

UNIVERZITA PALACKÉHO V OLOMOUCI

FILOZOFICKÁ FAKULTA

Katedra asijských studií

MAGISTERSKÁ DIPLOMOVÁ PRÁCE

**APLIKACE MENZERATH-ALTMANNOVA ZÁKONA NA MLUVENOU
ČÍNŠTINU – PŘÍZVUKOVÝ TAKT**

**APPLYING OF THE MENZERATH-ALTMANN LAW TO THE SPOKEN
CHINESE – STRESS UNIT**

Vypracovala: Bc. Denisa Schusterová

Vedoucí práce: Mgr. Ondřej Kučera

Olomouc 2013

Prohlašuji, že jsem magisterskou diplomovou prací na téma *Aplikace Menzerath-
Altmannova zákona na mluvenou čínštinu – přízvukový takt* vypracovala samostatně
za použití uvedených pramenů a literatury.

Děkuji Mgr. Ondřeji Kučerovi za odborné vedení mé práce a cenné připomínky, mé vřelé poděkování patří také Mgr. Martině Benešové Ph.D. za konzultace v oblasti matematiky a lingvistiky.

ANOTACE

Cílem této diplomové práce je aplikovat a ověřit platnost Menzerath-Altmanova zákona na současnou mluvenou čínštinu se zvláštním zřetelem na nejnižší jazykovou úroveň. Výzkum probíhal ve dvou fázích. V první fázi byl MAL aplikován na dva výběrové soubory. Zkoumané jazykové úrovně v první fázi výzkumu byly tyto: jazyková úroveň tvořená replikou (konstrukt, jehož délka je měřena v počtu výpovědí) a výpovědí (konstituent, jehož délka je měřena v průměrném počtu přízvukových taktů), jazyková úroveň tvořená výpovědí (konstrukt, jehož délka je měřena v počtu přízvukových taktů) a přízvukovým taktem (konstituent, jehož délka je měřena v průměrném počtu slabik) a jazyková úroveň tvořená přízvukovým taktem (konstrukt, jehož délka je měřena v počtu slabik) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek). Výsledky z první fáze výzkumu platnost MALu na nejnižší jazykové úrovni nepotvrdily.

Ve druhé fázi výzkumu je pozornost věnována nejnižší jazykové úrovni, kde konstruktem je tentokrát jazyková jednotka segment. MAL je v této fázi aplikován na třetí výběrový soubor, který je od předešlých dvou zcela odlišný.

Práce je rozdělena na teoretickou a praktickou část. V teoretické části je popsán celý průběh výzkumu, od stanovení hypotézy, přes volbu výběrových souborů, jejich přepis do psané podoby, stanovení zkoumaných jazykových jednotek a popis problémů, které jejich vymezení představovalo. V praktické části jsou pak prezentovány výsledky v pořadí od nejvyšší po nejnižší jazykovou úroveň. Dosažené výsledky jsou demonstrovány v grafech a tabulkách, pod nimi vždy následuje diskuse.

Klíčová slova: Menzerath-Altmanův zákon, mluvená čínština, kvantitativní lingvistika, jazykové úrovně, segmentace

OBSAH

ANOTACE.....	6
OBSAH.....	7
SEZNAM TABULEK.....	9
SEZNAM GRAFŮ.....	11
EDIČNÍ POZNÁMKA	13
ÚVOD.....	14
1. LINGVISTIKA	17
1.1 Strukturalismus.....	17
1.2 Matematická lingvistika	18
1.2.1 Kvantitativní lingvistika	18
1.2.2 Algebraická lingvistika.....	20
2. MENZERATH-ALTMANNŮV ZÁKON.....	21
3. FONETICKÝ SYSTÉM ČINŠTINY.....	24
4. POPIS VÝBĚROVÝCH SOUBORŮ.....	26
4.1 Výběrové soubory A a B	26
4.2 Výběrový soubor C	27
5. DEFINOVÁNÍ JAZYKOVÝCH JEDNOTEK.....	28
5.1 Výběrové soubory A a B	28
5.1.1 Replika.....	29
5.1.2 Výpověď.....	30
5.1.3 Přízvukový takt.....	31
5.1.4 Slabika	32
5.1.5 Hláška	33
5.2 Výběrový soubor C	33
5.2.1 Kólon	34
5.2.2 Segment	34
6. JAZYKOVÉ ÚROVNĚ.....	36
7. PŘEPIS VÝBĚROVÝCH SOUBORŮ.....	37

7.1 Kontrast mezi zhe (če), she (še) vs. zhi (č'), shi (š')	37
7.2 Souhlásky neaspirované a aspirované	37
7.3 Frikativy	38
7.4 Iniciála q	38
7.5 Samohlásky	38
8. PREZENTACE VÝSLEDKŮ	40
8.1 Jazyková úroveň U1: replika (konstrukt, jehož délka je měřena v počtu výpovědí) a výpověď (konstituent, jehož délka je měřena v průměrném počtu přízvukových taktů).....	41
8.2 Jazyková úroveň U2: výpověď (konstrukt, jehož délka je měřena v počtu přízvukových taktů) a přízvukový takt (konstituent, jehož délka je měřena v průměrném počtu slabik).....	46
8.3 Jazyková úroveň U3: přízvukový takt (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek)	52
8.4 Jazyková úroveň U3Š: segment (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek) 58	
8.5 Jazyková úroveň U3/U3Š: přízvukový takt/segment (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek) – odebrání některých pozorování.....	62
8.6 Jazyková úroveň U2Š-U1Š: kólon (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek) 65	
9. ZÁVĚR	68
10. RESUMÉ	72
11. KNIŽNÍ A INTERNETOVÉ ZDROJE	73

SEZNAM TABULEK

Tabulka 1: Přehled foneticky problematických aspektů	39
Tabulka 2: U1, výběrový soubor A: konstrukty x_1 – délka repliky (měřena v počtu výpovědí), z_1 – frekvence konstruktů, konstituenty y_1 – průměrná délka výpovědi (měřena v počtu přízvukových taktů)	41
Tabulka 3: U1, výběrový soubor B, verze 1: konstrukty x_1 – délka repliky (měřena v počtu výpovědí), z_1 – frekvence konstruktů, konstituenty y_1 – průměrná délka výpovědi (měřena v počtu přízvukových taktů).....	42
Tabulka 4: U1, výběrový soubor B, verze 2: konstrukty x_1 – délka repliky (měřena v počtu výpovědí), z_1 – frekvence konstruktů, konstituenty y_1 – průměrná délka výpovědi (měřena v počtu přízvukových taktů).....	43
Tabulka 5: Přehled parametrů b MAL a koeficientů determinace na U1	45
Tabulka 6: U2, výběrový soubor A: konstrukty x_2 – délka výpovědi (měřená v počtu přízvukových taktů), z_2 – frekvence konstruktů, konstituenty y_2 – průměrná délka přízvukového taktu (měřena v počtu slabik)	47
Tabulka 7: U2, výběrový soubor B, verze 1: konstrukty x_2 – délka výpovědi (měřená v počtu přízvukových taktů), z_2 – frekvence konstruktů, konstituenty y_2 – průměrná délka přízvukového taktu (měřena v počtu slabik)	48
Tabulka 8: U2, výběrový soubor B, verze 2: konstrukty x_2 – délka výpovědi (měřená v počtu přízvukových taktů), z_2 – frekvence konstruktů, konstituenty y_2 – průměrná délka přízvukového taktu (měřena v počtu slabik).....	50
Tabulka 9: Přehled parametrů b MAL a koeficientů determinace na U2	51
Tabulka 10: U3, výběrový soubor A: konstrukty x_3 – délka přízvukových taktů (měřena v počtu slabik), z_3 – frekvence konstruktů, konstituenty y_3 – průměrná délka slabiky (měřena v počtu hlásek).....	53
Tabulka 11: U3, výběrový soubor B: konstrukty x_3 – délka přízvukových taktů (měřena v počtu slabik), z_3 – frekvence konstruktů, konstituenty y_3 – průměrná délka slabiky (měřena v počtu hlásek).....	55
Tabulka 12: Přehled parametrů b MAL a koeficientů determinace na U3	57
Tabulka 13: U3Š, výběrový soubor C: konstrukty x_3 – délka segmentu (měřena v počtu slabik), z_3 – frekvence konstruktů, konstituenty y_3 – průměrná délka slabiky (měřena v počtu hlásek)	59
Tabulka 14: Přehled parametru b MAL a koeficientu determinace na U3Š	60

Tabulka 15: U2Š-U1Š, výběrový soubor C: konstrukty x_a – délka kolónů (měřena v počtu slabik), f_a – frekvence konstruktů, konstituenty y_a – průměrná délka slabiky (měřena v počtu hlásek) 65

Tabulka 16: Přehled parametru b MAL a koeficientu determinace na U2Š-U1Š..... 66

SEZNAM GRAFŮ

Graf 1: U1 replika vs. výpověď – soubor A: vizualizace souboru dat v Tabulce 2, pomocí programu R.....	42
Graf 2: U1 replika vs. výpověď – soubor B, verze 1: vizualizace souboru dat v Tabulce 3, pomocí programu R.....	43
Graf 3: U1 replika vs. výpověď – soubor B, verze 2: vizualizace souboru dat v Tabulce 4, pomocí programu R.....	44
Graf 4: U2 výpověď vs. přízvukový takt – soubor A: vizualizace souboru dat v Tabulce 6, pomocí programu R.....	48
Graf 5: U2 výpověď vs. přízvukový takt – soubor B, verze 1: vizualizace souboru dat v Tabulce 7, pomocí programu R.....	49
Graf 6: U2 výpověď vs. přízvukový takt – soubor B, verze 2: vizualizace souboru dat v Tabulce 8, pomocí programu R.....	51
Graf 7: U3 přízvukový takt vs. slabika – soubor A: vizualizace souboru dat v Tabulce 10, pomocí programu R	54
Graf 8: U3 přízvukový takt vs. slabika – soubor A: vizualizace souboru dat v Tabulce 10, pomocí programu Microsoft Excel	54
Graf 9: U3 přízvukový takt vs. slabika – soubor B: vizualizace souboru dat v Tabulce 11, pomocí programu R	56
Graf 10: U3 přízvukový takt vs. slabika – soubor B: vizualizace souboru dat v Tabulce 11, pomocí programu Microsoft Excel	56
Graf 11: U3Š segment vs. slabika – soubor C: vizualizace souboru dat v Tabulce 13, pomocí programu R.....	59
Graf 12: U3Š segment vs. slabika – soubor C: vizualizace souboru dat v Tabulce 13, pomocí programu Microsoft Excel	60
Graf 13: U3 přízvukový takt vs. slabika – soubor A: vizualizace souboru dat v Tabulce 10 po odebrání bodu $x = 1$, pomocí programu R	63
Graf 14: U3 přízvukový takt vs. slabika – soubor B: vizualizace souboru dat v Tabulce 11 po odebrání bodu $x = 1$, pomocí programu R	63
Graf 15: U3Š segment vs. slabika - soubor C: vizualizace souboru dat v Tabulce 13 po odebrání bodu $x=7$, pomocí programu R	64

Graf 16: U2Š-U1Š kólon vs. slabika - soubor C: vizualizace soubor dat v Tabulce 15 pomocí programu R.....	66
---	----

EDIČNÍ POZNÁMKA

V práci je použita česká transkripce, pinyin a zjednodušené znaky.

ÚVOD

Tématem mé diplomové práce je aplikace Menzerath-Altmanova zákona na současnou mluvenou čínštinu se zvláštním zřetelem na nejnižší jazykovou úroveň tvořenou přízvukovým taktem, respektive segmentem (konstrukt, jehož délka je měřena v počtu slabik) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek).

Výzkum probíhal ve dvou fázích. První fáze výzkumu byla realizována v rámci projektu studentské grantové soutěže IGA FF_2012_035 *Segmentace pro testování Menzerath-Altmanova zákona a hypotéz s ním souvisejících*. Během tohoto výzkumu byl MAL aplikován na současnou mluvenou a psanou čínštinu. Já jsem byla členem té části týmu, která se zabývala mluvenou čínštinou.

V první fázi výzkumu jsme s kolegyní aplikovaly MAL na dva výběrové soubory současné mluvené čínštiny. MAL byl v této fázi zkoumán na následujících jazykových úrovních: na jazykové úrovni tvořené replikou (konstrukt, jehož délka je měřena v počtu výpovědí) a výpovědí (konstituent, jehož délka je měřena v průměrném počtu přízvukových taktů), dále na jazykové úrovni tvořené výpovědí (konstrukt, jehož délka je měřena v počtu přízvukových taktů) a přízvukovým taktem (konstituent, jehož délka je měřena v průměrném počtu slabik) a na jazykové úrovni tvořené přízvukovým taktem (konstrukt, jehož délka je měřena v počtu slabik) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek).

V druhé výzkumné fázi byl MAL aplikován na třetí výběrový soubor současné mluvené čínštiny. Zkoumané jazykové úrovně pak byly tyto: jazyková úroveň tvořená kólonem (konstrukt, jehož délka je měřena v počtu segmentů) a segmentem (konstituent, jehož délka je měřena v průměrném počtu slabik) a jazyková úroveň tvořená segmentem (konstrukt, jehož délka je měřena v slabikách) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek).

Zmiňované tři výběrové soubory se od sebe odlišují především tím, že první dva zachycují spontánní, nepřipravenou řeč, zatímco třetí výběrový soubor zachycuje předem připravenou promluvu. Výběrové soubory A a B byly vytvořeny za účelem

pobavit televizní diváky, oproti tomu výběrový soubor C byl vytvořen pro studijní účely. To se samozřejmě projevuje na kvalitě promluv všech tří výběrových souborů.

Výběrové soubory A a B pojí společné téma, a to hudba. Oba mluvčí jsou hudebníci, každý z nich se však zabývá jiným žánrem hudby. Mluvčí výběrového souboru A je popová zpěvačka, zatímco mluvčím výběrového souboru B je rapový zpěvák. Promluva výběrového souboru A je charakteristická konstantním rychlým tempem řeči. U výběrového souboru B se pak tempo promluvy často zrychluje a opět zpomaluje. Charakteristickým znakem promluvy výběrového souboru B je také emocionalita projevu.

První dva výběrové soubory jsme rozčlenily samy, třetí výběrový soubor je pak segmentace podle Oldřicha Švarného.

Pro výše uvedený typ souborů jsme se rozhodly proto, že by podle našeho názoru bylo zajímavé komparovat promluvu zcela spontánní s promluvou, která je předem připravená. Zajímalo nás, zda a případně do jaké míry se budou lišit výsledky výpočtů provedené na souborech zmiňované povahy.

Původní hypotéza byla taková, že MAL se projeví na všech jazykových úrovních současné mluvené čínštiny. Výsledky výpočtů provedených v první fázi výzkumu však ukázaly, že na nejnižší jazykové úrovni se MAL neprojevil, ukázalo se však, že hodnoty u obou výběrových souborů oscilují v rámci pásu.

V druhé fázi výzkumu jsem se tedy zaměřila především na nejnižší jazykovou úroveň s cílem potvrdit hypotézu, že na nejnižší jazykové úrovni tentokrát však tvořené segmentem (konstrukt, jehož délka je měřena v počtu slabik) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek) se vždy projeví oscilace v rámci pásu.

Má diplomová práce je rozdělená do dvou hlavních oddílů, a to na teoretickou a praktickou část. Jak už vyplývá ze samotného názvu, tato práce v sobě zahrnuje dvě disciplíny, jednak čínštinu, jednak kvantitativní lingvistiku. V teoretické části proto věnuji jednu kapitolu stručnému popisu a historii kvantitativní lingvistiky. V další kapitole se pak věnuji Menzerath-Altmanovu zákonu samotnému. Součástí teoretické části je dále metodologie - popis zvolených výběrových souborů, jejich

segmentace a problémů s ní spojených, a popis a definice zkoumaných jazykových úrovní.

V praktické části se pak věnuji demonstraci dosažených výsledků pomocí grafů a tabulek, a následně diskusi nad výsledky, a jejich možnými příčinami. Praktická část je členěna podle zkoumaných jazykových úrovní, přičemž nejprve uvádím výsledky z první fáze výzkumu, tj. výběrové soubory A a B, poté následují výsledky z druhé fáze výzkumu, tj. výběrový soubor C. Po závěru pak následuje resumé a seznam použitých zdrojů a literatury.

Ve své diplomové práci pracuji zejména s primárními prameny, kterými jsou výběrové soubory A, B a C. V teoretické části se také ale opírám o mnohé sekundární zdroje. Co se týče informací ohledně kvantitativní lingvistiky a jejich metod, vycházím především z díla Marie Těšitelové. Při zpracovávání MALu pak vycházím především z díla Gabriela Altmanna a Lud'ka Hřebíčka. Co se týče mluvené čínštiny, vycházím především z díla Oldřicha Švarného.

1. LINGVISTIKA

Lingvistika, nebo také jazykověda, je věda, jak ostatně vypovídá její český název, zabývající se studiem jazyka. Vzhledem k faktu, že existuje velké množství jazyků a vzhledem k složitosti jazyka obecně, se lingvistika dělí na množství podoborů. Dělicím kritériem je pak konkrétní jazyk, či určité části jazyka obecně.¹

„Podle prvního kritéria tak rozlišujeme např. bohemistiku, rusistiku, anglistiku, hispanistiku atd., nebo slavistiku, germanistiku, romanistiku atd. Tyto lingvistické obory se zabývají buď studiem jednoho jazyka, nebo studiem celé skupiny příbuzných jazyků a jejich vzájemných vztahů.“²

„Podle druhého hlediska se jazykověda dělí na fonetiku a fonologii (které se zabývají zvukovou stránkou jazyka), gramatiku (kde morfologie studuje slovní tvary a syntax větnou stavbu), lexikologii a lexikografii (slovní zásoba), sémantiku (význam), dialektologii (zeměpisné nebo společenské rozvrstvení jazyka), stylistiku (sloh ústních i písemných projevů) atd.“³

Vzhledem k tomu, že ve své diplomové práci využívám některých metod kvantitativní lingvistiky, považuji za nutné alespoň stručně nastínit její historický vývoj.

