určitého intervalu, čo by mohlo naznačovať linearitu a pravidelnosť vo vzťahu slabiky a hlásky. V rámci tohto výskumu sa medzi jednotkami na tejto najnižšej úrovni nepreukázala žiadna výrazná zhoda s MAL, na základe čoho bola v rámci U1 pôvodná hypotéza vyvrátená. Rovnako ako v predchádzajúcich prípadoch je ale nutné, aby boli ďalšie otázky, ktoré sa objavili v priebehu výskumu, preverené na väčšom množstve

výberových súborov a aby boli novo vzniknuté hypotézy overené na základe ďalších pozorovaní. V rámci tohto výskumu sa teda pôvodná hypotéza stanovujúca podliehanie čínštiny zákonitostiam MALu potvrdila len na úrovni U2. Zároveň sa na tejto úrovni prejavila tendencia umelého rytmu ovplyvňovať ten prirodzený, čo však samozrejme nabáda k ďalšiemu výskumu pre lepšie a detailnejšie pochopenie a preskúmanie daného fenoménu.

## **Resumé**

The aim of this paper was to verify the validity of the Menzerath-Altmann Law applied to contemporary spoken Chinese. The research described in this paper investigated the mutual relationship between the predefined language units, with an emphasis on the relationship between a statement and a stress unit which was measured in syllables. The original hypothesis stated that Chinese language would be subject to the patterns of the MAL in every mutual relationship on all the three tested language levels. In addition, the second hypothesis stated that the rhythm of an artificial music genre would have an influence on the natural rhythm of speech. This hypothesis was tested within the mutual relationship of the statement and the stress unit. Based on this research it was discovered that the first hypothesis was not entirely correct as the acquired results demonstrated that the Chinese language was only subject to the patterns of the MAL within the relationship between the statement and the stress unit. Nevertheless, the second hypothesis seems to be valid which means that the rhythm of the artificial music genre influences the natural rhythm of speech.

## Zoznam literatúry

ALTMANN, Gabriel. **Prolegomena to Menzerath's law.** *Glottometrika*, č.2, 1980. s. 1–10.

ANDRES, Jan et al. **Optimization of parameters in the Menzerath-Altmann law.** *Mathematica*, 2012, roč. 51, č. 1, s. 5-27.

ANDRES, Jan. BENEŠOVÁ, Martina et al. **Methodological Note on the Fractal Analysis of Texts.** *Journal of Quantitative Linguistics*, [databáza online]. 2012, roč. 19, č. 1, s. 1-31. [cit. 2012-03-21]. Dostupné z: <<http://dx.doi.org/10.1080/09296174.2011.608604>>.

ANDRES, Jan. **On a Conjecture about the Fractal Structure of Language.** *Journal of Quantitative Linguistics*, [databáza online]. 2010, roč.17, č. 2, s.101-122. [cit. 2013-03-21]. Dostupné z: <<http://www.tandfonline.com/doi/abs/10.1080/09296171003643189>>.

BARTHES, Roland. **Úvod do strukturální analýzy vyprávění.** *Znak, struktura, vyprávění.* Brno: Host, 2002, s. 9-43.

BEDNAŘÍKOVÁ, Božena. *Slovo a jeho konverze.* Olomouc: Univerzita Palackého v Olomouci, 2009. ISBN 978-80-244-2220-6.

BENEŠOVÁ, Martina. *Kvantitativní analýza textu se zvláštním zřetelem k analýze fraktální.* Olomouc, 2011. disertační práce (Ph.D.). UNIVERZITA PALACKÉHO V OLOMOUCI. Filozofická fakulta.

CAO, Jianfen. **Rhythm of Spoken Chinese.** [cit. 2012-04-06]. Dostupné z: <[http://ling.cass.cn/yuyin/report/files/2000\\_4.pdf](http://ling.cass.cn/yuyin/report/files/2000_4.pdf)>.

CAO, Jianfen. **The Rhythm of Mandarin Chinese.** [cit. 2012-04-08]. Dostupné z: <[http://www.researchgate.net/publication/228722757\\_The\\_rhythm\\_of\\_Mandarin\\_Chinese?citationList=incoming](http://www.researchgate.net/publication/228722757_The_rhythm_of_Mandarin_Chinese?citationList=incoming)>.

CRAMER, Irene. **The Parameters of the Altmann-Menzerath Law.** *Journal of Quantitative Linguistics*, [online] 2005, roč. 12, č. 1. s. 41-52. [cit. 2012-03-28]. Dostupné z: <<http://www.tandfonline.com/doi/pdf/10.1080/09296170500055301>>.

ČERNÝ, Jiří. *Dějiny lingvistiky.* Olomouc: Votobia, 1996. ISBN 80-85885-96-4.

ČERNÝ, Jiří. *Stručné dějiny lingvistiky III.* Olomouc: Univerzita Palackého, 1981.

DANEŠ, František a spol. *O mluveném slově.* Praha: Státní pedagogické nakladatelství, 1954.

DANEŠ, František. *Jazyk a text I : výbor z lingvistického díla Františka Daneše. Část I.* Praha : Univerzita Karlova, 1999. ISBN 8085899701.

DANEŠ, František. *Jazyk a text I : výbor z lingvistického díla Františka Daneše. Část 2.* Praha : Univerzita Karlova, 1999. s. 249-471. ISBN 8085899701.

DANEŠ, František. **Příspěvek k rozboru významové výstavby výpovědi.** *Studie a práce lingvistické. I.* 1. vyd. Praha : Nakladatelství Československé akademie věd, 1954.

