

UNIVERZITA PALACKÉHO V OLOMOUCI

Filozofická fakulta

Katedra asijských studií

MAGISTERSKÁ DIPLOMOVÁ PRÁCE

Frekvenční korpusová analýza slovní zásoby HSK

Corpus-Based Frequency Analysis of Selected HSK Vocabulary

OLOMOUC 2016, Gabriela Tan

Vedoucí diplomové práce: Doc. Mgr. David Uher, PhD.

Kopie zadání DP ze STAGU

Prohlašuji, že jsem diplomovou práci na téma: „Frekvenční korpusová analýza vybrané slovní zásoby HSK“ vypracovala samostatně pod dohledem vedoucího diplomové práce a uvedla jsem všechny použité podklady a literaturu.

Ve Valašském Meziříčí dne 25. 4. 2016

Podpis.....

Anotace

Příjmení a jméno autora:	Tan Gabriela
Název katedry a fakulty:	Katedra asijských studií, Filozofická fakulta
Název diplomové práce:	Frekvenční korpusová analýza vybrané slovní zásoby HSK
Vedoucí diplomové práce:	Doc. Mgr. David Uher, PhD.
Počet stran:	81
Počet znaků:	135 418 (bez mezer), 116 087 (včetně mezer)
Počet příloh:	1
Počet sekundárních zdrojů:	78
Klíčová slova:	HSK, Hanyu Shuiping Kaoshi, frekvenční analýza, slovní zásoba, korpus, frekvence, Hanban, studium čínštiny

Tato diplomová práce se věnuje analýze vývoje vybraných úrovní doporučené slovní zásoby zkoušky z čínštiny jako cizího jazyka (HSK). Tematicky se zaměřuje na porovnání seznamu slovíček pro začátečníky starší formy zkoušky s novým soupisem lexémů HSK4 a se seznamem téže úrovně přepracovaným v roce 2012. V teoretické části je pojednáváno o vývoji zkoušky, kontroverzi změn kvalifikace úrovní testů a také o problematice tvorby slovníčku pro studenty cizího jazyka. V praktické části jsou na základě korpusů LCMC, TORCH2009, SUBTLEX-CH a BCC(微信) analyzovány frekvence výskytu lexikálních jednotek v seznamech a jejich distribuce v textech různých žánrů. Frekvenční analýzou soupisů bylo zjištěno, že starší podoba slovníčku HSK lépe než ostatní soupisy odráží užití jazyka v psané i hovorové standardní čínštině.

Ráda bych poděkovala vedoucímu mé práce panu docentu Mgr. Davidu Uhrovi, Phd. za vstřícnost, pomoc a podporu při zpracování mé diplomové práce i během celého mého studia.

Obsah

1. ÚVOD	11
2. METODOLOGIE	13
3. ZHODNOCENÍ LITERATURY	15
4. TEORETICKÁ ČÁST	16
4.1. Vznik zkoušky HSK	16
4.1.1 Vývoj původní zkoušky HSK.....	16
4.1.2 Vývoj zkoušky 改进版 HSK.....	18
4.1.3 Vývoj nové zkoušky HSK.....	20
4.2. HSK v České republice.....	22
4.3. Formát zkoušky HSK	23
4.3.1 Formát zkoušky HSK (基础)	23
4.3.2 Formát zkoušky 新 HSK4.....	25
4.4. Hodnocení a kritika zkoušky HSK.....	26
4.4.1 Hodnocení a kritika staré formy zkoušky HSK.....	26
4.4.2 Hodnocení a kritika zkoušky 改进版 HSK.....	27
4.4.3 Hodnocení a kritika nové zkoušky HSK.....	28
4.5. Specifika čínštiny a slovní zásoba.....	29
4.6. Výběr lexémů seznamu slovní zásoby.....	32
4.6.1 Role frekvence při sestavování slovníčku.....	32
4.6.2 Základní slovní zásoba.....	33
4.7. Obsah seznamu slovní zásoby	35
4.7.1 Sémantická průhlednost lexémů.....	37
4.7.2 Využití souvislostí mezi lexikálními jednotkami.....	38
4.7.3 Lexikální přístup.....	38
4.7.4 Role morfémů při akvizici čínštiny.....	39
5. PRAKTICKÁ ČÁST	40
5.1. Obecná charakteristika seznamu slovní zásoby HSK 甲级.....	40
5.1.1 Kvantitativní posouzení seznamu HSK 甲级.....	41