Než se dostanu k lingvistice matematické, pod kterou právě kvantitativní lingvistika spadá, zmíním se nejdříve o strukturalismu.

1.1 Strukturalismus

Dvacáté století znamenalo převratnou dobu pro lingvistiku. V jeho druhém a třetím desetiletí vznikl v Ženevě strukturalismus. Strukturalismus byl v jazykovědě směrem revolučním, a to z toho důvodu, že nahlížel na jazyk jako na systém, jehož jednotlivé

¹ ČERNÝ, Jiří. *Dějiny lingvistiky*. Olomouc: Votobia, 1996, 517 s, s. 16.

² ČERNÝ, *Dějiny*, s. 16.

³ ČERNÝ, *Dějiny*, s. 16.

části nelze zkoumat izolovaně. Do té doby zkoumali jazykovědci všechny aspekty jazyka samostatně, nezávisle jeden na druhém. Zakladatelem strukturalismu je Ferdinand de Saussure (švýcarský jazykovědec žijící v letech 1857-1913). Strukturalismus významně ovlivnil následující generace lingvistů a směr, kterým se dále lingvistika ubírala.⁴

Důvod, proč se zde o strukturalismu zmiňuji je především ten, že strukturalismus zavedl pojem lingvistika, tím se teoretický obor odlišil od filologie a byl zdůrazněn rozdíl mezi vědeckým a předvědeckým pojetím jazyka.⁵

1.2 Matematická lingvistika

1.2.1 Kvantitativní lingvistika

V následující části bych ráda definovala pojem matematická lingvistika a stručně popsala historii kvantitativní lingvistiky. Nejdříve je třeba se zmínit o lingvistice matematické. Matematickou lingvistikou dnes rozumíme pomezí disciplínu, která podle některých lingvistů stojí na rozhraní matematiky a jazykovědy. Podle jiných je to část lingvistiky využívající matematických metod. Matematická lingvistika vznikla na konci padesátých a na začátku šedesátých let dvacátého století.

Matematických metod se ale v lingvistice užívalo již mnohem dříve, dokonce již ke konci devatenáctého století. Metody tenkrát užívané však dnes označujeme spíše jako metody kvantitativní nebo statistické. Kvantitativní lingvistika tedy navazuje na určitou tradici, přičemž její základy byly v naší lingvistice položeny již ve 30. letech dvacátého století a jsou spjaty s pracemi jazykovědců, jako byli Vilém Mathesius, Bohumil Trnka a Josef Vachek.⁶

Kořeny kvantitativní lingvistiky u nás však sahají hlouběji do historie. Jak píše Těšitelová:

⁴ ČERNÝ, *Dějiny*, s. 128.

⁵ HŘEBÍČEK, Luděk. *Vyprávění o lingvistických experimentech s textem*. Vyd. 1. Praha: Academia, 2002, 195 s., s. 36.

⁶ TĚŠITELOVÁ, Marie. *Kvantitativní lingvistika*. 1. vyd. Praha: Univerzita Karlova, 1987, s. 9.

„pokud jde o uplatnění kvantitativních metod při studiu slovní zásoby, vzato ovšem se značnou mírou volnosti, můžeme hledat začátky u našeho velkého pedagoga J.A. Komenského; ten ve svém díle *Janua linguarum reserva* ukázal, jak možno znalostí o počtu slov v jazyce a o jejich opakovaném výskytu (frekvenci) využít k ekonomickému a efektivnímu rozvíjení slovní zásoby u žáků.“⁷

K tradici kvantitativní lingvistiky u nás Těšitelová dále píše: „Z hlediska metodologického stojí dále za pozornost, že v 80. letech minulého století aplikoval počet pravděpodobnosti na studium jazykových jevů matematik A. Seydler; šlo o boj o tzv. Rukopisy – královédvorský a zelenohorský.“⁸

V současnosti je běžně užíván termín kvantitativní lingvistika, ale někteří lingvisté dávají s ohledem na převažující metodu přednost pojmu statistická lingvistika. Těšitelová definuje matematickou lingvistiku následovně:

„Matematická lingvistika zkoumá jazyk, přirozený a umělý, s pomocí metod matematických, popř. logických. Jde o metody jednak kvantitativní (jako je statistika, počet pravděpodobnosti, matematická statistika apod.), jednak nekvantitativní (jako je algebra, teorie grafů, topologie, matematická logika apod.). Metod kvantitativních využívá v zásadě kvantitativní lingvistika, metod nekvantitativních, kvalitativních algebraická lingvistika.“⁹

Kvantitativní lingvistika nám dovoluje rozpoznat jazykové zákonitosti, a to právě pomocí sběru kvantitativních dat o konkrétních jazykových jevech a vztazích mezi nimi. Na podkladě takových údajů pak lze vytvářet modely, nebo také porovnávat celý systém takových modelů a skutečností.¹⁰

⁷ TĚŠITELOVÁ, *Kvantitativní*, s. 9.

⁸ TĚŠITELOVÁ, *Kvantitativní*, s. 9.

⁹ TĚŠITELOVÁ, *Kvantitativní*, s. 7.

¹⁰ TĚŠITELOVÁ, *Kvantitativní*, s. 9.

Co se týče aplikace metod kvantitativní lingvistiky na čínštinu obecně, kvantitativní lingvistika nemá zatím v Číně příliš dlouhou tradici a dá se říct, že Číňané se s tímto oborem teprve seznamují.¹¹

1.2.2 Algebraická lingvistika

Jak jsem zmínila výše, druhým odvětvím matematické lingvistiky je lingvistika algebraická. Vznik algebraické lingvistiky se datuje do druhé poloviny padesátých let dvacátého století. Algebraická lingvistika dnes zahrnuje celou řadu metod, které se od sebe v některých případech značně liší. Spojuje je však fakt, že se pomocí formální logiky snaží exaktně popsat funkční hodnoty jednotek, s kterými se pracuje při analýze přirozených jazyků.¹²

¹¹ Liu Haitao, Huang Wei. **Jiliang yuyanxue de xianzhuang, lilun yu fangfa.** *Zhejiang daxue xuebao* (renwen shehui kexue ban) 2012, roč. 42, č. 2. ISSN 1008-942X.

¹² ČERNÝ, *Dějiny*, s. 280.

2. MENZERATH-ALTMANNŮV ZÁKON

Jelikož nástrojem našeho výzkumu je MAL, následující kapitolu proto věnuji jeho popisu a stručné historii.

Menzerath-Altmannův zákon je považován za jeden z fundamentálních jazykových zákonů. Je založen na následujícím principu: čím delší je slovo, tím kratší jsou v průměru jeho slabiky. Tento vztah nejprve pozoroval v roce 1928 Paul Menzerath. Jeho významný objev však zůstal více než půl století bez povšimnutí. Menzerathova myšlenka tak zůstala na dlouhou dobu nevyužita.¹³

Teprve až v roce 1980 se Gabriel Altmann (německý lingvista slovenského původu) rozhodl publikovat článek, ve kterém formuloval Menzerathův zákon. Altmann dále zavedl dva nové pojmy: jazykový konstrukt a jazykový konstituent. Altmann následně tento zákon testoval v různých jazycích pro jednotky různých jazykových úrovní.¹⁴ „Ověřil, že mezi těmito entitami existuje následující charakteristický vztah: Čím delší je v jazyce nějaký konstrukt, tím kratší jsou v průměru jeho konstituenty.“¹⁵

Algebraická podoba zákona, kterou odvodil Altmann:

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

kde x je délka konstruktů

y je délka konstituentů

A , b jsou reálné parametry

Zpřesňující matematická formulace tohoto zákona je následující:

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

kde x je opět délka konstruktů

y délka konstituentů

¹³ HŘEBÍČEK, *Vyprávění*, s. 53.

¹⁴ HŘEBÍČEK, *Vyprávění*, s. 53.

¹⁵ HŘEBÍČEK, *Vyprávění*, s. 53.

A, b, c jsou reálné parametry¹⁶

Z obsahu Menzerath-Altmanova zákona vyplývá, že je jím v podstatě určeno kritérium pro jazykovou úroveň. Chceme-li uvažovat o nějakých jednotkách jako o jazykové úrovni, musí být požadováno splnění principu, který je tímto zákonem vyjádřen.¹⁷

„Dvě jazykové jednotky různého druhu mohou být uvažovány jako vztah dvou vzájemně hierarchizovaných úrovní, čili mohou být uvažovány jako konstrukt a konstituent... V textu je realizován řetězec:

... ↔ KONSTRUKT/KONSTITUENT ↔ KONSTRUKT/KONSTITUENT
↔ KONSTRUKT/KONSTITUENT ...

Jinak řečeno, každá jazyková entita vůči všem vyšším jazykovým úrovním je konstituentem a vůči všem nižším úrovním je konstruktem. Jazykové úrovně jsou do sebe zasunuty jedna do druhé. Mezi jazykovými úrovněmi existuje podobnost vyjádřená Menzerathovým-Altmanovým zákonem.“¹⁸

„Jazykové úrovně známe z pozorování konkrétních jazyků. Víme, že existují tyto jednotky:

hláska – morf – (slabika) – slovo – syntaktická konstrukce – věta

Máme tedy možnost kdykoliv ověřit platnost tohoto zákona.“¹⁹ Konstituent závisí na konstruktu, ale i naopak konstrukt na konstituentu. Konstrukt i konstituent jsou relativní pojmy. Konstrukt představuje jazykovou jednotku na vyšší úrovni, zatímco

¹⁶ HŘEBÍČEK, *Lectures on text theory*. Prague: Oriental Institute, 1997, s. 22.

¹⁷ HŘEBÍČEK, *Vyprávění*, s. 59.

¹⁸ HŘEBÍČEK, *Vyprávění*, s. 59-60.

¹⁹ HŘEBÍČEK, Luděk. **Struktura textu a sémantika**. *AntropoWebzin* [online]. 2006, 1,2, [cit. 2013-03-01]. ISSN 1801-8807. Dostupné z: <<http://antropologie.zcu.cz/>> s. 37.

konstituent reprezentuje jazykovou jednotku na nižší úrovni. Například slovo je konstruktem vzhledem ke slabice a konstituentem vzhledem ke klauzi.²⁰

²⁰ HŘEBÍČEK, *Vyprávění*, s. 37.

3. FONETICKÝ SYSTÉM ČÍNŠTINY

Vzhledem k faktu, že předmětem našeho bádání je aplikace MALu na současnou mluvenou čínštinu, ráda bych se v této kapitole věnovala stručnému popisu čínského fonetického systému. Jelikož byla čínština vždy písmem znakovým, tedy písmem, ve kterém se fonetické změny vůbec neodrážejí, je vývoj mluvené čínštiny poněkud zastřen.²¹ V této kapitole se tedy nebudu věnovat historickému vývoji mluvené čínštiny, ale pouze fonetickému systému současné čínštiny.

Čínština je jazyk slabičný, mající daný, omezený soubor slabik. Tím se značně liší od češtiny, která disponuje neomezeným setem slabik, a za pomoci nejrůznějších kombinací hlásek mohou vznikat slabiky nové. V češtině „formu slova tvoří jeho hlásková stavba. V naprosté většině případů vytváří slovo skupina hlásek, např. úl, oko, šest, letět.“²²

Základní stavební jednotkou v čínštině při tvoření slov je však slabika. Pokud tedy v čínštině vznikají nová slova, pak pouze za použití oněch limitovaných, již existujících slabik. V čínštině nelze nová slova vytvářet náhodnou, libovolnou kombinací hlásek. A tak i při vzniku nových slov v čínštině pomocí výpůjček z cizích jazyků, je daný výraz převeden pomocí čínských slabik.

Čínštinu lze charakterizovat jako jazyk s tónickým přízvukem.²³ Tón považujeme za součást slabiky. Slabika se v čínštině dělí na dvě hlavní části: iniciálu a finálu, přičemž slabika nemusí obsahovat iniciálu. Povinným komponentem slabiky je finála. „Iniciál napočítáme toliko 21, finál 34. Všech foneticky rozlišených slabik ohlédneme-li od rozlišující funkce tónů a rozlišení pomocí sufixace /-r/) napočítáme toliko něco přes 400; všechny slabiky se pak řadí v několik málo zcela jednoduchých typů: souhláska + jednoduchá samohláska, souhláska + vzestupná dvojhláska, souhláska + sestupná dvojhláska apod.“²⁴ I s tóny je slabik ale jen 1200, protože ne každá slabika se realizuje ve všech čtyřech tónech. Pro podrobnější informace

²¹ ŠVARNÝ, Oldřich a David UHER. *Hovorová čínština: úvod do studia hovorové čínštiny*. Olomouc: Univerzita Palackého, 2001, 190 s., s. 12.

²² HAUSER, Přemysl. *Nauka o slovní zásobě*, Praha: Státní pedagogické nakladatelství, 1980, 192 s., s. 9.

²³ ŠVARNÝ, *Úvod*, s. 37.

²⁴ ŠVARNÝ, *Úvod*, s. 16.

ohledně fonetického systému čínštiny viz. *Hovorová čínština: úvod do studia hovorové čínštiny*.

4. POPIS VÝBĚROVÝCH SOUBORŮ

4.1 Výběrové soubory A a B

Na samotném počátku výzkumu stálo rozhodnutí zvolit vhodné výběrové soubory pro plánovaný experiment. Vzhledem k faktu, že předmětem našeho bádání byla současná hovorová čínština, bylo třeba vybrat takový výběrový soubor, který byl produkován rodilým mluvčím. Dále bylo nutné, aby mluvený text splňoval následující kritéria: text měl být ucelený, předem nepřipravený, tj. mluvčí neměl text číst, ale pronášet ho spontánně. Promluva měla být pokud možno co nejvíce přirozená, aby mohla odrážet přirozený tok a tempo řeči, a poskytnout tak co nejlepší materiál k výzkumu, aby bylo dosaženo co nejpřesnějších výsledků. Co se týče délky výběrových souborů, stanovily jsme na počátku minimum 1000 slabik.

Abychom docílily lepšího srovnání, rozhodly jsme se vybrat takové výběrové soubory, které byly tematicky podobné. Pojícím prvkem byla v tomto případě hudba, respektive fakt, že oba mluvčí výběrových souborů byli hudebníci. Nahrávky, kterých bylo použito jako výběrových souborů, byly vytvořeny pro čínský televizní hudební pořad. Tyto dvě nahrávky byly zpracovány v rámci projektu studentské grantové soutěže IGA FF_2010_42 *Korpus hovorové čínštiny*.

Jak jsem zmínila výše, pro náš experiment jsme nakonec zvolily dva díly čínského televizního hudebního pořadu. Jde o typ programu, ve kterém vždy konkrétní hudebník či hudebníci hovoří o hudbě, kterou se zabývají. Jednotlivé vstupy mluvčího jsou prokládány písněmi daného hudebníka. Z popisu pořadu tedy vyplývá, že se nejedná o zcela přirozenou promluvu, respektive jedná, ta je však nastříhána a uměle upravena pro účely pořadu. Text výběrového souboru, který jsme v našem projektu označily jako výběrový soubor A, byl produkován popovou zpěvačkou. Oproti tomu text souboru označeného jako výběrový soubor B byl promluvou rapového zpěváka. Naším cílem bylo aplikovat MAL na promluvy právě těchto dvou hudebníků a zjistit tak, zdali má hudební žánr vliv na řeč mluvčích, kteří mohou být právě hudebním žánrem, který zpívají, ovlivněni.

4.2 Výběrový soubor C

Pro svou diplomovou práci jsem se rozhodla zvolit ještě jeden výběrový soubor. Kromě vzorku A a vzorku B, které jsem popsala výše, jsem si vybrala ještě vzorek C. Vzorek C se oproti dvěma dalším liší svou povahou.

Jako poslední výběrový soubor jsem zvolila zvukové stopy (č. 136, 137, 142, 143, 165, 170, 171, 172, 175, 176, 177) z *Gramatiky hovorové čínštiny v příkladech* (GHČP). V přepisu těchto stop v GHČP, který je pořízen podle magnetofonové nahrávky, je zachycena prozodická realizace každé věty a souvětí. Tyto krátké pasáže jsou produkovány rodilým mluvčím, a tudíž splňují jednu z podmínek, které jsem si stanovila již při volbě prvních dvou výběrových souborů.

Třetí výběrový soubor se ale odlišuje především tím, že text není spontánní. Jedná se o předem připravený soubor vět, které jsou následně zvukově realizovány rodilou mluvčí. Ta sice předem zná obsah vět, nicméně se je při nahrávání snaží vyslovit tak, jak by měly zaznít ve spontánním, předem nepřipraveném projevu, snaží se je vyslovit jako normu.

Vzorek C se tedy od vzorku A a B odlišuje především tím, pro koho je určen. Vzorky A a B jsou zábavní populární televizní pořady, jejichž cílem je pobavit diváky. To se také odráží na jazyce a projevu mluvčích. Oba mluvčí realizují svou promluvu spontánně. Výběrový soubor A je specifický především rychlým tempem řeči. Mluvčí tohoto vzorku pronáší svou řeč tak rychle, že se velmi často „zakoktavá“ a tak je její promluva obohacena velkým množstvím hezitačních zvuků. Mluvčí vzorku B velice často v průběhu své promluvy používá vulgarismů, a počet cizích slov, zejména anglických, je také nemalý. Soubor B je také výrazně ovlivněn emocionalitou mluvčího. Oproti souboru A, kde mluvčí hovoří ve velice rychlém tempu konstantně, mluvčí souboru B právě díky zvýšené emocionalitě tempo své promluvy velice často mění, a mnohdy nečekaně přechází od pomalého k rychlému tempu řeči.