DANEŠ, František. *Věta a text.* Praha: Academia, 1985.

DOKOUPIL, Martin. *Jazykověda.* Vrbátky: M. Dokoupil, 2012. ISBN 978-80-260-2129-2.

HRBÁČEK, Josef. **Promluva.** *Naše řeč*, [online] 2003, roč. 86, č. 1. [cit. 2012-03-28]. Dostupné z: <<http://nase-rec.ujc.cas.cz/archiv.php?art=7705>>.

HŘEBÍČEK, Luděk. **Filologie versus lingvistika.** *Vesmír* [online] 2008, roč. 87, č. 7. [cit. 2012-04-15]. Dostupné z: <<http://www.vesmir.cz/clanek/filologie-versus-lingvistika>>.

HŘEBÍČEK, Luděk. *Lectures on text theory.* Prague: Oriental Institute, 1997. ISBN 80-85425-26-2.

HŘEBÍČEK, Luděk. *Vyprávění o lingvistických experimentech s textem.* Praha: Academia, 2002. ISBN 80-200-0973-6.

HŘEBÍČEK, Luděk. *Text in Semantics: The Principle of Compositenes.* Prague: Academy of Sciences of Czech Republic, Oriental Institute, 2007.

CHOW, Ivan. **Research in the Prosodic Structure of Mandarin.** [cit. 2012-04-08]. Dostupné z: <<http://www.semioticon.com/virtuals/talks/chow.htm>>.

JAKLOVÁ, Alena. *Úvod do textové lingvistiky.* České Budějovice: Pedagogická fakulta, 1991. ISBN 80-7040-035-8.

*Jazykověda = Linguistica.* Ostrava: Ostravská univerzita, 2004 . Sborník prací Filozofické fakulty Ostravské univerzity. ISBN 80-7042-664-0.

KANE, Daniel. *Knížka o čínštině.* Mirošovice: DesertRose, 2009. ISBN 978-80-903296-1-4.

KRČMOVÁ, Marie: *Fonetika* [online] Brno: Filosofická fakulta MU, 2007. [cit. 2013-04-10] Dostupné z: <[is.muni.cz/elportal/estud/ff/js07/fonetika/materialy/ch04.html](http://is.muni.cz/elportal/estud/ff/js07/fonetika/materialy/ch04.html)>.

KUŁACKA, Agnieszka. **The Coefficients in the Formula for the Menzerath-Altman Law.** *Journal of Quantitative Linguistics*, [databáza online]. 2010. roč. 17, č. 4, s. 257-268. [cit. 2012-03-27]. Dostupné z: <<http://dx.doi.org/10.1080/09296174.2010.512160>>.

LIN, Yen-Hwei. *The Sounds of Chinese.* Cambridge: Cambridge University Press, 2007, 312 p. ISBN: 978-0-521-60398-0.

- MATHESIUS, Vilém. *Jazyk, kultura a slovesnost*. Praha: Odeon, 1982.
- MÜLLEROVÁ, Olga. *Mluvený text a jeho syntaktická výstavba*. Praha: Academia, 1994. ISBN 80-200-0489-0.
- PACKARD, Jerome L. *The morphology of Chinese: a linguistic and cognitive approach*. Cambridge: Cambridge University Press, 2006. ISBN 978-052-1026-109.
- PACKARD, Jerome Lee. *New approaches to Chinese word formation: morphology, phonology and the lexicon in modern and ancient Chinese*. New York: Mouton de Gruyter, 1998. ISBN 31-101-5109-X.
- PING, Chen. *Modern Chinese*. Cambridge: Cambridge University Press, 1999.
- SCHUSTEROVÁ, Denisa. *Aplikace Menzerath-Altmanova zákona na mluvenou čínštinu - přízvukový takt*. Olomouc, 2013. diplomová práce (Mgr.) UNIVERZITA PALACKÉHO V OLOMOUCI. Filozofická fakulta.
- SCHUSTEROVÁ, Denisa. ŠČIGULINSKÁ Jana et al. **An Application of the Menzerath-Altman Law to Contemporary Spoken Chinese**. *Czech and Slovak Linguistic Review*, 2013, č. 2, (v tlači).
- ŠIMKOVÁ, Mária. Výberový slovník termínov z korpusovej lingvistiky. [2013-03-04] Dostupné z: <http://korpus.juls.savba.sk/attachments/what/2006-simkova-vyberovy%20slovník%20terminov.pdf>
- ŠVARNÝ, Oldřich a UHER, David. *Hovorová čínština: úvod do studia hovorové čínštiny*. Olomouc: Univerzita Palackého, 2001. ISBN 80-244-0298-X.
- ŠVARNÝ, Oldřich a kol. *Gramatika hovorovej čínštiny v príkladoch*. Bratislava: Vydavateľstvo Univerzity Komenského, 1993.
- ŠVARNÝ, Oldřich a kol. *Hovorová čínština v príkladech 3*. Olomouc: Vydavatelství Univerzity Palackého, 1998.
- TĚŠITELOVÁ, Marie. *Kvantitativní lingvistika*. Praha: Univerzita Karlova, 1987.
- TŘÍSKOVÁ, Hana. *Transkripce čínštiny: Tabulky a návody*. Praha: Česko-čínská společnost, 1999. ISBN 80-902515-3-6.
- WIMMER, Gejza. *Úvod do analýzy textov*. Bratislava: Veda, 2003. ISBN 80-224-0756-9.
- XING, Janet Zhiqun. *Studies of Chinese linguistics: functional approaches*. Hong Kong: Hong Kong University Press, 2009. ISBN 96-220-9964-5.