5.1.2	Kvalitativní posouzení seznamu HSK 甲级.....	48
5.1.3	Změny seznamu HSK 甲级 po reformě.....	50
5.2.	Obecná charakteristika seznamu slovní zásoby 新 HSK4 (2010)	52
5.2.1	Kvantitativní posouzení seznamu 新 HSK4 (2010)	54
5.2.2	Kvalitativní posouzení seznamu 新 HSK4 (2010)	57
5.2.3	Změny v seznamu 新 HSK4 (2010)	59
5.3.	Obecná charakteristika seznamu slovní zásoby 新 HSK4 (2012)	61
5.3.1	Kvantitativní posouzení seznamu 新 HSK4 (2012)	63
5.3.2	Kvalitativní posouzení seznamu 新 HSK4 (2012)	64
5.3.3	Změny v seznamu 新 HSK4 (2012)	65
5.4.	Shrnutí.....	67
6.	ZÁVĚR.....	70
	Summary.....	72
	Seznam literatury a dalších zdrojů.....	73
	Seznam příloh.....	81

Seznam grafů a tabulek

Tabulka č. 1: Přehled úrovní staré verze zkoušky HSK

Tabulka č. 2: Přehled úrovní nové verze zkoušky HSK

Tabulka č. 3: Porovnání úrovní nové zkoušky HSK a TOCFL podle SERR

Tabulka č. 4: Slovíčka ze soupisu HSK 甲级 v mikroblogových příspěvcích

Tabulka č. 5: Slovíčka ze soupisu HSK4 (verze 2010) v mikroblogových příspěvcích

Graf č. 1: Shoda soupisu HSK 甲级 s 1 000 nejfrekventovanějšími lexémy v korpusech

Graf č. 2: Výskyt lexémů soupisu HSK 甲级 ve filmové databázi a textových žánrech

Graf č. 3: Průměrná frekvence seznamu HSK 甲级 v textových žánrech

Graf č. 4: Průměrná frekvence všech lexikálních jednotek v seznamech

Graf č. 5: Složení seznamů slovní zásoby podle délky slov

Graf č. 6: Frekvenční rozložení lexémů v seznamech slovní zásoby

Graf č. 7: Průměrná frekvence lexémů odstraněných ze seznamu HSK 甲级

Graf č. 8: Shoda soupisu HSK4 verze 2010 s 1 000 nejfrekventovanějšími lexémy v korpusech

Graf č. 9: Výskyt lexémů soupisu HSK4 (2010) ve filmové databázi a textových žánrech

Graf č. 10: Průměrná frekvence seznamu HSK4 (2010) v textových žánrech

Graf č. 11 Průměrná frekvence lexémů přidávaných do seznamu HSK4 (2010)

Graf č. 12 Frekvenční rozložení lexémů odebraných a přidávaných do seznamu HSK4 (2010)

Graf č. 13: Průměrná frekvence lexémů po změnách v seznamu HSK4 (2010)

Graf č. 14: Shoda soupisu HSK4 verze 2012 s 1 000 nejfrekventovanějšími lexémy v korpusech

Graf č. 15: Výskyt lexémů soupisu HSK4 (2010) ve filmové databázi a textových žánrech

Graf č. 16: Průměrná frekvence seznamu HSK4 (2012) v textových žánrech

Graf č. 17: Průměrná frekvence lexémů po změnách v seznamu HSK4 (2012)

Seznam zkratek

ČLR Čínská lidová republika

HSK *Hànyǔ Shuǐpíng Kǎoshì* (Zkouška z čínštiny jako cizího jazyka)

HSKK *Hànyǔ Shuǐpíng Kǒuyǔ Kǎoshì* (Ústní zkouška z čínštiny jako cizího jazyka)

SERR Společný evropský referenční rámec pro hodnocení jazyků

TOCFL *Test Of Chinese as a Foreign Language* (Zkouška z čínštiny jako cizího jazyka)

TOEFL *Test of English as a Foreign Language* (Zkouška z angličtiny jako cizího jazyka)

Ediční poznámka

V této práci jsou čínské znaky psány ve své zjednodušené podobě. V případě uvedení výslovnosti znaků je přepis znaků vždy v transkripci pinyin. Přepis čínských slov je uváděn v transkripci pinyin, pokud se nejedná o místní názvy, které mají v českém jazyce zavedený častěji používaný ekvivalent.