5. DEFINOVÁNÍ JAZYKOVÝCH JEDNOTEK

5.1 Výběrové soubory A a B

Dalším důležitým problémem, který bylo na začátku experimentu třeba vyřešit, byl výběr metod, kterých bude používáno během analýzy výběrových souborů a určení jazykových jednotek a úrovní, které budou zkoumány.

„V promluvě lze vyčlenit tyto artikulačně-akustické jednotky: výpověď, promluvový úsek (výpovědní úsek, taktovou skupinu, kólon), přízvukový takt (fonetické slovo), slabiku. Takt a slabika jsou jednotky zvukové formy, jejichž souvislost s členěním významovým není bezprostřední, i když takt je velmi často totožný se slovem a slovo může být jednoslabičné. K hlásce se dospívá až analýzou slabik, samostatnou zvukovou jednotkou není. Vyšší jednotky jsou ohraničeny hraničními signály.“²⁵ K samotnému výběru jednotek souboru Těšitelová píše:

„v kvantitativní lingvistice přejímáme obvykle pojetí běžné v lingvistice a podle potřeby – se zřetelem k aplikovaným metodám – je upravujeme, modifikujeme. Vzhledem k tomu, že pro aplikaci statistických metod je základní podmínkou jednoznačné vymezení jednotky souboru, volíme – při nesnadném vedení přesné hranice mezi jednotlivými jazykovými jevy – maximálně vyhovující definici a tu potom důsledně dodržujeme. Platí pak dále zásada, že se každá jednotka někde započítává a žádná jednotka se nepočítá dvakrát.“²⁶

Na počátku našeho výzkumu jsme se rozhodly, že se v průběhu celého našeho bádání budeme striktně držet pouze fonetického hlediska. Později jsme si však uvědomily, že je to cíl nemožný a reálně neuskutečnitelný. Abych to objasnila, je třeba nejprve zmínit se o jazykových jednotkách, které jsme se rozhodly v první fázi výzkumu zkoumat. Konkrétně to tedy byly následující jednotky: replika, výpověď, přízvukový takt, slabika, hláska. Z těchto pěti jazykových jednotek ne všechny jsou jednotky odrážející výhradně fonetické hledisko.

²⁵ KARLÍK, Petr. *Encyklopedický slovník češtiny*. Praha: NLN, 2002, 604 s., s. 102.

²⁶ TĚŠITELOVÁ, *Kvantitativní*, s. 12.

Dvě nejvyšší jednotky, tedy replika a výpověď zcela jistě zohledňují také hledisko syntaktické. Hranice repliky a výpovědi se neurčují pouze podle fonetických hledisek, ale také na základě syntaktických hledisek, či obsahové soudržnosti. Ostatní tři nižší úrovně jsou fonetické.

Níže uvedené definice jazykových jednotek jsou podle Krčmové. Krčmová následující definice aplikuje především na češtinu. Vzhledem k tomu, že se v našem výzkumu zabýváme současnou mluvenou čínštinou, která je od češtiny typologicky odlišná, rozhodly jsme se stanovit pro jednotlivé úrovně vlastní definice. U některých úrovní jsme k tomu byly donuceny typem výběrového souboru. Konkrétně pak při stanovení definice pro jazykovou úroveň replika.

5.1.1 Replika

Krčmová definuje repliku (promluvu) jako:

„celek řeči mezi dvěma absolutními pauzami. Z mimozvukových prostředků ji sjednocuje obsahová soudržnost (koherence) a jediný mluvčí. Z fonetického hlediska je dialog, ev. polylog sledem promluv (vzájemně navazujících obsahem), zatímco z textového hlediska jde o jediný (dialogický, polylogický) text. Délka promluvy je potenciálně neomezená.“

Replika může být také definována jako:

„základní jednotka členění dialogického textu; souvislý úsek textu, který pronese jeden z účastníků rozhovoru, aniž by ho vystřídal mluvčí jiný. Obvykle je tvořena jednou nebo několika výpověďmi; může být velmi rozsáhlá (např. vyprávění nepřerušované posluchači), nebo v živém přirozeném dialogu velmi krátká.“²⁷

Na základě výše uvedených definic jsme se rozhodly stanovit promluvu jako úsek mezi dvěma vsunutými hudebními klipy. Tyto jednotlivé úseky jsou celky mezi dvěma absolutními pauzami, jsou obsahově koherentní a pronášeny jediným mluvčím, zároveň jsou však v podstatě součástí dialogu, jelikož mluvčí jakoby

²⁷KARLÍK, *Encyklopedický*, s. 374.

odpovídal na dané otázky, které jsou však ze souboru odstraněny. V pozdějším stadiu výzkumu se ukázalo, že frekvence replik byla nízká. Domnívám se však, že na nejvyšší úrovni k tomuto problému dochází velmi často.

5.1.2 Výpověď

Podle Krčmové je výpověď „obvykle menší než promluva, může však být i minimální promluvou. Z fonetického hlediska je vyznačena větnou intonací a relativními pauzami, mnohdy jen potenciálními. Promluva je kohézní (souvislá). Je zvukově členěna větným přízvukem a intonací a je tedy nositelkou intonačních prostředků modulace řeči“.

Podle výše uvedené definice výpovědi je patrné, že nejen fonetické hledisko, ale také sémanticko-syntaktický aspekt je velmi důležitý při stanovení hranic výpovědi. Tak jsme také při segmentaci výběrových souborů učinily a výpověď jsme pro náš výzkum stanovily jako sémanticko-syntaktický ucelený celek.

V případě segmentace vzorku B jsme však narazily na problém, a to vulgarismy. Mluvčím výběrového souboru B je rapový zpěvák, jehož řeč je protkána množstvím vulgárních a sémanticky prázdných slov. Konkrétně šlo o to, zda je počítat jako součást výpovědi, či ne. Na jednu stranu totiž některé z vulgarismů fungují jako atributy a rozvíjejí některá podstatná jména. Na druhou stranu však v některých případech nehrají z hlediska významového obsahu textu žádnou roli a v promluvě zpěváka se objevují spíše jako vsuvky zdůrazňující jeho ztotožnění se s určitou skupinou, v tomto případě konkrétně s rapery, pro jejichž mluvu jsou vulgarismy postrádající jakýkoli sémantický význam typické.

U vzorku B jsme se proto rozhodly pro dvě různé verze segmentace. V první verzi považujeme vulgarismy a sémanticky prázdná slova za součást výpovědi. V druhé verzi však tato sémanticky prázdná slova a vulgarismy vyčleňujeme a nahlížíme na ně jako na samostatné výpovědi.

5.1.3 Přízvukový takt

Jak uvádí Krčmová, takt je „skupina slabik patřící k jednomu slovnímu přízvuku. Jde o jednotku shodnou se slovem, může však být i větší. K jediné slabice charakterizované přízvukem se přičleňuje několik slabik, které přízvuk nemají.“²⁸

Černý definuje přízvukový takt následovně: „...z hlediska přízvuků se tedy věty dají dělit do mluvnických taktů, tj. rytmických jednotek, které obsahují jedno nebo několik slov a je pro ně charakteristický jeden hlavní přízvuk.“²⁹

K hranici taktu dále Krčmová udává: „Hranice taktu je totožná s hranicí některého slova, které takt tvoří... Hranice taktu je (navíc) signalizována potenciální pauzou, kromě toho se v některých jazycích na hranici taktů při neutrální a pečlivé výslovnosti neprovádějí hláskové změny, jež jsou uvnitř taktu pravidlem (vzájemné přízpusobování konsonantů a vokálů, asimilační změny).“³⁰

Podle Krčmové takt také někdy bývá označován jako fonetické slovo, a to proto, že je základní zvukovou jednotkou přirozeně vyslovované řeči. Z toho vyplývá že, „mezi taktem a slovem ve významu lexikálním a morfologickém může být korespondence: plnovýznamové slovo může být vždy samostatným taktem (pomocná slova mají tuto schopnost jen potenciálně, jsou-li samostatně vyslovena nebo jsou-li jádrem sdělení, jinak bývají bez přízvuku). Slovo je minimálním taktem.“³¹

Krčmová uvádí, že „délka taktu je dána jazykovou zvyklostí a souvisí i s obvyklou délkou slova. Obecně lze říci, že takty nejsou nikdy příliš dlouhé, v češtině mají často např. 2-3 slabiky.“³²

Výše uvedené definice přízvukového taktu jsou stanoveny převážně pro češtinu. Švarný ve své segmentaci čínského textu pojem přízvukový takt nepoužívá, používá pojem segment. „Segmentům v čínštině svou délkou a určitou významovou celistvostí odpovídají v našich jazycích přibližně takty. V plynulé řeči však naše takty jsou navzájem odsazeny prakticky neznatelně a i dlouhé sledy slabik

²⁸ KRČMOVÁ, Marie. *Fonetika* [online] Brno: Filosofická fakulta MU, 2007 [17. 01. 2013] Dostupné z: is.muni.cz/el/portal/estud/ff/js07/fonetika/materialy/ch04.html.

²⁹ ČERNÝ, Jiří. *Úvod do studia jazyka*. Olomouc: Rubico, 2008, 248 s, s. 89.

³⁰ KRČMOVÁ

³¹ KRČMOVÁ

³² KRČMOVÁ

vyslovujeme v češtině nebo slovenštině „spojitě“, bez mezer.“³³ Podle Švarného je průměrná délka segmentů variabilní a závislá na tempu řeči a „v průměru kolísá mezi 2,5 až 4,5 slabiky.“³⁴

Švarný také upozorňuje na to, že v pomalejší řeči se délka přízvukového taktu, respektive segmentu zkracuje. Při rychlejším tempu řeči se délka přízvukového taktu prodlužuje.³⁵

My jsme při našem výzkumu za takt považovaly skupinu slabik s minimálně jedním přízvukem, která je oddělená od ostatních pauzou.

Co se týče dvou následujících jazykových jednotek, jmenovitě slabiky a hlásky, vymezení jejich hranic nepředstavuje problém. Ohledně těchto dvou jazykových hladin však vedou jazykovědci spory o tom, která z nich je tou nejnižší základní jednotkou mluvené řeči.

5.1.4 Slabika

Černý ve svém „Úvodu do studia jazyka“ tvrdí, že „otázka podstaty slabiky nebyla dodnes uspokojivě vyřešena a vedou se dokonce spory o to, zda základní jednotkou mluvené řeči je slabika nebo hláska.“³⁶ Také Krčmová píše „slabika je nejjednodušší a nejtěsnější možnou artikulační jednotou funkčních prvků řeči, která vyhovuje dorozumívání.“³⁷ Dále se zmiňuje, že:

„Fonetickou podstatu slabiky se však dosud nepodařilo jednoznačně popsat: vždy zůstává jistý počet hláskových kombinací, jež jsou uživateli hodnoceny jako slabiky, ale teoretickým vymezením neodpovídají. Žádný z principů vymezení popsanych ve starší literatuře nedostačuje sám pro poznání podstaty slabiky, neboť při artikulaci i vnímání se uplatňují společně a v jednotlivých typech slabik vždy některý vystupuje do popředí.“³⁸

³³ ŠVARNÝ, Oldřich a kol.: *Gramatika hovorovej čínštiny v príkladoch*. Bratislava: Vydavateľstvo Univerzity Komenského, 1993, s. 57.

³⁴ ŠVARNÝ, *Gramatika*, s. 24.

³⁵ ŠVARNÝ, *Gramatika*, s. 24.

³⁶ ČERNÝ, *Úvod*, s. 86.

³⁷ KRČMOVÁ

³⁸ KRČMOVÁ

„Slabika je nejmenší reálně existující artikulačně-akustická jednotka řeči. Jde o nejjednodušší a nejtěsnější možnou artikulační jednotu funkčních prvků řeči (hláska), která vyhovuje dorozumívání.“³⁹

Z výše uvedených definic je patrné, že jazykovědci zastávají dvojí názor. Pro některé je nejmenší jazykovou (fonetickou) jednotkou hláska, pro jiné slabika.

5.1.5 Hláska

Hláska je definována jako „elementární zvukový segment řeči, základní jednotka fonetického popisu jazyka. Je určitým stupněm zobecnění nesmírně proměnlivé zvukové reality lidské řeči.“⁴⁰

Podle Hály jsou hlásky „plynulými řadami zvuků, jež představují základní (elementární) a dále už nedělitelné útvary řeči.“⁴¹

5.2 Výběrový soubor C

Výběrové soubory A a B se od výběrového souboru C odlišují také členěním textu. Jak uvádí Švarný, délka segmentu v čínštině v průměru osciluje mezi 2,5 – 4,5 slabiky, přičemž v rychlejší řeči se délka segmentu prodlužuje. Délka kól se pak pohybuje v rozpětí 6 – 7 slabik.⁴²

Vzorek A a B (především pak vzorek A) jsou specifické rychlým tempem řeči. Přízvukový takt je dlouhý v průměru 5 slabik a stojí tak na pomezí mezi segmentem a kólonem. Hranice mezi segmentem a kólonem se tak stírají.

V kontrastu k předcházejícím dvěma souborům stojí výběrový soubor C. Jak jsem zmínila výše, jsou to nahrávky pořízené k učebnici GHČP. Vzhledem k faktu, že se jedná o učební materiál a to především pro studenty čínštiny – začátečníky, je tomu

³⁹ KARLÍK, *Encyklopedický*, s. 404.

⁴⁰ KARLÍK, *Encyklopedický*, str. 169.

⁴¹ HÁLA, Bohuslav. *Fonetika v teorii a v praxi*. Praha: Státní pedagogické nakladatelství, 1975, 482 s., s. 109.

⁴² ŠVARNÝ, *Gramatika*, s. 24.

také přizpůsobena promluva mluvčího. Jedná se výhradně o čínský text, který není spontánní, ale je pronášen tak, jak by měl v korektním jazyce znít. V nahrávce výběrového souboru C se neobjevují žádná cizí slova, vulgarismy, ani hezitační zvuky.

Jak jsem zmínila výše, námi zkoumané jazykové hladiny byly: hláska, slabika, přízvukový takt, výpověď a promluva. Švarného členění textu se oproti tomu našemu poměrně liší. V následující části bych proto popsala vzorek C a způsob, jakým text Švarný rozčlenil a poukázala na rozdíly mezi jeho a naším členěním.

5.2.1 Kólon

Švarný dále souvětí a věty člení na kóla:

„tj. ucelené významové úseky s ucelenou rytmickou strukturou, zakončené přeryvem (tj. pauzou trvajícím zlomek až $\frac{3}{4}$ sec. nebo více) a charakterizované neukončující, tj. stoupavou modifikací poslední tónické slabiky. Přeryvy, tj. hranice kól, se z největší části objevují v souvětích na předělu vět (i když ne vždy) a nezřídka se objevují i po preverbálních jmenných větných členech (podmět, určení času) i po vazbách prepozičních sloves s předmětem; objevují se i mezi jednotlivými jmennými členy výřtů. Méně často se vyskytují mezi predikativem (nebo predikativním komplexem) a postverbálními jmennými členy (příležitostně tehdy, je-li některý, případně oba komplexy, poměrně dlouhý).“⁴³

5.2.2 Segment

Jednotkou nižší než kólon je pak segment:

„kóla, pokud nejsou jednoslabičná nebo velmi krátká, se podle svého lexikálního složení a gramatické stavby rozpadají na menší skupiny slabik... tyto sledy mohou zůstat jako samostatnější jednotky, segmenty,

⁴³ ŠVARNÝ, *Gramatika*, s. 24.

nebo se navzájem i různě kombinují a těsněji spojují, takže vznikají segmenty složitější.“⁴⁴

Švarný tedy dělí text na hlásky, slabiky, segmenty, kóla, věty.

Vzorek C tedy reprezentuje umělý text, při jehož realizaci se klade snaha o co nejpřesnější odraz čínštiny tak, jak by opravdu měla být realizována. Srovnání těchto dvou, respektive tří promluv se pak odráží na výsledku, kdy se navíc ukazuje, že v rychlejší řeči se zvyšuje počet taktů, tzn. v námi segmentovaném textu jsou takty v průměru delší než segmenty u Švarného, a je to dáno právě tím, že oba mluvčí mluví velmi rychle.

Soubory A a B jsme si vybraly především proto, že jsou vhodné zejména pro srovnání, které jsme chtěly realizovat (a to srovnání takové, zda hudební žánr ovlivní tendenci řeči). Soubory A a B jsou svou strukturou vhodné pro segmentaci na námi zvolené jazykové úrovni, i když u nejvyšší úrovni je to diskutabilní. U vzorku A a B mluvčí odpovídají na kladené otázky, svým způsobem je to ale kompaktní promluva, sjednocená především tématem.

Vzorek C jsem si pro svou diplomovou práci vybrala proto, že jsem chtěla srovnat měření aplikovaná na přirozeném, respektive co nejpřirozenějším, a umělém textu.

⁴⁴ ŠVARNÝ, *Gramatika*, s. 24.