1 ÚVOD

Objektem této diplomové práce je vývoj vybraných úrovní slovní zásoby zkoušky HSK (Zkouška z čínštiny jako cizího jazyka) doporučené organizací Hanban (Kancelář pro výuku čínštiny jako cizího jazyka při ministerstvu školství ČLR). HSK je určena dospělým cizincům, kteří chtějí mít certifikát o úrovni své jazykové kompetence v čínském jazyce. Zkouška HSK prošla od svých začátků v 80. letech 20. stol. dlouhým vývojem. Roku 1997 byla založena úroveň zkoušky HSK pro začátečníky (Ren 2001, s. 63). Nyní byla zkouška rozdělena na tři úrovně, od začátečnické přes základní/středně pokročilou po pokročilou, ke kterým byla přiřazena doporučená slovní zásoba k absolvování zkoušky. V roce 2009 byly forma, obsah i úrovně testu přepracovány. Doporučená slovní zásoba byla obsahově výrazně zkrácena a přepracována tak, aby lépe odpovídala potřebám studentů (Hankao Guoji 2009a). Po dvou letech byla v roce 2012 doporučená slovní zásoba dále upravena a od té doby až do současnosti zůstala beze změny. HSK se během svého vývoje stala nejrozšířenější mezinárodní standardizovanou zkouškou z čínského jazyka na světě (Meyer 2014, s. 15).

Tematicky se práce zaměřuje na porovnání jednotlivých seznamů slovní zásoby pro začátečníky v její starší formě (*HSK jiǎjí* HSK 甲级) v rozsahu 1 033 lexikálních jednotek s novou slovní zásobou HSK4 (*HSK sì jí* HSK 四级) a se seznamem stejné úrovně přepracovaným v roce 2012, které mají oba rozsah 1 200 lexikálních jednotek. Tyto úrovně seznamů slovní zásoby byly zvoleny pro jejich srovnatelný rozsah a také s ohledem k prohlášení organizace Fachverband Chinesisch, podle níž HSK4 spadá do úrovně A2 – znalostí začátečníků podle společného evropského referenčního rámce hodnocení znalosti cizích jazyků (Fachverband Chinesisch 2010). Podle výzkumu, který provedl Florian Meyer (2009, s. 40), odpovídá začátečnická úroveň starého testu HSK také úrovní A2 SERR. Srovnatelnost výsledků obou verzí testu potvrzuje rovněž prohlášení organizace Hanban (Hankao Guoji 2009b).

Důvodem výběru tématu je potřeba bližšího poznání praktické využitelnosti vybraných úrovní slovní zásoby zkoušky HSK pro účely studia čínského jazyka. Seznam slovní zásoby HSK je výsledkem mnohaleté práce odborníků, je veřejně přístupný a není přímo závislý na učebnicích. Díky tomu může být používán nejen jako pomůcka ke složení zkoušky, ale také k systematickému osvojení si čínského jazyka.

Předmětem práce je zkoumání změn v obsahu slovní zásoby HSK zejména s ohledem na frekvenci výskytu lexikálních jednotek v psaném i hovorovém jazyce a na změny ve výskytu jednotlivých lexikálních jednotek v různých textových žánrech.