6. JAZYKOVÉ ÚROVNĚ

V průběhu projektu jsme zkoumaly tři jazykové úrovně:

1, jazykovou úroveň tvořenou replikou (konstrukt, jehož délka je měřená v počtu výpovědí) a výpovědí (konstituent, jehož délka je měřená v průměrném počtu přízvukových taktů)

2, jazykovou úroveň tvořenou výpovědí (konstrukt, jehož délka je měřená v počtu přízvukových taktů) a přízvukovým taktem (konstituent, jehož délka je měřená v průměrném počtu slabik)

3, jazykovou úroveň tvořenou přízvukovým taktem (konstrukt, jehož délka je měřená v počtu slabik) a slabikou (konstituent, jehož délka je měřená v průměrném počtu hlásek)

V našich diplomových pracích se zaměřujeme především na nejnižší a střední úroveň, nepovažovaly jsme proto za nutné u výběrového souboru C provést výpočty na nejvyšší úrovni. Vzorek C byl pro zkoumání nejvyšší úrovně absolutně nevhodný. Jak jsem již avizovala výše, vzorek C obsahuje oproti vzorku A a B tyto jazykové jednotky: hláska, slabika, segment, kólon, věta. Na vzorku C jsme tedy zkoumaly následující jazykové úrovně:

1, jazykovou úroveň tvořenou kólonem (konstrukt, jehož délka je měřená v počtu segmentů) a segmentem (konstituent, jehož délka je měřená v průměrném počtu slabik)

2, jazykovou úroveň tvořenou segmentem (konstrukt, jehož délka je měřena v počtu slabik) a slabikou (konstituentem, jehož délka je měřena v průměrném počtu hlásek).

7. PŘEPIS VÝBĚROVÝCH SOUBORŮ

Po selekci výběrových souborů následoval jejich přepis do písemné podoby. Rozhodovaly jsme se mezi českou transkripcí a standardní čínskou transkripcí pinyin. S ohledem na fakt, že jednou z námi stanovených jazykových hladin byla hláska, rozhodly jsme se výběrové soubory přepsat do české transkripce čínštiny vytvořené prof. Švarným. Při rozhodování jsme zohlednily především pohodlnost počítání hlásek z textu přepsaného do české transkripce, protože právě ta odráží jednotlivé fonémy co nejdříveji.

Při segmentaci textu jsme narazily na několik problémů spojených především s počítáním hlásek, respektive s tím, zda některé hlásky počítat jako jednu nebo dvě. Níže tedy představuji nejčastější problémy a demonstruji je jak na pinyinu, tak na české transkripci.

7.1 Kontrast mezi zhe (če), she (še) vs. zhi (č'), shi (š')⁴⁵

Česká transkripce zachycuje zvukovou realizaci výše uvedených slabik přesněji. Při zvukové realizaci slabiky *zhe* a *zhi*, jsou v obou případech produkovány dvě hlásky. Zatímco pinyin tyto hlásky zaznamenává pomocí tří písmen, česká transkripce používá písmena dvě anebo písmeno a apostrof (v našem chápání je apostrof rovnocenný písmenu), což je pro nás při zpracovávání souborů rozhodně výhodnější. Nepopíráme, že se slabiky *zhe* a *zhi* liší kvalitou samohlásky. Vzhledem k tomu, že během našeho výzkumu pracujeme s počtem hlásek ve slabice a ne s délkou hlásek, počítáme u těchto slabik vždy dvě hlásky

7.2 Souhlásky neaspirované a aspirované

Pinyin zapisuje neznělé souhlásky *p*, *t*, *k* pomocí znělých variant *b*, *d*, *g* a neznělé souhlásky *p*, *t*, *k* pak používá pro zápis jejich aspirovaných variant. Oproti tomu česká transkripce zapisuje neznělé souhlásky pomocí neznělých souhlásek a u

⁴⁵ Varianta bez závorčky je uvedena v pinyin, v závorce jsou pak slabiky přepsány v české transkripci.

aspirovaných variant pak přidává písmeno *ch*, např. neaspirované *ban* (*pan*) a aspirované *pan* (*pchan*).

7.3 Frikativy

Frikativy *sh* a *zh* jsou v pinyině přepisovány pomocí dvou písmen, a to i přesto, že odrážejí pouze jednu hlásku. Česká transkripce, která používá k zápisu pouze jedno písmeno *š* a *č*, odráží tento fenomén lépe. V tomto případě je tedy česká transkripce také vhodnější.

7.4 Iniciála *q*

Iniciála *q* je jediný případ, kdy je pinyin efektivnější než česká transkripce. Pinyin zaznamenává tento zvuk pomocí jednoho písmena, česká transkripce jej zaznamenává pomocí tří písmen (*čch*).

7.5 Samohlásky

Poslední problém, který se v souvislosti s přepisem souboru objevil, se týká trojhlásek *ui* a *iu*. I když jsou tyto trojhlásky zvukově realizovány jako tři fonémy, pinyin je zachycuje pomocí dvou písmen. Česká transkripce používá k přepisu těchto trojhlásek písmena tři, proto je v tomto případě víc efektivní než pinyin.

Níže uvádím tabulku, v které srovnávám výhody počítání hlásek z textu přepsaném v české transkripci a textu přepsaném v pinyině. Z tabulky jasně vyplývá česká transkripce jako lepší volba.

Tabulka 1: Přehled foneticky problematických aspektů⁴⁶

problematický aspekt	pinyin (počet písmen/počet hlásek)	česká transkripce (počet písmen/počet hlásek)
finála i vs. e	<i>zhi</i> (3/2) <i>zhe</i> (3/2)	<i>č'</i> (2/2) <i>če</i> (2/2)
souhlásky aspirované a neaspirované	<i>ban</i> (3/3) <i>pan</i> (3/4)	<i>pan</i> (3/3) <i>pchan</i> (4/4)
frikativy	<i>zhe</i> (3/2)	<i>če</i> (2/2)
iniciála 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)

Jak je patrné z Tabulky 1 a z textu nad ní, ve 4 z 5 případů je pro naše účely efektivnější česká transkripce a v jednom případě pak transkripce pinyin. Z tohoto důvodu jsme se rozhodly přepsat texty výběrových souborů do české transkripce.

⁴⁶ 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) s. 4.

8. PREZENTACE VÝSLEDKŮ

Následující část mé diplomové práce je věnována analýze dosažených výsledků, demonstraci tabulek a grafů, jejich komentáři a zamyšlení se nad možnými příčinami a vysvětlením výsledků se zvláštním zřetelem na jazykovou úroveň tvořenou přízvukovým taktem/segmentem (konstrukt, jehož délka je měřena v počtu slabik) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek).

Nejprve uvádím výsledky, kterých bylo dosaženo v první fázi výzkumu, poté následují výsledky z druhé fáze výzkumu, tedy výsledky, kterých bylo dosaženo aplikací MALu na zvukových stopách v GHČP. U každé jazykové úrovně nejprve uvádím tabulky, dále grafy vytvořené pomocí programu R (u jazykové úrovně tvořené přízvukovým taktem/segmentem – konstrukt, jehož délka je měřena v počtu slabik – a slabikou – konstituent, jehož délka je měřena v průměrném počtu hlásek navíc přidávám grafy vytvořené pomocí programu Microsoft Excel), pod kterými hodnotím výsledky, které jsou demonstrovány právě v tabulkách a grafech jednotlivých výběrových souborů. Jako poslední následuje diskuse nad dosaženými výsledky.

Výsledky uvádím postupně od nejvyšší jazykové úrovně tvořené replikou (konstrukt, jehož délka je měřena v počtu výpovědí) a výpovědí (konstituent, jehož délka je měřena v průměrném počtu přízvukových taktů), přes jazykovou úroveň tvořenou výpovědí (konstrukt, jehož délka je měřena v počtu přízvukových taktů) a přízvukovým taktem (konstituent, jehož délka je měřena v průměrném počtu slabik) až po nejnižší jazykovou úroveň tvořenou přízvukovým taktem (konstrukt, jehož délka je měřena v počtu slabik) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek).

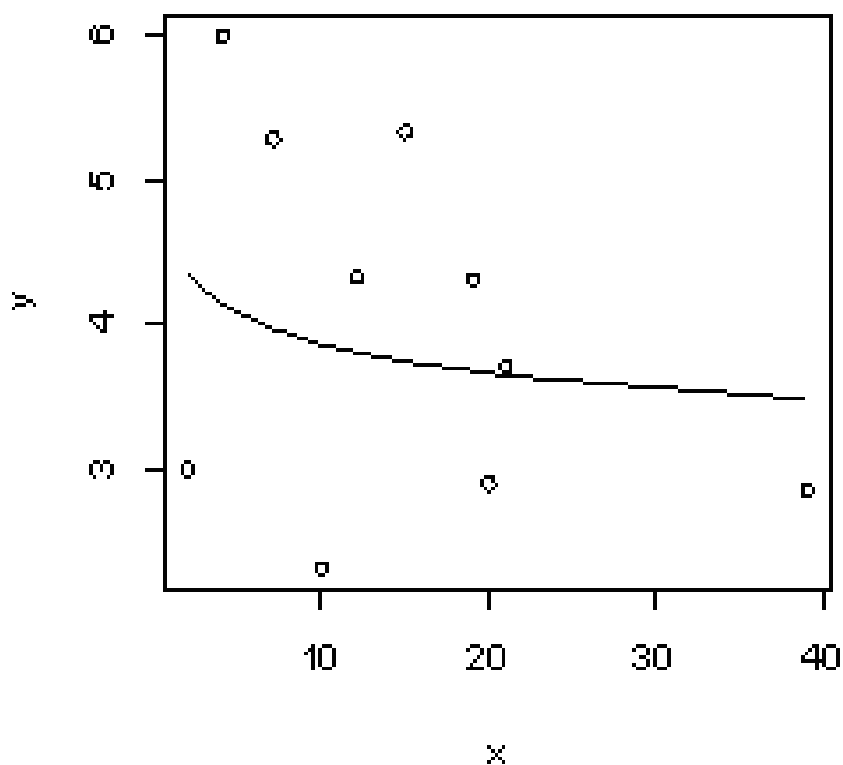
8.1 Jazyková úroveň U1: replika (konstrukt, jehož délka je měřena v počtu výpovědí) a výpověď (konstituent, jehož délka je měřena v průměrném počtu přízvukových taktů)

Na nejvyšší jazykové úrovni byl zkoumán vzájemný vztah mezi konstruktem, tj. replikou a konstituentem, tj. výpovědí. Výpověď zároveň v následující nižší jazykové úrovni vystupuje jako konstrukt vůči svému konstituentu, tedy přízvukovému taktu. Výrazným rysem nejvyšší jazykové úrovně byl nízký počet konstruktů, což mohlo s velkou pravděpodobností ovlivnit konečné výsledky. O tomto problému se však zmiňuji níže v diskusi nad výsledky.

Tabulka 2: U1, výběrový soubor A: konstrukty x_1 – délka repliky (měřena v počtu výpovědí), z_1 – frekvence konstruktů, konstituenty y_1 – průměrná délka výpovědi (měřena v počtu přízvukových taktů)

délka repliky ve výpovědích x_1	frekvence z_1	průměrná délka výpovědi v přízvukových takttech y_1
2	2	3
4	1	6
7	1	5,28571429
10	1	2,3
12	1	4,33333333
15	1	5,33333333
19	1	4,31578947
20	1	2,9
21	1	3,71428571
39	1	2,84615385

Jak lze vyčíst z Tabulky 2, nejvyšší úroveň se vyznačuje velmi nízkou frekvencí konstruktů. V případě výběrového souboru A je jich konkrétně 11. Délka replik se pohybuje od 2 – 39 výpovědí, průměrná délka výpovědi měřená v přízvukových takttech pak osciluje mezi 2,3 – 6.



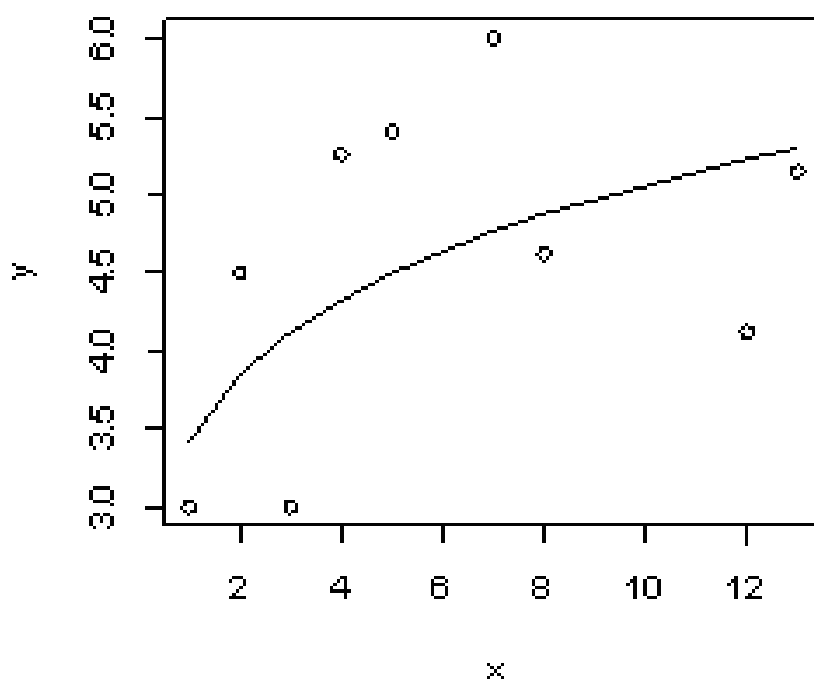
Graf 1: U1 replika vs. výpověď – soubor A: vizualizace souboru dat v Tabulce 2, pomocí programu R⁴⁷

Tabulka 3: U1, výběrový soubor B, verze 1: konstrukty x_1 – délka repliky (měřena v počtu výpovědí), z_1 - frekvence konstruktů, konstituenty y_1 – průměrná délka výpovědi (měřena v počtu přízvukových taktů)

délka repliky ve výpovědích x_1	frekvence z_1	průměrná délka výpovědi v přízvukových taktch y_1
1	3	3
2	1	4,5
3	4	3
4	1	5,25
5	1	5,4
7	1	6
8	1	4,625
12	2	4,125
13	1	5,153846

⁴⁷ SCHUSTEROVÁ, *An Application*, s. 8.

Jak lze vyčíst z Tabulky 3, i výběrový soubor B verze 1 se vyznačuje nízkou frekvencí konstruktů, zde jich můžeme napočítat 15. Délka replik se zde pohybuje od 1 – 13 výpovědí. Průměrná délka výpovědi v přízvukových taktech se pak pohybuje v rozmezí 3 - 6.



Graf 2: U1 replika vs. výpověď – soubor B, verze 1: vizualizace souboru dat v Tabulce 3, pomocí programu R⁴⁸

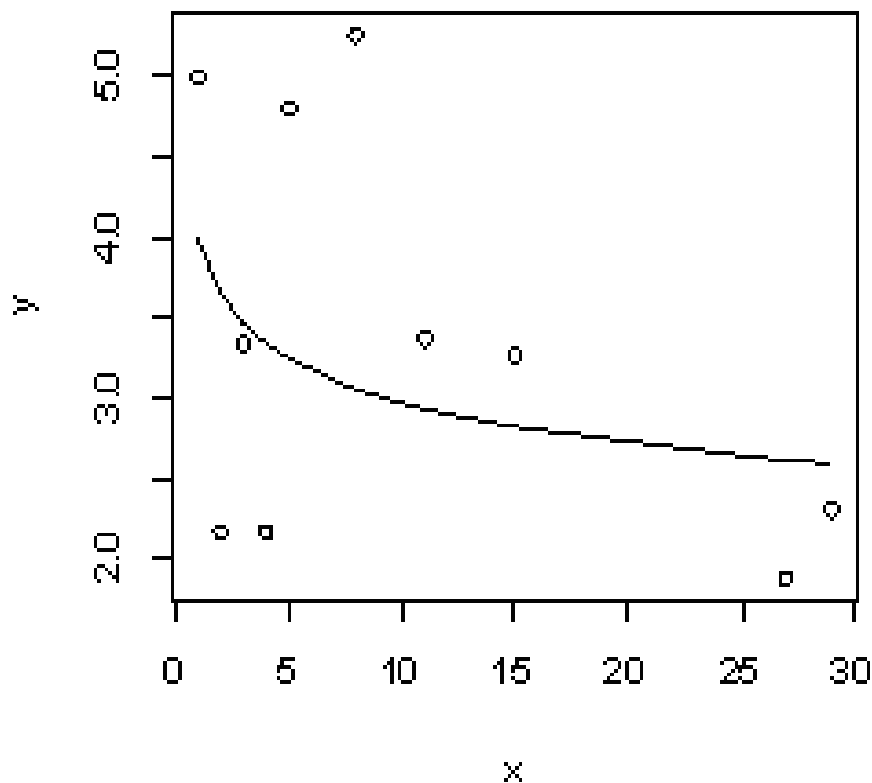
Tabulka 4: U1, výběrový soubor B, verze 2: konstrukty x_1 – délka repliky (měřena v počtu výpovědí), z_1 – frekvence konstruktů, konstituenty y_1 – průměrná délka výpovědi (měřena v počtu přízvukových taktů)

délka repliky ve výpovědích x_1	frekvence z_1	průměrná délka výpovědi v přízvukových taktech y_1
1	1	5
2	3	2,16666667
3	1	3,33333333
4	3	2,16666667

⁴⁸ SCHUSTEROVÁ, *An Application*, s. 9.

5	2	4,8
8	1	5,25
11	1	3,363636364
15	1	3,266666667
27	1	1,888888889
29	1	2,310344828

Z Tabulky 4 je patrné, že i u výběrového souboru B verze 2 se na nejvyšší úrovni setkáváme s nízkou frekvencí konstruktů, v tomto případě jich je rovněž 15. Délka výpovědi se pohybuje v rozmezí od 1 – 29 replik. Průměrná délka výpovědi v přízvukových taktech osciluje od 1,8 – 5,25.



Graf 3: U1 replika vs. výpověď – soubor B, verze 2: vizualizace souboru dat v Tabulce 4, pomocí programu R⁴⁹

Z Grafů 1, 2, a 3 lze vypožorovat, že na nejvyšší jazykové úrovni tvořené replikou (konstrukt, jehož délka je měřená v počtu výpovědí) a výpovědi (konstituent, jehož délka je měřená v průměrném počtu přízvukových taktů) se prokazuje pouze mírná

⁴⁹ SCHUSTEROVÁ, *An Application*, s. 9.

klesající tendence v případech 1. A a 1. B2. V případě 1.B1 se klesající tendence neprojevuje vůbec, naopak se zde prokázala tendence stoupající.

Tabulka 5: Přehled parametrů b MAL a koeficientů determinace na U1⁵⁰

		Parametr b	Koeficient determinace R^2 (%)
Výběrový soubor A		0.0748	4.16
Výběrový soubor B	Verze 1	-0.1717	34.1
	Verze 2	0.1279	13.97

Koeficienty determinace jsou pak tyto: 1. A $R^2 = 4.16\%$, 1. B1 $R^2 = 34.1\%$ a 1. B2 $R^2 = 13.97\%$. U výběrového souboru B verze 1 se vyskytla záporná hodnota parametru b. Z výše uvedených údajů a grafů vyplývá, že na nejvyšší jazykové úrovni se MAL neprojevil. Příčin pro takové výsledky může být hned několik.

Jednou z nich je nedostatečně vysoká frekvence konstruktů na této úrovni. Ani u jednoho z výběrových souborů jejich počet nedosahuje 20, konkrétně v souboru A jich je 11, v souboru B.1 jich je 15 a v souboru B.2 také 15. Jsem toho názoru, že problém podobné povahy však nastane na nejvyšší úrovni u většiny experimentů, a to z toho důvodu, že je složité najít text, který by se skládal z vysokého počtu promluv. Velice těžké je to pak u mluveného textu. V takovém případě je skutečně náročné najít text, který by byl realizován pouze jedním mluvčím. Pakliže takový text najdeme, vyvstává druhý problém, a to ten, že taková promluva není většinou realizována spontánně, přinejmenším je dopředu promyšlená, popř. naučená z paměti.

Jak už jsem ale naznačovala výše, výběrové soubory, které jsme zvolily, dostatečně splňovaly kritéria pro další, nižší jazykové úrovně. Jsme si tedy vědomy toho, že pro nejvyšší úroveň tento soubor nebyl zcela adekvátně zvolen, na druhou stranu to však byl nejvhodnější soubor z těch, které jsme měly k dispozici a pro nižší jazykové úrovně je naprosto vyhovující.

Další možná příčina by pak mohla spočívat v určení jednotek na této jazykové úrovni, respektive ve vynechání jedné z jednotek. Jak jsem se již zmiňovala výše, v

⁵⁰ SCHUSTEROVÁ, *An Application*, s. 10.

lingvistice existují mnohé různé definice jak pro pojem výpověď, tak pro pojem replika. U některých definic se hranice výpovědi a repliky překrývají. Jedna výpověď může být dokonce samostatná replika. Je proto dost možné, že kdybychom mezi výpověď a repliku vložily ještě jednu hladinu, výsledek by mohl být diametrálně odlišný. Během experimentu se prokázalo, že v budoucích výzkumech je nutné stanovit více jazykových jednotek jak na vyšších, tak na nižších jazykových úrovních.

Jako další příčinu, která se nám jeví pravděpodobnou, bych chtěla uvést sloučení syntaktického a fonetického kritéria při určování hranic a také nedostatečnou definici pojmu výpověď a replika.

Poslední možná příčina pak nemá nic společného s ostatními výše zmíněnými. Jak už jsem se zmínila dříve při popisu výběrových souborů, hudební pořad, který jsme zvolily jako výběrové soubory, je nastříhán a dále upravován pro potřeby jeho tvůrců. Tyto umělé zásahy mohou mít zcela jistě vliv na konečné výsledky našich výpočtů prováděných právě na souborech, jako jsou výběrové soubory A a B.

Vzhledem k výsledkům zmiňované nejvyšší jazykové úrovně si myslím, že tato úroveň si zaslouží další samostatný výzkum. MAL se na této úrovni nepotvrdil ani u jednoho z výběrových souborů. Můžeme jen odhadovat, co je příčinou takových výsledků. Domnívám se však, že je nutné dále provádět výzkum právě na této úrovni. Při pohledu na Grafy 1, 2 a 3 vidíme, že hodnoty jednotlivých pozorování jsou chaoticky rozmístěny a jen stěží z nich lze vyzpozorovat nějakou konkrétní tendenci. V budoucím výzkumu je nutné zvolit nové, zcela odlišné výběrové soubory a navýšit počet zkoumaných jazykových jednotek.

8.2 Jazyková úroveň U2: výpověď (konstrukt, jehož délka je měřena v počtu přízvukových taktů) a přízvukový takt (konstituent, jehož délka je měřena v průměrném počtu slabik)

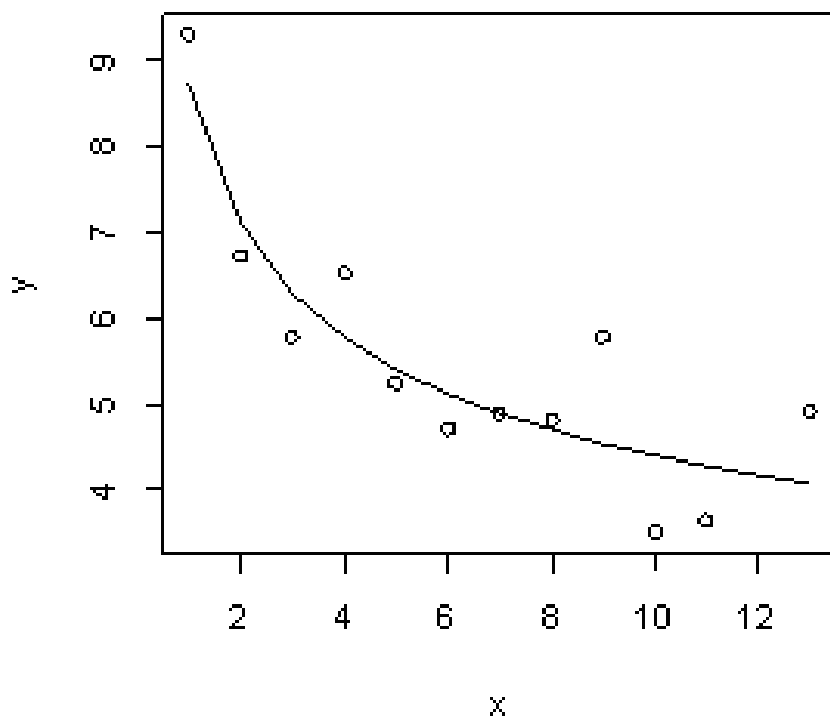
Na prostřední jazykové úrovni byl zkoumán vzájemný vztah mezi konstruktem, tj. výpovědí a konstituentem, tj. přízvukovým taktem. Přízvukový takt zároveň vystupuje v následující nejnižší jazykové úrovni jako konstrukt vůči svému konstituentu, tj. slabice. Jazyková úroveň výpověď vs. přízvukový takt je zajímavá

především tím, že je to dle výsledků našeho výzkumu úroveň, která je ovlivněna rytmem řeči. Dále se o tomto jevu zmiňuji v diskusi nad výsledky.

Tabulka 6: U2, výběrový soubor A: konstrukty x_2 – délka výpovědi (měřená v počtu přízvukových taktů), z_2 – frekvence konstruktů, konstituenty y_2 – průměrná délka přízvukového taktu (měřena v počtu slabik)

délka výpovědi v taktách x_2	frekvence z_2	průměrná délka přízvukového taktu v slabikách y_2
1	20	9,3
2	39	6,730769
3	26	5,794872
4	21	6,52381
5	14	5,257143
6	13	4,705128
7	7	4,897959
8	4	4,8125
9	2	5,777778
10	1	3,5
11	2	3,636364
13	2	4,923077

Z Tabulky 6 lze vyčíst, že délka výpovědi se pohybuje od 1 – 13 přízvukových taktů. Počet konstruktů je zde podstatně vyšší, než tomu bylo u předchozí úrovně, napočítáme jich zde 151. Délka přízvukového taktu ve slabikách se pohybuje v rozmezí 3,5 – 9,3.



Graf 4: U2 výpověď vs. přízvukový takt – soubor A: vizualizace souboru dat v Tabulce 6, pomocí programu R⁵¹

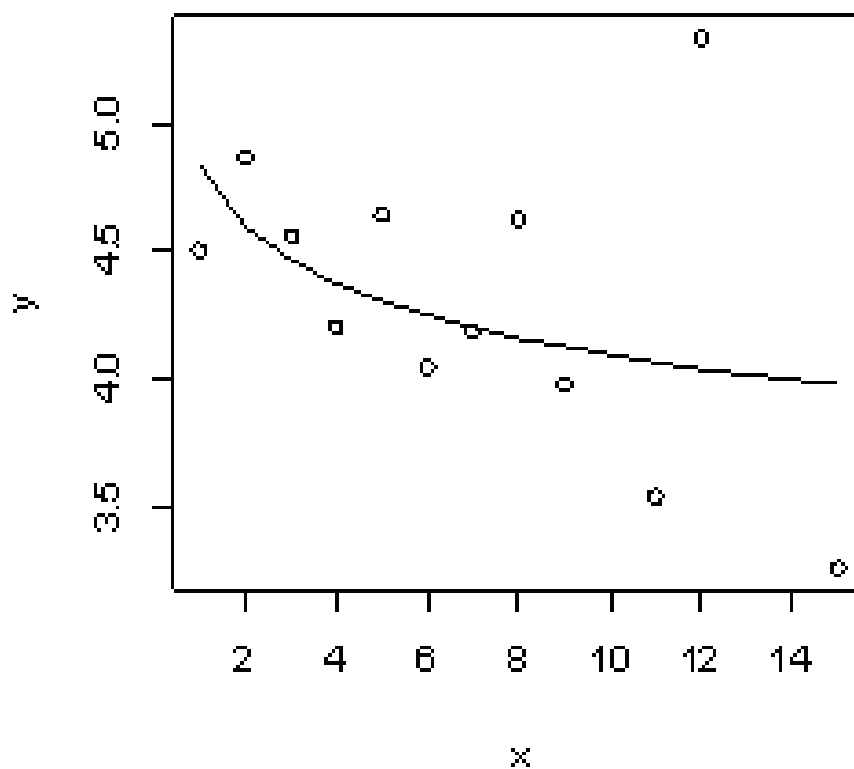
Tabulka 7: U2, výběrový soubor B, verze 1: konstrukty x_2 – délka výpovědi (měřená v počtu přízvukových taktů), z_2 – frekvence konstruktů, konstituenty y_2 – průměrná délka přízvukového taktu (měřena v počtu slabik)

délka výpovědi v takttech x_2	frekvence z_2	průměrná délka přízvukového taktu v slabikách y_2
1	2	4,5
2	15	4,8666667
3	15	4,5555556
4	17	4,2058824
5	5	4,64
6	10	4,05
7	3	4,1904762
8	1	4,625
9	6	3,9814815
11	1	3,5454545

⁵¹ SCHUSTEROVÁ, *An Application*, s. 10.

12	1	5,3333333
15	1	3,2666667

Z Tabulky 7 lze vypočítat, že u výběrového souboru B verze 1 nacházíme i výpověď o délce 15 přízvukových taktů. Počet konstruktů, tj. 77, je sice menší než u výběrového souboru A, je však vyšší než u předcházející nejvyšší jazykové úrovně, která je tvořena replikou (konstrukt, jehož délka je měřena v počtu výpovědí) a výpovědí (konstituent, jehož délka je měřena v průměrném počtu hlásek). Délka přízvukového taktu ve slabikách se pohybuje v rozmezí 3,26 – 5,33.



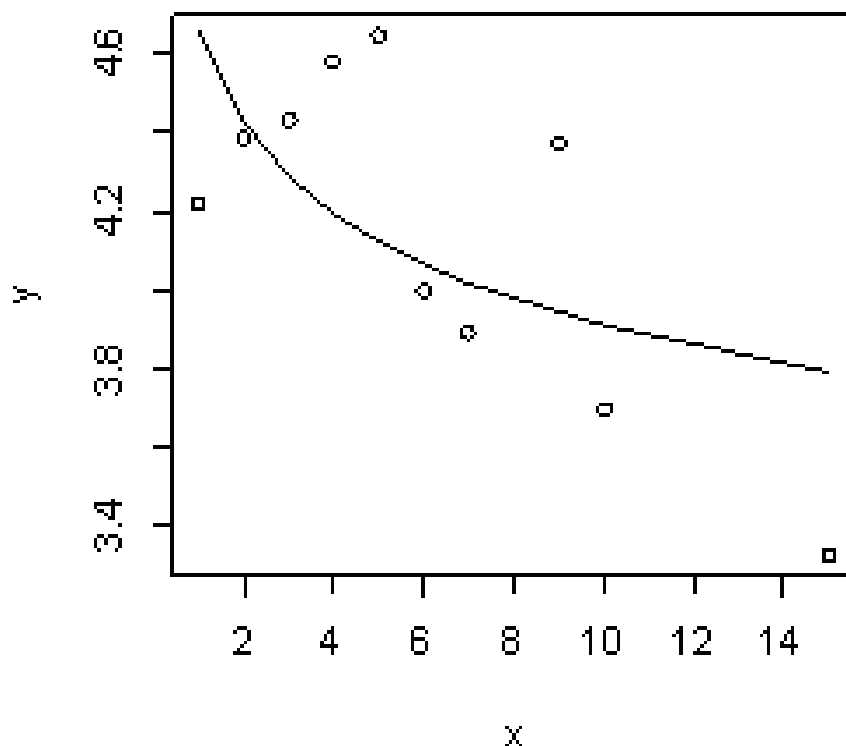
Graf 5: U2 výpověď vs. přízvukový takt – soubor B, verze 1: vizualizace souboru dat v Tabulce 7, pomocí programu R⁵²

⁵² SCHUSTEROVÁ, *An Application*, s. 11.

Tabulka 8: U2, výběrový soubor B, verze 2: konstrukty X_2 – délka výpovědi (měřená v počtu přízvukových taktů), Z_2 – frekvence konstruktů, konstituenty Y_2 – průměrná délka přízvukového taktu (měřena v počtu slabik)

délka výpovědi v taktach X_2	frekvence Z_2	průměrná délka přízvukového taktu v slabikách Y_2
1	37	4,216216
2	30	4,383333
3	18	4,425926
4	17	4,573529
5	5	4,64
6	4	4
7	4	3,892857
9	3	4,37037
10	1	3,7
15	1	3,333333

Z Tabulky 8 je očividné, že i u výběrového souboru B verze 2 se opět vyskytuje výpověď čítající 15 přízvukových taktů, počet konstruktů se však zde zvýšil na 120. Navýšení počtů konstruktů je zde dáno tím, že ve verzi 2 výběrového souboru B jsme považovaly vulgarismy za samostatné výpovědi. Oproti tomu ve verzi 1 jsme vulgarismy považovaly za součást předcházející výpovědi. Délka přízvukového taktu v slabikách se pohybuje v rozmezí 3,33 – 4,64.



Graf 6: U2 výpověď vs. přízvukový takt – soubor B, verze 2: vizualizace souboru dat v Tabulce 8, pomocí programu R⁵³

Tabulka 9: Přehled parametrů b MAL a koeficientů determinace na U2⁵⁴

		Parametr b	Koeficient determinace R^2 (%)
Vzorek A		0.2982	73.11
Vzorek B	Verze 1	0.2737	22.38
	Verze 2	0.0761	34.78

Tendence zobrazená v Grafu 4 je očividně klesající, jinými slovy, vztah mezi výpovědí jako konstruktem a přízvukovým taktem jako konstituentem je nepřímo úměrný. Procentuální shoda matematického modelu se získanými daty je $R^2 = 73,11\%$ (viz. Tab. 9). Mluvčí tohoto vzorku je popová zpěvačka, jejíž řeč není patrně ovlivněná hudebním žánrem, kterému se věnuje. Z Grafu 5 lze vidět, že u

⁵³ SCHUSTEROVÁ, *An Application*, s. 11.

⁵⁴ SCHUSTEROVÁ, *An Application*, s. 12.

výběrového souboru B verze 1 se projevuje pouze velmi mírná klesající tendence a z Grafu 6 pak vidíme, že u výběrového souboru B verze 2 se klesající tendence neprojevuje vůbec. Z Grafu 6 navíc vidíme, že tendence nejprve stoupá a poté klesá.

Možným vysvětlením pro tyto výsledky je u výběrového souboru B verze 1 fakt, že mluvčím tohoto souboru je rapový zpěvák, jenž je hudebním žánrem, který zpívá, značně ovlivněn. Už ze samotného poslechu jeho řeči lze rozpoznat, že řeč je pronášená v určitém rytmu charakteristickém pro rapovou hudbu. Hudební žánr, kterým se dotyčný mluvčí zabývá, zcela jistě ovlivnil styl jeho vyjadřování a to mohlo mít také dopad na výsledky našeho výzkumu.