Zdrojem informací jsou hlavně odborné články a publikace zabývající se korpusovou lingvistikou, lexikografií a lexikologií. Primárními zdroji jsou seznamy slovní zásoby publikované organizací Hanban a korpusy čínského jazyka. Při analýze frekvence a distribuce v textu se budu opírat o korpus LCMC (*The Lancaster Corpus of Mandarin Chinese*). Jde o vyvážený korpus obsahující jeden milion slov vyvinutý na univerzitě v Lancasteru. Využiji také korpus TORCH2009 (*Texts of Recent Chinese 2009*), což je také vyvážený korpus, který pojímá přes milion slov, ovšem obsahuje texty publikované o 18 let později než LCMC. Pro výzkum užití lexikálních jednotek v hovorové čínštině využiji SUBTLEX-CH. To je korpus čínských televizních a filmových titulků vyvinutý na univerzitě v Gentu, obsahující 33,5 milionu slov. Dalším zdrojem frekvencí bude korpus BCC(微信) (*wēixìn*), což je čínský korpus mikrobloginových příspěvků.

Metody výzkumu vycházejí z výše uvedených zdrojů. Nejdříve zmapuji změny ve slovní zásobě jednotlivých seznamů. Poté použiji korpusy k prozkoumání frekvencí všech lexikálních jednotek v seznamech slovní zásoby. Zjištěná data poté z kvantitativního a kvalitativního hlediska porovnam.

Podle výchozí hypotézy by současná verze slovní zásoby HSK4 v porovnání s její původní a starší verzí HSK na začátečnické úrovni měla být z hlediska akvizice a orientování se v moderním čínském jazyce pro studenty usilující o dosažení daných úrovní nejvhodnější. Záměrem práce je zjistit, v jakých ohledech se slovní zásoba HSK pro začátečníky změnila a jak osvojení si těchto slovních zásob ovlivní porozumění studenta běžné čtené a hovorové standardní čínštině. Přínosem práce bude zejména poznatek, zda osvojení si slovní zásoby HSK daných úrovní umožní studentům skutečnou schopnost porozumět modernímu čínskému jazyku podle kritérií definovaných tvůrci zkoušky.

2 METODOLOGIE

Výchozími podklady pro analýzu slovní zásoby jsou seznamy slovíček publikované v sylabech HSK. Pro starší verzi HSK 甲级 čerpám ze seznamu v příručce HSK *Zhōngguó Hànyǔ Shuǐpíng Kǎoshì Yìngshì Zhǐnán (jīchǔ)* HSK 中国汉语水平考试应试指南 (基础) [Průvodce testováním čínské zkoušky HSK (začátečnická úroveň)] z roku 2000, pro původní verzi slovní zásoby HSK4 čerpám z příručky *Xīn Hànyǔ Shuǐpíng Kǎoshì Dàgāng HSK Sì Jí* 新汉语水平考试大纲 HSK 四级 (Sylabus čtvrté úrovně nového testu z čínštiny jako cizího jazyka) a jako podklad k aktuální verzi slovní zásoby HSK používám seznam publikovaný na oficiálních stránkách Hanbanu (Hankao Guoji, ©2013–2016a). Tyto seznamy v pořadí abecedně podle výslovnosti v pinyinu seřadím a zmapuji změny v jejich obsahu. Označím lexikální jednotky, které byly při reformách seznamů odebrány a přidány.

Následně všechny výskyty lexikálních jednotek vyhledám v korpusech. Největší korpus, který používám, je BCC(微信). BCC je rozsáhlý korpus zahrnující 15 miliard slov vyvinutý na Pekingské jazykové univerzitě (*Běijīng Yǔyán Dàxué* 北京语言大学). Zahrnuje mnoho různých subkorpusů, které je možné samostatně prohledávat. Mnou používaný subkorpus je souhrnem mikroblových příspěvků o rozsahu 3 miliardy slov. V mikroblovích lze psát příspěvky o maximální délce 140 znaků. Čínské mikroblogy umožňují nejen přímé interakce na sociálních sítích, ale nabízejí také funkce podobné diskusním fórům (Fu a Chau 2013, s. 5). Mikroblovové platformy s sebou přinesly nový druh specifické komunikace, která má blíže k hovorovému jazyku (Tian 2014, s. 1). Zkoumání frekvencí slovíček korpusu internetových příspěvků mi umožní podívat se na užití moderní čínštiny z nového úhlu.