V obou případech lze shodu považovat za nedostatečnou. Pro modely zobrazené v Grafech 5 a 6 jsou koeficienty determinace tyto: $R^2 = 22,38\%$ a $R^2 = 34,78\%$ (viz. Tab. 9). Domníváme se, že důvod proč se u jazykové úrovně tvořené výpovědí (konstrukt, jehož délka je měřena v počtu přízvukových taktů) a přízvukovým taktem (konstituent, jehož délka je měřen v průměrném počtu slabik) u výběrového souboru A MAL projevil a u výběrového souboru B neprojevil je ten, že zmiňovaná úroveň, je úrovní, která je ovlivňována rytmem řeči. Touto jazykovou úrovní se ve své diplomové práci zabývá kolegyně Ščigulinská.⁵⁵

8.3 Jazyková úroveň U3: přízvukový takt (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek)

Na nejnižší jazykové úrovni byl zkoumán vzájemný vztah mezi konstruktem, tj. přízvukovým taktem a konstituentem, tedy slabikou, jež je zároveň konstruktem vůči svému konstituentu, tj. hlásce. Na nejnižší jazykové úrovni přízvukový takt vs. slabika se vyskytnul zajímavý fenomén. Původní hypotéza našeho výzkumu byla taková, že se MAL projeví na všech zkoumaných jazykových úrovních současné mluvené čínštiny. Výsledky první fáze výzkumu však ukázaly, že nejnižší jazyková úroveň přízvukový takt vs. slabika MAL neodráží. Ukázalo se, že se na zmiňované

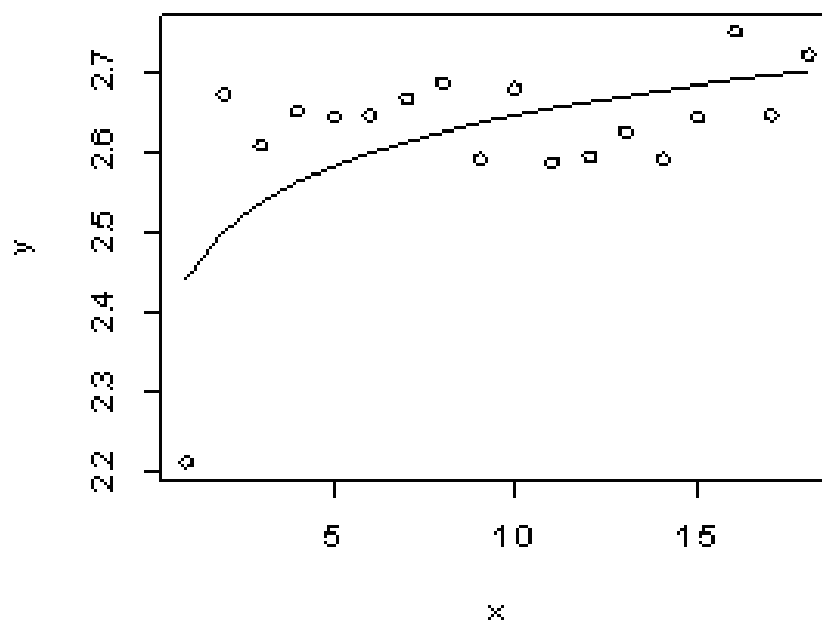
⁵⁵ ŠČIGULINSKÁ, Jana. *Aplikácia Menzerath-Altmanovho zákona na súčasnú hovorenú čínštinu na úrovni výpovede*. Olomouc, 2013. Magisterská práce (Mgr.). UNIVERZITA PALACKÉHO V OLOMOUCI. Filozofická fakulta

úrovni projevuje lineární tendence, respektive že hodnoty bodů obou výběrových souborů A a B na nejnižší jazykové úrovni oscilují v rámci pásu.

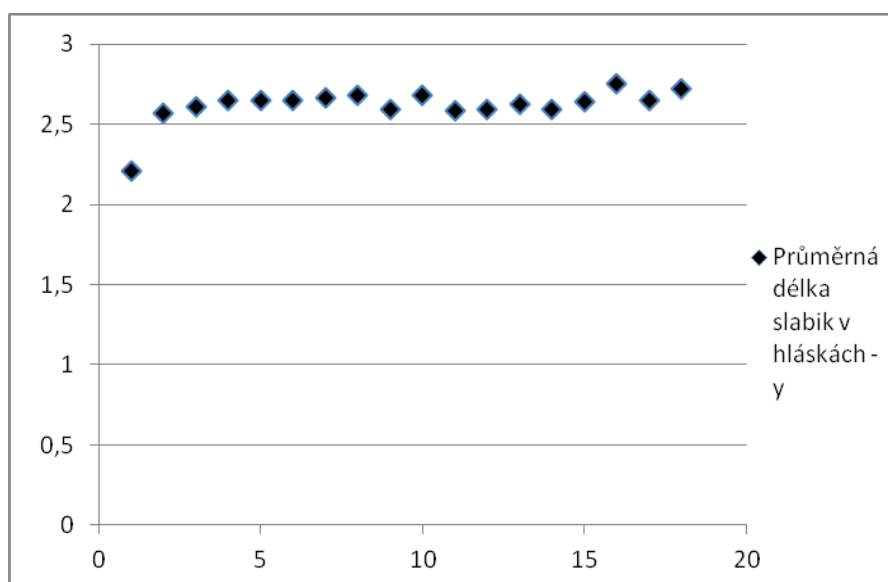
Tabulka 10: U3, výběrový soubor A: konstrukty x_{α} – délka přízvukových taktů (měřena v počtu slabik), z_{α} – frekvence konstruktů, konstituenty y_{α} – průměrná délka slabiky (měřena v počtu hlásek)

délka přízvukových taktů v slabikách x_{α}	frekvence z_{α}	průměrná délka slabik v hláskách y_{α}
1	19	2,210526316
2	76	2,565789474
3	73	2,607305936
4	69	2,652173913
5	84	2,645238095
6	65	2,646153846
7	48	2,666666667
8	36	2,684027778
9	33	2,592592593
10	24	2,679166667
11	9	2,585858586
12	8	2,59375
13	9	2,623931624
14	7	2,591836735
15	3	2,644444444
16	2	2,75
17	2	2,647058824
18	1	2,722222222

Z Tabulky 10 je patrné, že ojediněle délka některých přízvukových taktů dosáhla až 18 slabik. Počet konstruktů u výběrového souboru A je 568, tedy mnohonásobně vyšší, než tomu bylo u nejvyšší jazykové úrovně tvořené replikou (konstrukt, jehož délka je měřena v počtu výpovědí) a výpovědí (konstituent, jehož délka je měřena v průměrném počtu přízvukových taktů). Délka slabiky v hláskách se pohybuje v rozmezí 2,21 – 2,75.



Graf 7: U3 přízvukový takt vs. slabika – soubor A: vizualizace souboru dat v Tabulce 10, pomocí programu R⁵⁶



Graf 8: U3 přízvukový takt vs. slabika – soubor A: vizualizace souboru dat v Tabulce 10, pomocí programu Microsoft Excel

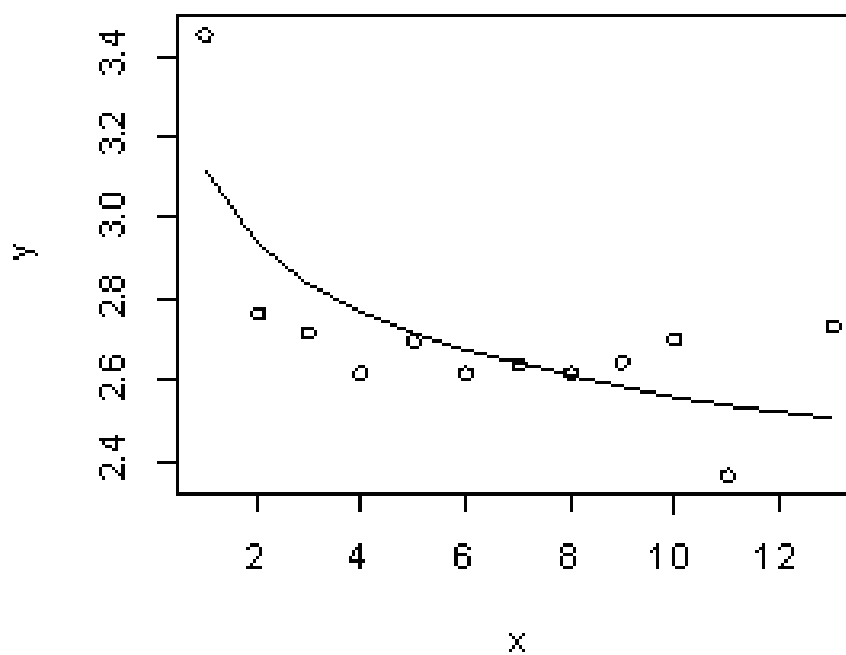
⁵⁶ SCHUSTEROVÁ, *An Application*, s. 13.

U této úrovně uvádím grafy dva. Z Grafu 7 lze vidět, že tendence je stoupající, z Grafu 8 je však lépe patrná zmiňovaná oscilace v rámci pásu s výjimkou prvního bodu, který viditelně vychyluje. Tato odchylka je způsobená právě množstvím hezitačních zvuků, jež jsou pro promluvu mluvčí výběrového souboru A charakteristické.

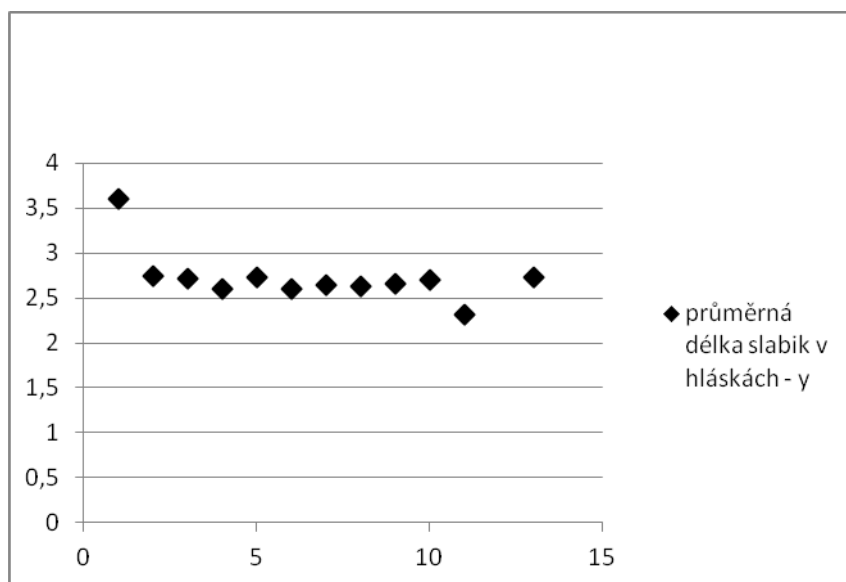
Tabulka 11: U3, výběrový soubor B: konstrukty x_{a} – délka přízvukových taktů (měřena v počtu slabik), z_{a} – frekvence konstruktů, konstituenty y_{a} – průměrná délka slabiky (měřena v počtu hlásek)

délka přízvukových taktů v slabikách x_{a}	frekvence z_{a}	průměrná délka slabik v hláskách y_{a}
1	10	3,6
2	82	2,75
3	47	2,716312057
4	63	2,603174603
5	48	2,725
6	37	2,603603604
7	25	2,64
8	20	2,63125
9	8	2,652777778
10	3	2,7
11	2	2,318181818
13	2	2,730769231

U výběrového souboru B dosáhla délka přízvukového taktu v některých případech 13 slabik, počet konstruktů je pak 347. Délka slabiky v hláskách se pohybuje v rozmezí 2,31 -3,6.



Graf 9: U3 přízvukový takt vs. slabika – soubor B: vizualizace souboru dat v Tabulce 11, pomocí programu R⁵⁷



Graf 10: U3 přízvukový takt vs. slabika – soubor B: vizualizace souboru dat v Tabulce 11, pomocí programu Microsoft Excel

⁵⁷ SCHUSTEROVÁ, *An Application*, s. 13.

V tomto případě rovněž uvádím grafy dva. Z prvního grafu (Graf 9) je patrná mírná klesající tendence, když se však podíváme na druhý graf (Graf 10), můžeme si opět všimnout oscilace v rámci pásu, kde opět značně vychyluje první bod. V případě výběrového souboru B je odskok prvního bodu způsoben cizími slovy, kterých mluvčí během své promluvy používá.

Tabulka 12: Přehled parametrů b MAL a koeficientů determinace na U3⁵⁸

	Parametr b	Koeficient determinace R^2 (%)
Vzorek A	-0.0349	37.53
Vzorek B	0.0855	56.65

Jak je patrné z Grafů 7, 8, 9, 10 oba výsledky mají podobnou, konstantní tendenci, s výše zmiňovanými výjimkami bodů $x_3=1$. Domnívám se, že vysvětlení pro takový výsledek se nachází v struktuře čínštiny coby slabičného jazyka. Jak jsem již uváděla výše, čínština disponuje omezeným setem slabik, kdy mimo rámec tohoto užívaného setu nevznikají již v čínštině žádné nové slabiky.

Proto se v obou případech, jak u výběrového souboru A, tak u výběrového souboru B, jedná o oscilaci v rámci pásu. Jak lze vidět z Tabulky 10 hodnota bodů y_3 se u výběrového souboru A pohybuje v rozmezí 2,56 - 2,72, v případě výběrového souboru B, jak lze vyčíst z Tabulky 11 se pak pohybuje v rozpětí 2,60 - 2,75, s výjimkami bodů $x_3 = 1$ u obou výběrových souborů a $x_3 = 11$ u výběrového souboru B. Tyto výjimky mohou být způsobeny vnějšími faktory. U výběrového souboru A se jedná o několik případů, kdy mluvčí svou promluvu proložila hezitačními zvuky či gramatickými částicemi, které jsou v těchto konkrétních případech tvořeny jednou hláskou. U vzorku B můžeme pozorovat výkyv způsobený nemalým počtem cizích (anglických) slov, kterými mluvčí svou řeč velmi často obohacuje. Používá je především z toho důvodu, aby vyjádřil svou příslušnost k raperské komunitě. Vzhledem k tomu, že anglická slova nejsou součástí lexika čínštiny, považuji tento jev za narušení zpěvákovy promluvy. Angličtina na rozdíl od

⁵⁸ SCHUSTEROVÁ, *An application*, s. 14.

čínštiny není slabičný jazyk, a proto i struktura anglických slov se od struktury čínských slov liší. Časté užívání anglických slov tak narušuje lineární tendenci. Vychýlení u bodu $x_{\text{a}} = 11$ je způsobeno sice nízkou, ale přesto přítomnou frekvencí jedenáctislabičných přízvukových taktů.

Z porovnání Grafů 7, 8, 9 a 10 tedy vyplývá, že MAL se na nejnižší jazykové úrovni, která je tvořená přízvukovým taktem (konstrukt, jehož délka je měřena ve slabikách) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek) neprojevil pravděpodobně kvůli tomu, že je čínština slabičný jazyk.

8.4 Jazyková úroveň U3Š: segment (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek)

Dále uvádím výsledky na nejnižší jazykové úrovni z druhé fáze výzkumu, tj. výsledky segmentace výběrového souboru C. Jelikož zkoumané jazykové jednotky výběrového souboru C se odlišují od jazykových jednotek jazykové úrovně U3 výběrových souborů A a B, zvolila jsem pro tuto úroveň pojmenování U3Š.

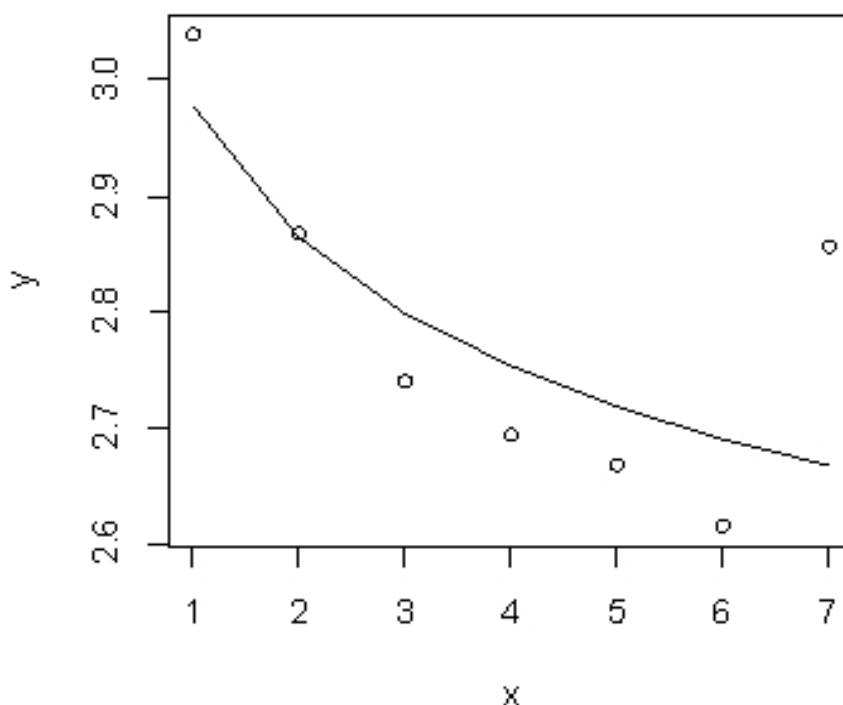
Na této jazykové úrovni byl zkoumán vztah mezi konstruktem, tj. segmentem a konstituentem, tedy slabikou, která zároveň vystupuje jako konstrukt vůči svému konstituentu, tj. hlásce. Jak už jsem zmiňovala výše, na samém počátku výzkumu stála hypotéza, že MAL se na nejnižší jazykové úrovni projeví. Tato hypotéza však po první části výzkumu potvrzena nebyla. Podle výpočtů provedených na výběrových souborech A a B se ukázalo, že na nejnižší jazykové úrovni tvořené přízvukovým taktem (konstrukt, jehož délka je měřena v počtu slabik) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek) se projevuje oscilace v rámci pásu. Hypotéza pro druhou fázi výzkumu byla ta, že na nejnižší jazykové úrovni se oscilace v určitém pásu projeví vždy.

Výpočty provedené v druhé fázi výzkumu na výběrovém souboru C však přinesly zajímavé výsledky. Křivka prokazuje klesající tendenci.