Korpus LCMC je naopak nepřiliš obsáhlý, ale obsahově velmi dobře vyvážený (McEnery a Xiao 2004, s. 1175). Jedná se o korpus psané moderní čínštiny o rozsahu jednoho milionu slov. Obsahuje texty publikované v PRC kolem roku 1991, které jsou rozděleny do patnácti kategorií. Tři kategorie pocházejí z tisku (úvodníky, reportáže, recenze), jedna je z oblasti naučné literatury (akademické texty) a zbylé dvě jsou řazeny podle žánrů pokrývajících důležitá literární témata.

LCMC mi poslouží k porovnání s korpusem TORCH2009 (*Texts Of Recent Chinese*), který je přibližně stejného rozsahu (1 289 373 slov), ovšem s texty publikovanými v PRC v roce 2009. TORCH byl vyvinut roku 2013 na Beijing Foreign Studies University. Stejně jako korpus LCMC je rozdělen do čtyř obecných žánrů (fikce, próza, naučná literatura a tisk), do kterých je zařazeno patnáct subžánrů. Rozsahem a strukturou se podobá korpusu LCMC. V obou korpusech budu sledovat frekvenci lexémů ze seznamů slovní zásoby a jejich distribuci v jednotlivých žánrech.

Posledním korpusem, z něhož chci čerpat frekvence mluvené čínštiny, je SUBTLEX-CH, což je rozsáhlý korpus televizních a filmových titulků. Frekvence slov získané na základě tohoto korpusu podle autorů dobře reprezentují variaci i běžné každodenní užití jazyka (Cai a Brysbaert 2010, s. 2). I zde lze sledovat frekvenci slov či morfémů a jejich distribuci v korpusu. Původním záměrem bylo sledovat frekvenci užití lexémů přímo v korpusu mluvené čínštiny, ty však bohužel v současné době buď nemají dostatečně reprezentativní rozsah standardní moderní čínštiny, nebo nejsou kvůli obsahu vázanému autorskými právy volně k dispozici. Jazyk televize a filmů sice není spontánním přirozeným jazykovým projevem, ale stále se jedná o projev hovorové čínštiny, která se z hlediska stylu, syntaxe a lexika od psané čínštiny může výrazně lišit.

Práce je rozdělena na teoretickou a praktickou část. V praktické části je popsáno provedení výše uvedených kroků a kvantitativní vyhodnocení získaných dat. Následovat bude kvalitativní analýza jednotlivých seznamů z hlediska aktuálnosti a praktičnosti jejich slovní zásoby a komparace, který z nich je pro uživatele seznamů nejvhodnější pro osvojení si současného čínského jazyka. Dále bude vyhodnoceno, zda po osvojení si slovní zásoby mohou studenti dosáhnout stupně kompetence uváděného tvůrci zkoušky na daných úrovních.

Teoretická část bude pojednávat o vývoji a formátu zkoušky HSK, o jejím zhodnocení, avšak také o kritice, které testy byly a dosud jsou vystaveny. Dále bude stručně popsána problematika slovní zásoby při akvizici čínského jazyka s ohledem na komplikovanost a osobitost lexémů v čínštině. V závěru shrnu, co bylo zjištěno, informace z toho vyplývající a stručně zrekapituluji obsah celé práce.

3 ZHODNOCENÍ LITERATURY

K porozumění problematice vyhodnocování jazykových dat pomocí korpusů jsem využila publikací, jež napsal a editoval František Čermák, který se korpusové lingvistice věnuje zejména z pozice lexikografa. Jeho knihy mi poskytly cenné podněty, jakým způsobem lze seznamy slovní zásoby analyzovat.

Co se týče jednotlivých korpusů, informace o nich jsem čerpala primárně od jejich tvůrců. Ve svých odborných článcích vysvětlují metody tvorby korpusů, jejich složení, funkce, i doporučení ohledně jejich používání, takže jsem si mohla vybrat odpovídající korpusy, které vyhovují mým účelům.

Další teoretické informace, které využiji při praktické analýze a komparaci seznamů slovní zásoby i v teoretické části práce, jsem získala z internetových databází, odborných a publicistických článků. V databázi elektronických informačních zdrojů UP jsou přístupné odborné články týkající se lexikografie, lexikologie a akvizice cizího jazyka.