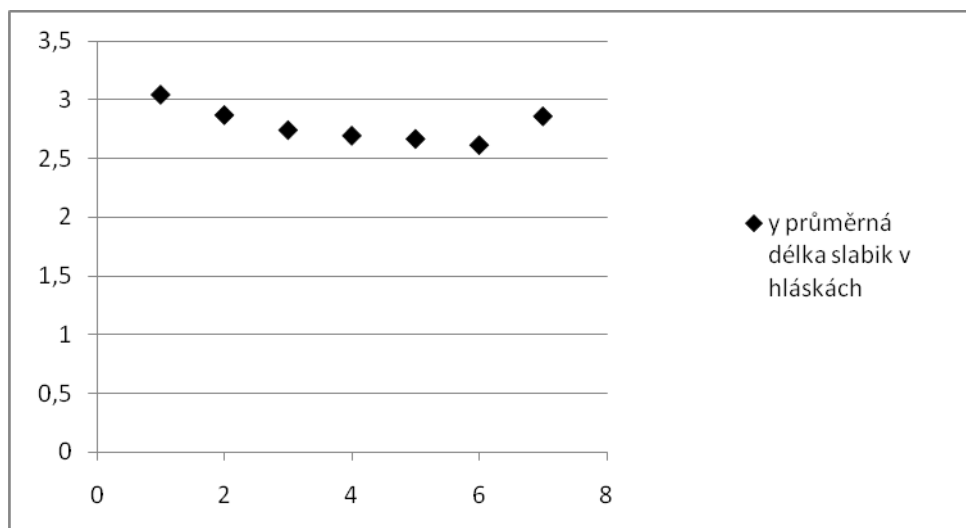
Tabulka 13: U3Š, výběrový soubor C: konstrukty x_a – délka segmentu (měřena v počtu slabik), z_a – frekvence konstruktů, konstituenty y_a – průměrná délka slabiky (měřena v počtu hlásek)

délka segmentu v slabikách x_a	frekvence z_a	průměrná délka slabiky v hláskách y_a
1	51	3,039215686
2	430	2,86744186
3	351	2,740740741
4	222	2,693693694
5	69	2,666666667
6	19	2,614035088
7	2	2,857142857

Z Tabulky 13 lze vyčíst, že délka segmentu dosáhla v některých případech 7 slabik. Počet konstruktů na této úrovni je pak 1144. Výběrový soubor C tedy disponuje nejvyšším počtem konstruktů na nejnižší jazykové úrovni ze všech tří souborů. Průměrná délka slabiky v hláskách se pohybuje v rozmezí 2,61 - 3,03, tedy v o trochu vyšším rozmezí, než tomu je u výběrových souborů A a B.



Graf 11: U3Š segment vs. slabika – soubor C: vizualizace souboru dat v Tabulce 13, pomocí programu R



Graf 12: U3Š segment vs. slabika – soubor C: vizualizace souboru dat v Tabulce 13, pomocí programu Microsoft Excel

Z Grafů 11 a 12 je očividná klesající tendence, která potvrzuje shodu s MALem.

Tabulka 14: Přehled parametru b MAL a koeficientu determinace na U3Š

	Parametr b	Koeficient determinace R^2 (%)
Vzorek C	0,0569	55,6

V prezentaci výsledků nejnižší jazykové úrovně u výběrových souborů A a B poukazují na oscilaci v rámci pásu. Po první fázi výzkumu jsem byla přesvědčena o tom, že ať budu zkoumat jakýkoli text, ať bude mluvčím kdokoli, výsledek bude pokaždé téměř identický. Následné výpočty provedené na přepisu promluvy z GHČP však tuto mou hypotézu nepotvrdily.

Křivka v Grafu 11 velmi pěkně kopíruje tendenci MALu. Chtěla bych však především upozornit na procentuální shodu s MALem, která je, jak lze vyčíst z Tabulky 14, $R = 55,6\%$. Při pohledu na Graf 11 je patrná jasná tendence MALu, která je narušená pouze posledním bodem $x_7 = 7$, který očividně odskakuje. U tohoto bodu se jedná o sedmi-slabičné segmenty, které jsou v celém vybraném vzorku výjimkou, o tom svědčí i velmi nízká frekvence výskytu, tj. 2. Při pohledu na graf 12 pak také vidíme klesající tendenci, i když zde není tak zřetelná jako u Grafu 11.

U výběrových souborů A a B tvořily nejnižší jazykovou úroveň přízvukový takt (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek). Právě na této úrovni se projevila oscilace v rámci pásu. U výběrového souboru C pak nejnižší jazykovou úroveň tvořily segment (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek). Tato úroveň však prokázala očividnou shodu s MALem.

Mé tvrzení, že nejmenší jazyková úroveň neodráží MAL z důvodu slabičnosti čínštiny tak bylo narušeno výsledky výpočtů provedených na vzorku C, kdy se potvrdila očividná tendence s MALem.

V tomto bodě však považuji za nutné znovu připomenout zásadní rozdílnost výběrových souborů. Především pak výběrový soubor C, kdy se jedná o velmi nestandardní soubor. Právě to, že mluvená řeč zde není přirozená, může mít velký vliv na tak rozdílné výsledky.

Ale také samotné jednotlivé oddíly, tj. zvukové stopy, z GHČP, respektive jejich neintegrita, může mít na konečné výsledky značný vliv. Jednotlivé oddíly totiž vždy spojuje určitý gramatický rys čínštiny, který je právě těmito větami demonstrován. Z popisu výběrového souboru C tedy vyplývá, že se nejedná o tematicky ucelený text. Tento fakt také nesporně mohl ovlivnit výsledky.

Aplikace MALu na nejnižší jazykovou úroveň tvořenou přízvukovým taktem, respektive segmentem (konstrukt, jehož délka je měřena v počtu hlásek) a slabikou (konstituent, jehož délka je měřena v průměrném počtu slabik) přinesla celkově velmi zajímavé výsledky. U dvou výběrových souborů poukázala na určitou lineární tendenci. U jednoho výběrového souboru se pak projevila klesající tendence. Podle mého názoru je ze slabičnosti čínštiny pravděpodobné, že se na zmiňované nejnižší jazykové úrovni bude lineární tendence projevovat velmi často. Dalším výzkumem je však nutné zjistit, proč se u jednoho výběrového souboru prokázala shoda, respektive je nezbytné zjistit, zda se nejednalo jen o náhodu a zda tato klesající tendence nebyla dána právě tím, že soubor, na kterém se shoda projevila, je umělý.

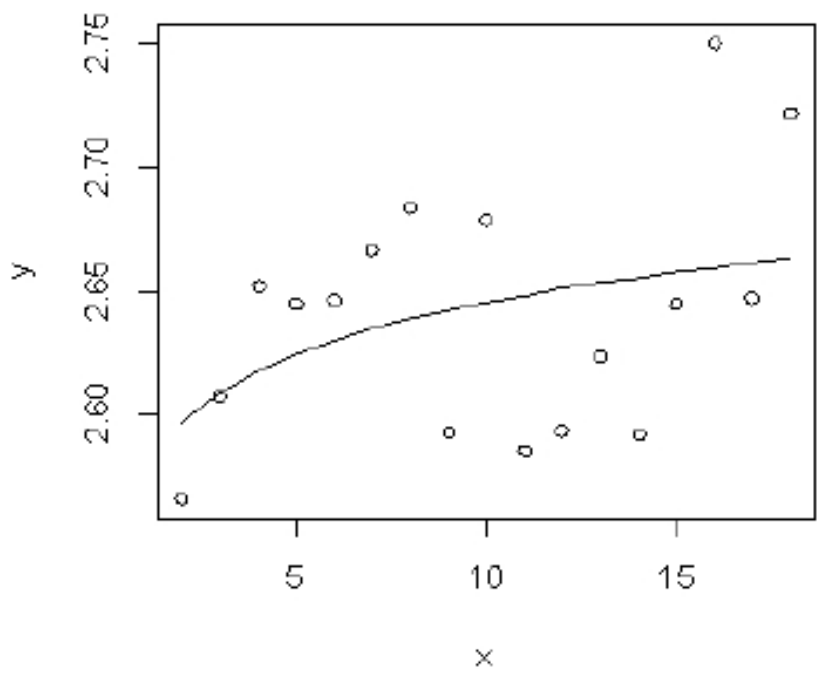
Když postavíme výběrové soubory A a B do kontrastu s výběrovým souborem C, musíme se zamyslet také nad způsobem segmentace zmiňovaných souborů.

Vzhledem k tomu, že na nejnižší jazykové úrovni stojí v pozici konstruktů jazykové jednotky přízvukový takt a segment, tedy jazykové jednotky, které jsou subjektivní, musíme vzít v úvahu také jejich různé členění. Vzhledem k tomu, že se jedná o jazykové jednotky striktně fonetické, nastává občas problém s určením jejich hranic, respektive s tím, že každý by jejich hranice na základě empirické zkušenosti určit jinak.

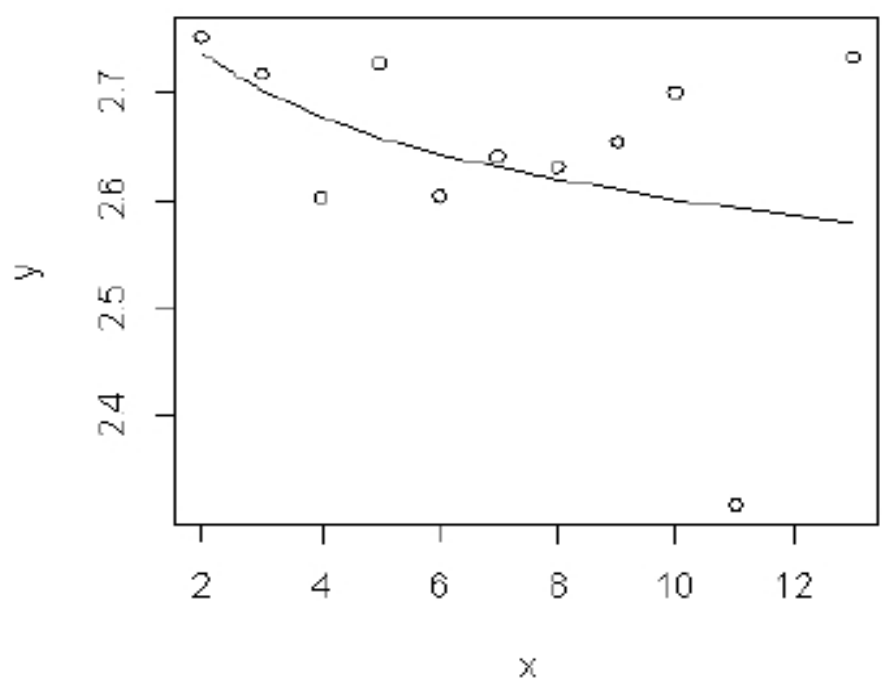
8.5 Jazyková úroveň U3/U3Š: přízvukový takt/segment (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek) – odebrání některých pozorování

Dále uvádím grafy po odebrání některých bodů. Ve statistice je běžně zažitou praxí, že se některá pozorování, zvláště ty body, jež se svými hodnotami od zbylých bodů odchyľují, vynechávají. Dovolila jsem si tento krok učinit zejména proto, že body, které jsem odebrala, byly právě těmi body, které reprezentují hezitační zvuky, vulgarismy a víceslabičné segmenty. Hezitační zvuky, vulgarismy a víceslabičné segmenty nejsou běžnou součástí promluvy. Podle mého názoru jsou to jevy specifické, a tudíž není nutné, aby byly do výpočtů zahrnuty. Dále mě k tomuto kroku vedl fakt, že frekvence výskytu zmiňovaných bodů je velice nízká.

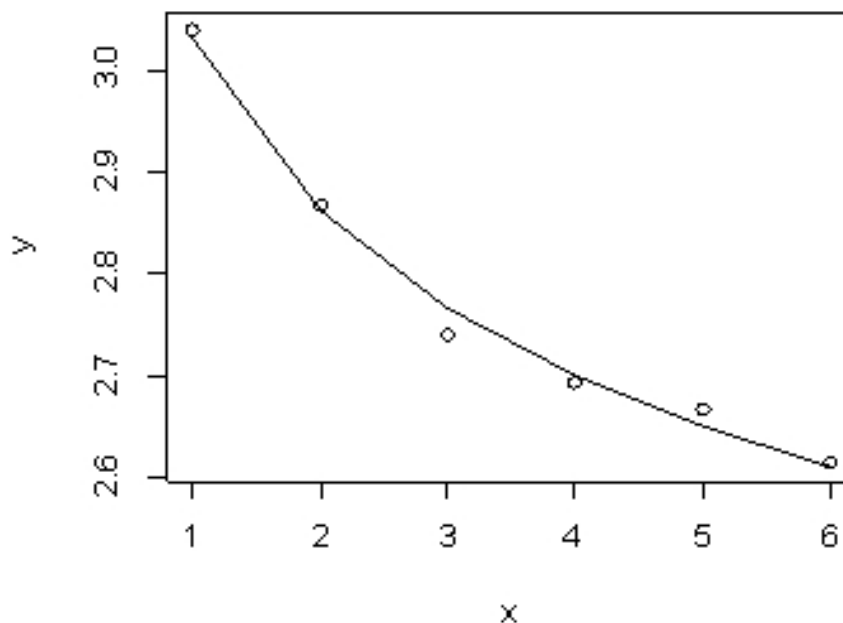
U vzorku A a B jsem vynechala vždy body $x = 1$. U vzorku C se potom jednalo o bod $x = 7$.



Graf 13: U3 přízvukový takt vs. slabika – soubor A: vizualizace souboru dat v Tabulce 10 po odebrání bodu $x = 1$, pomocí programu R



Graf 14: U3 přízvukový takt vs. slabika – soubor B: vizualizace souboru dat v Tabulce 11 po odebrání bodu $x = 1$, pomocí programu R



Graf 15: U3Š segment vs. slabika - soubor C: vizualizace souboru dat v Tabulce 13 po odebrání bodu $x=7$, pomocí programu R

Po odebrání některých pozorování nenastala u výběrových souborů A, B ani C žádná dramatická změna ve vzhledu grafů. U souboru C se však rapidně zvýšilo procento koeficientu determinace. Procentuální shoda s MALEM se díky tomuto kroku zvýšila z 55,6% na 99,1%. Znamená to tedy, že pokud vyčleníme z promluvy sedmi-slabičné segmenty, které se v celém souboru objevily pouze dvakrát, pak křivka vykazuje téměř identickou shodu s MALEM. Zde znovu připomínám, že se jedná o výběrový soubor C, tedy soubor, který prezentuje promluvu, která není spontánní, ale naopak je předem připravená. Je zajímavé, že se MAL projevil právě na výběrovém souboru takové charakteru. V budoucím výzkumu proto plánuji provést aplikaci MALu na soubory nejrůznějších typů. Je třeba provádět aplikaci MALu na spontánní promluvy, ale zároveň také znovu provést aplikaci MALu na soubor, který prezentuje předem připravenou promluvu. A to právě z toho důvodu, aby se prokázalo, zda se jednalo o náhodu, nebo zda právě soubory prezentující předem připravenou promluvu budou i nadále prokazovat shodu s MALEM. Během našeho výzkumu byl MAL aplikován pouze na tři výběrové soubory. Z takového množství tedy není možné činit závěry.

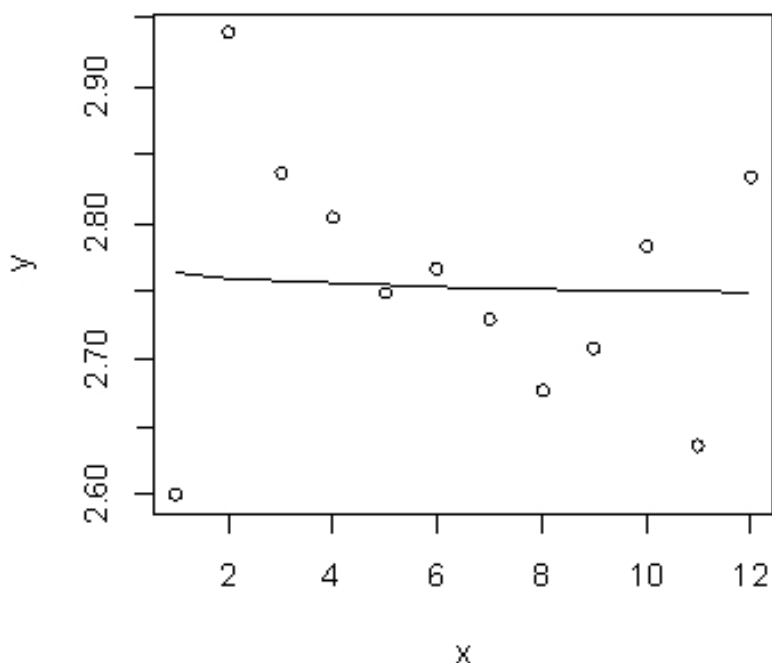
8.6 Jazyková úroveň U2Š-U1Š: kólon (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek)

Pro doplnění uvádím Tabulky 15 a 16 a Graf 16, ve kterých prezentuji výsledky, které vzešly z výpočtů prováděných opět na výběrovém souboru C. Jedná se však o jazykovou úroveň tvořenou kólonem (konstrukt, jehož délka je měřena v počtu slabik) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek). Z popisu této jazykové úrovně vyplývá, že byla vynechána jedna jazyková jednotka, a to segment (připomeňme si, že Švarný člení text na tyto jazykové hladiny: hláska, slabika, segment, kólon, věta). Záměrně jsem vynechala jazykovou jednotku: segment. A to z toho důvodu, abych zjistila, jak se tento krok projeví na konečných výsledcích.

Tabulka 15: U2Š-U1Š, výběrový soubor C: konstrukty \mathcal{X}_2 – délka kolónů (měřena v počtu slabik), \mathcal{Y}_2 – frekvence konstruktů, konstituenty \mathcal{Y}_2 – průměrná délka slabiky (měřena v počtu hlásek)

délka kolónů ve slabikách \mathcal{X}_2	frekvence \mathcal{X}_2	průměrná délka slabik v hláskách \mathcal{Y}_2
1	5	2,6
2	33	2,939393939
3	84	2,837301587
4	135	2,803703704
5	142	2,749295775
6	88	2,767045455
7	64	2,729910714
8	51	2,676470588
9	29	2,708812261
10	6	2,783333333
11	2	2,636363636
12	1	2,833333333

V Tabulce 15 vidíme, že kólon je v tomto vzorku tvořen až 12 slabikami. Počet konstruktů na této úrovni je 640. Průměrná délka slabiky měřená v počtu hlásek osciluje v rozmezí 2,6 - 2,93.



Graf 16: U2Š-U1Š kólon vs. slabika - soubor C: vizualizace soubor dat v Tabulce 15 pomocí programu R

Z Grafu 16 je očividná linearita způsobená pravděpodobně vynecháním jedné z jazykových jednotek.

Tabulka 16: Přehled parametru b MAL a koeficientu determinace na U2Š-U1Š

	Parametr b	Koeficient determinace R^2 (%)
Vzorek C	0,0021	0,2

Experiment s vynechanou jazykovou jednotkou jsem se rozhodla ve své diplomové práci uvést z toho důvodu, že bych na něm chtěla demonstrovat, jak důležité je stanovení jazykových jednotek na počátku výzkumu.

Při analýze výsledků jazykové úrovně tvořené replikou (konstrukt, jehož délka je měřena v počtu výpovědí) a výpovědí (konstituent, jehož délka je měřena v průměrném počtu přízvukových taktů) výběrových souborů A a B jsem naznačila, že jednou z příčin, proč se na zmiňované úrovni MAL nepotvrdil, může být nedostatečný počet jazykových jednotek. Právě na grafu 16 můžeme vidět, že vynecháním jazykové jednotky segment, se projevila absolutně lineární tendence.

V budoucích výzkumech je proto nezbytné v počátečním stadiu výzkumu věnovat dostatek prostoru definování jazykových jednotek a následně také jazykových úrovní. Vynechání některé z jazykových jednotek může mít nesmírný vliv na konečné výsledky.

9. ZÁVĚR

Cílem mé diplomové práce bylo zjistit, zda se MAL odráží na jednotlivých jazykových úrovních současné mluvené čínštiny se zvláštním zřetelem na nejnižší jazykovou úroveň tvořenou přízvukovým taktem, respektive segmentem (konstrukt, jehož délka je měřena v počtu slabik) a slabikou (konstituent, jehož délka je měřena v průměrném počtu hlásek). Původní hypotéza byla taková, že MAL se odrazí na všech jazykových úrovních.

V první fázi byl výzkum prováděn testováním na dvou výběrových souborech. Jednalo se o nahrávky pořízené pro čínský televizní hudební pořad. Mluvčími výběrových souborů byli popová zpěvačka a rapový zpěvák. U těchto dvou výběrových souborů se jednalo o spontánní promluvy. Výběrový soubor, kde mluvčím byla popová zpěvačka, byl nazván výběrový soubor A, a soubor, kde mluvčím byl rapový zpěvák, byl nazván výběrový soubor B.

MAL byl v první fázi aplikován na následující jazykové úrovně: replika (konstrukt, jehož délka je měřena ve výpovědích) a výpověď (konstituent, jehož délka je měřena v průměrném počtu přízvukových taktů), výpověď (konstrukt, jehož délka je měřena v přízvukových taktech) a přízvukový takt (konstituent, jehož délka je měřena v průměrném počtu slabik), přízvukový takt (konstrukt, jehož délka je měřena v slabikách) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek).

Výzkum přinesl tyto výsledky. U nejvyšší jazykové úrovně (replika vs. výpověď) je patrná pouze mírně klesající tendence. MAL se tedy na nejvyšší jazykové úrovni neprojevil. Takový výsledek může mít hned několik důvodů.

Jedním z nich je především nedostatečně vysoký počet konstruktů na této jazykové úrovni. Ani u jednoho vzorku nepřesáhnul počet konstruktů na jazykové úrovni replika vs. výpověď hodnotu 20.

Jako druhou příčinu pak uvádím nedostatečný počet zkoumaných jazykových jednotek, respektive vynechání některé z nich. Považuji proto za nutné se při dalších

experimentech zamyslet nad navýšením jazykových jednotek a paralelně s tímto krokem tak navýšit i počet jazykových úrovní.

Poslední příčinou by pak mohl být umělý zásah do promluvy. Hudební pořad, který jsme zvolily jako výběrové soubory A a B, je nastříhán pro účely jeho tvůrců. Tyto zásahy rovněž mohly ovlivnit výsledky výzkumu.

Vzhledem k tomu, že ani u výběrového souboru A, ani u výběrového souboru B se MAL neprojevil, a vzhledem k tomu, že jednotlivé body pozorování byly v grafu chaoticky rozmístěny, domnívám se, že nejvyšší jazyková úroveň, tj. replika vs. výpověď, si zaslouží samostatný výzkum. V budoucím výzkumu je nutné nejen navýšit počet zkoumaných jazykových jednotek, ale také zvolit odlišné výběrové soubory prezentující promluvy různých mluvčích.

U prostřední jazykové úrovně (výpověď vs. přízvukový takt) se MAL bezpochyby projevil na vzorku A a klesající tendence je patrná také u vzorku B verze 1, i když velmi mírná. U vzorku B verze 2 pak není patrná vůbec žádná klesající tendence. Domníváme se, že hlavní důvod, proč se shoda s MALEM u vzorku A projevila a vzorku B naopak ne, je ten, že mluvčím vzorku B je rapový zpěvák, jehož řeč může být výrazně ovlivněna hudebním stylem, kterým se zabývá.

Tak jako u předcházející nejvyšší jazykové úrovně představoval možnou příčinu neshody s MALEM nedostatečný počet zkoumaných jazykových jednotek, i u úrovně výpověď vs. přízvukový takt může být nedostatečné množství jazykových jednotek jednou z příčin nízké shody v případě výběrového souboru B.

U nejnižší jazykové úrovně přízvukový takt vs. slabika výběrových souborů A a B se MAL nepotvrdil. Na grafech se však projevila oscilace v rámci pásu s výjimkami bodů $x = 1$. Tuto odchylku si v případě vzorku A vysvětlujeme především hezitačními zvuky a gramatickými částicemi, které jsou tvořeny pouze jednou hláskou. V případě vzorku B se pak jedná o anglická slova, kterých mluvčí ve své promluvě často používá.

Právě při analýze výsledků první fáze výzkumu mě napadlo, že na nejnižší úrovni se tato oscilace v rámci pásu projeví vždy, ať se jedná o jakýkoli výběrový soubor současné mluvené čínštiny. Logickým argumentem pro toto tvrzení by pak byla slabičnost čínštiny.

Do druhé fáze výzkumu jsem tedy vstupovala s předpokladem, že na nejnižší úrovni se MAL neprojeví.

Pro druhou fázi výzkumu byl vybrán třetí výběrový soubor, výběrový soubor C. Výběrový soubor C se od výběrových souborů A a B odlišuje tím, že se nejedná o spontánní promluvu. Jako výběrový soubor byla zvolena nahrávka k učebnici GHČP, respektive některé její části. Problematika tohoto souboru spočívá také v tom, že v promluvě se nejedná o kohezní obsah, ale o jednotlivé tematicky oddělené věty. Na druhou stranu byl tento výběrový soubor vybrán záměrně. A to především z toho důvodu, že při segmentaci textu výběrového souboru C bylo použito odlišného členění, než jsme u segmentace výběrových souborů A a B použily my.

Vzhledem k tomu, že jsme se v našich diplomových pracích rozhodly zabývat dvěma nejnižšími jazykovými úrovněmi, tj. výpověď vs. přízvukový takt a přízvukový takt/segment vs. slabika, byl MAL na soubor C aplikován pouze na zmiňované dvě nejnižší úrovně, kdy jazykové jednotky v těchto úrovních obsažené se odlišovaly od jazykových jednotek obsažených ve výběrových souborech A a B.

V druhé fázi výzkumu pak byl MAL zkoumán na následujících úrovních: kólon (konstrukt, jehož délka je měřena v počtu segmentů) a segment (konstituent, jehož délka je měřena v průměrném počtu hlásek) a na úrovni segment (konstrukt, jehož délka je měřena v počtu slabik) a slabika (konstituent, jehož délka je měřena v průměrném počtu hlásek).

Na úrovni segment vs. slabika výběrového souboru C, tedy na úrovni, která je paralelní s úrovní přízvukový takt vs. slabika výběrových souborů A a B, se MAL překvapivě projevil.

Takový výsledek tedy vyvrátil mou hypotézu, že se MAL na nejnižší úrovni neprojeví, ale že se kvůli slabičné povaze čínštiny na nejnižší jazykové úrovni projeví oscilace v rámci pásu.

Je však velmi pravděpodobné, že výsledek mohl být ovlivněn především tím, že výběrový soubor C prezentuje nespontánní, předem připravenou promluvu. Znovu bych chtěla připomenout, že výběrový soubor C představuje nahrávka pořízená k učebnici GHČP. Vzhledem k tomu, že se jedná o učební materiál, který je určen především začínajícím studentům čínštiny, je tomu přizpůsobeno i tempo a členění

řeči. Mluvčí tohoto souboru se snaží všechny věty vyslovit tak, jak by měly správně znít.

Tematická neucelenost výběrového souboru C také může být jednou z příčin, proč se na nejnižší úrovni u výběrového souboru C MAL projevil. Jak jsem již zmiňovala výše, jsme si vědomy, že vzorek C disponuje velkým množstvím nedokonalostí. Na druhou stranu je však třeba se zamyslet nad kontrasty výsledků u výběrových souborů A a B a výsledkem u výběrového souboru C. Je také nutné zohlednit slabičnost čínštiny a to, že tento její rys by mohl mít na projev MALu u hovorové čínštiny značný vliv. I když je výběrový soubor C v mnoha ohledech souborem nevhodným, jsem toho názoru, že byl přesto vhodně zvoleným výběrovým souborem. A to právě kvůli rozdílným výsledkům mezi výběrovým souborem C a výběrovými soubory A a B.

V budoucím výzkumu plánuji zjistit, proč se na nejnižší jazykové úrovni výběrových souborů A a B MAL neprojevil a u výběrového souboru C naopak projevil. Budoucím výzkumem je třeba také zjistit, zda mělo na platnost MALu u výběrového souboru C vliv to, že tento soubor je umělý, anebo zda to byla pouze náhoda a u jiných výběrových souborů podobného charakteru už se znovu MAL neprojeví.

Je také třeba znovu zmínit způsob segmentace výběrových souborů. Vzhledem k tomu, že jazyková jednotka přízvukový takt je jednotka velmi subjektivní, mohly být výsledky rozdílné v případě, že by segmentace textu byla provedena někým jiným.

Vzhledem k tomu, že MAL dosud nebyl na čínštinu aplikován, jsme teprve v rané fázi výzkumu. Během našeho bádání byl MAL aplikován pouze na tři výběrové soubory současné mluvené čínštiny. Z tak nízkého počtu experimentů samozřejmě nelze vyvozovat žádné konečné výsledky. V dalších výzkumech plánuji provést zkoumání na mnoha jiných výběrových souborech různého charakteru současné mluvené čínštiny, na promluvách produkovanými různými mluvčími.

Aplikací MALu na další výběrové soubory plánuji zjistit, proč se MAL na výběrových souborech odrážejících spontánní řeč na nejnižší úrovni neprojevil, a naopak na souboru odrážejícím předem připravenou řeč se MAL projevil.

10. RESUMÉ

The aim of this paper is to verify the validity of the Menzerath-Altmann Law (MAL) applied to contemporary spoken Chinese with a special reference to the language unit consisting of the stress unit (a construct, which is measured in syllables) and the syllable (a constituent, which is measured in the average numbers of phonemes). The original hypothesis stated that the tendency of the MAL would apply to the mentioned language unit. However, the results of the first part of the research did not confirm the hypothesis. It demonstrated that the individual observations oscillate within the linear area. The second hypothesis consequently stated that all the following research will prove the same phenomena. Within the second phase of the research, the MAL was tested with a different sample. The results from this particular phase of the research surprisingly demonstrated the tendency with the MAL. Over the course of the following research it will be therefore necessary to discover what the reason for the differences between the results is.

11. KNIŽNÍ A INTERNETOVÉ ZDROJE

ALTMANN, Gabriel. **Prolegomena to Menzerath's law.** *Glottometrika*, 1980, 2, p. 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*, [databáze online]. 2012, roč. 19, č. 1, s. 1-31. [cit. 2012-03-16]. 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*, [databáze online]. 2010, roč.17, č. 2, s.101-122. [cit. 2013-03-11]. Dostupné z: <<http://www.tandfonline.com/doi/abs/10.1080/09296171003643189>>.

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

ČERNÝ, Jiří. *Dějiny lingvistiky.* Olomouc: Votobia, 1996, 517 s. ISBN: 80-85885-96-4.

ČERNÝ, Jiří. *Úvod do studia jazyka.* Olomouc: Rubico, 2008, 248 s. ISBN: 978-80-7346-093-8.

DeFRANCIS, John. *The Chinese Language Fact and Fantasy.* Taipei: The Crane Publishing co., LTD, 1990.

DOKOUPIL, Martin. *Jazykověda.* Vrbátky: M. Dokoupil, 2012. 136 s. ISBN: 978-80-260-2129-2.

Doplnit diplomku od Janky

HÁLA, Bohuslav. *Fonetika v teorii a v praxi.* Praha: Státní pedagogické nakladatelství, 1975, 482 s.

- HÁLA, Bohuslav. *Slabika: její podstata a vývoj*. Praha: Nakladatelství československé akademie věd, 1956, 107 s.
- HAUSER, Přemysl. *Nauka o slovní zásobě*. Praha: Státní pedagogické nakladatelství, 1980, 192 s, s. 9.
- HŘEBÍČEK, Luděk. *Lectures on text theory*. Prague: Oriental Institute, 1997.
- HŘEBÍČEK, Luděk. **Struktura textu a sémantika**. *AntropoWebzin* [online]. 2006, 1,2 [cit. 2013-03-01]. ISSN 1801-8807. Dostupné z: <<http://antropologie.zcu.cz/>>
- HŘEBÍČEK, Luděk. *Text in semantics: The principle of compositenes*. Prague: Academy of Sciences of Czech Republic, Oriental Institute, 2007.
- HŘEBÍČEK, Luděk. *Vyprávění o lingvistických experimentech s textem*. Vyd. 1. Praha: Academia, 2002, 195 s. ISBN 80-200-0973-6.
- KANE, Daniel. *Knížka o čínštině*. Mirošovice: DesertRose, 2009. ISBN: 978-80-903296-1-4.
- KANE, Daniel. *The Chinese language: its history and current usage*. 3. repr. Tokyo: Tuttle Pub, 2006. ISBN 08-048-3853-4.
- KARLÍK, Petr. *Encyklopedický slovník češtiny*. Praha: NLN, 2002, 604 s. ISBN 80-710-6484-X.
- KRČMOVÁ, Marie: *Fonetika* [online] Brno: Filosofická fakulta MU, 2007 [17. 01. 2013] Dostupné z: is.muni.cz/el/portal/estud/ff/js07/fonetika/materialy/ch04.html.
- LIN, Yen-Hwei. *The Sounds of Chinese*. Cambridge: Cambridge University Press, 2007, 312 p. ISBN: 978-0-521-60398-0.
- Liu Haitao, Huang Wei. **Jiliang yuyanxue de xianzhuang, lilun yu fangfa**. *Zhejiang daxue xuebao* (renwen shehui kexue ban) 刘海涛, 黄伟. 计量语言学的现状、理论与方法. *浙江大学学报 (人文社会科学版)* 2012, roč. 42, č. 2. ISSN 1008-942X.
- MÜLLEROVÁ, Olga. *Mluvený text a jeho syntaktická výstavba*. Praha: Academia, 1994, 145 s. 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, 386 p. ISBN 31-101-5109-X.

PING, Chen (1999): *Modern Chinese*. Cambridge: Cambridge University Press.

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 tisku).

ŠČIGULINSKÁ, Jana. *Aplikácia Menzerath-Altmannovho zákona na súčasnú hovorenú čínštinu na úrovni výpovede*. Olomouc, 2013. Magisterská práce (Mgr.).
UNIVERZITA PALACKÉHO V OLOMOUCI. Filozofická fakulta

ŠVARNÝ, Oldřich a David UHER. *Hovorová čínština: úvod do studia hovorové čínštiny*. Olomouc: Univerzita Palackého, 2001, 190 s. ISBN 80-244-0298-X.

ŠVARNÝ, Oldřich a kol. (1998): *Hovorová čínština v příkladech 3*. Olomouc: Vydavatelství Univerzity Palackého.

ŠVARNÝ, Oldřich a kol.: *Gramatika hovorovej čínštiny v príkladoch*. Bratislava: Vydavateľstvo Univerzity Komenského, 1993, 261 s. ISBN: 80-223-0534-0.

ŤEŠITELOVÁ, Marie. *Kvantitativní lingvistika*. 1. vyd. Praha: Univerzita Karlova, 1987.

ŤEŠITELOVÁ, Marie. *Využití statistických metod v gramatice*. 1. vyd. Praha: Academia, 1980.

VOLÍN, Jan. *Statistické metody ve fonetickém výzkumu*. Praha: Nakladatelství Epoque, 2007, 343 s. ISBN: 9788087027547.

WIMMER, Gejza. *Úvod do analýzy textov*. Bratislava: Veda, 2003, 344 s. ISBN 80-224-0756-9.

XING, Janet Zhiqun. *Studies of Chinese linguistics: functional approaches*. Hong Kong: Hong Kong University Press, 2009, 235 p. ISBN: 96-220-9964-5.