Při zpracování teoretické části jsem čerpala z odborných článků a vědeckých prací týkajících se didaktiky čínského jazyka a zkoušky HSK. V databázích UP nejsou vhodné zdroje týkající se zkoušky HSK k dispozici, a tak jsem využívala také odborných a publicistických článků volně přístupných na internetu. Na téma HSK bylo v ČLR napsáno mnoho článků, ze kterých rovněž čerpám. K dohledání relevantních zdrojů jsem použila databázi CNKI (China National Knowledge Infrastructure), do níž měla univerzita po omezenou dobu volný přístup. Informace o HSK čerpám také z přímo z oficiálních webových stránek Hanbanu (*Hànkǎo Guójì* 汉考国际) a z knihovny Katedry asijských studií FF UP. V knihovně KAS je k vypůjčení mnoho příruček a sylabů testu HSK jak v nové, tak ve staré formě.

Při zpracování diplomové práce se z velké části opírám o zahraniční zdroje, protože informací a spolehlivých zdrojů o zkoušce HSK je v českém jazyce příliš málo. V tomto směru bude má diplomová práce přínosná i jako zdroj informací, které jsou v českém jazyce prakticky nedohledatelné.

4 TEORETICKÁ ČÁST

4.1 Vznik zkoušky HSK

HSK je standardizovaná zkouška vyvinutá za účelem posouzení úrovně ovládnutí čínštiny pro studenty s jiným mateřským jazykem (Wang 1993, s. 74). Vývoj standardizované zkoušky z čínského jazyka započal v Pekingu roku 1982. Roku 1984 byla založena *Hànyǔ Shuǐpíng Kǎoshì Shèjì Xiǎozǔ* 汉语水平考试设计小组 (*Skupina pro vývoj zkoušky čínštiny jako cizího jazyka*), která měla za úkol vyvinout standardizovaný test čínštiny. Test, který skupina navrhla, byl v letech 1985–1988 nejprve testován na poměrně malém počtu studentů (362–514) z 82 různých států (Liu 1989, s. 90). Posléze roku 1988 byla zkouška dále spravována a vyvíjena na Pekingské jazykové univerzitě (Ren 2001, s. 63). V roce 1990 byla zkouška poprvé oficiálně zavedena na univerzitách v Pekingu, Šanghaji, Wuhanu, Dalianu a Tianjinu. Od roku 1991 byla zavedena i v zahraničí, nejdříve v Singapuru, Austrálii a Japonsku (Wang 1993, s. 74).

4.1.1 Vývoj původní zkoušky HSK

Původní test HSK byl vytvořen po vzoru amerického testu TOEFL (Liu 1999, s. 18). Kromě vlivu TOEFL byl starší formát HSK také inspirován teorií a strukturou anglických standardizovaných testů (Liu 1989, s. 87). Cílem zkoušky bylo posoudit úroveň ovládnutí čínštiny zahraničních studentů zejména pro účely studia a pobytu v Číně. Také měl objektivně posoudit schopnost komunikovat v čínštině (Liu 1989, 84). Kromě toho byla zkouška používána pro určení úrovně ovládnutí čínštiny příslušníků etnických menšin v Číně (Song 1997, s. 69–70).

Jako nejlépe dostupná standardizovaná zkouška čínského jazyka pro cizince se HSK stala důležitým ukazatelem kompetence v čínském jazyce. Výsledky testu byly zohledněny při podávání žádostí o stipendium pro studium v Číně a důležitým faktorem pro přijetí ke studiu na čínských univerzitách. Jako ukazatel kompetence byla používána i při hledání práce v Číně nebo k doložení úrovně čínštiny pro překladatele (Wang 1993, s. 77).

Zpočátku se zkouška skládala pouze ze základní (*chūjí* 初级) a středně pokročilé (*zhōngjí* 中级) úrovně, podle dosažených bodů byly potom výsledky dále rozděleny na 8 stupňů (Liu 1989, s. 88). O dva roky později byla struktura zkoušky přepracována a rozdělena na tři úrovně, s nově připojenou úrovní pro pokročilé (*gāojí* HSK 高级) (Liu 1999, s. 19). V roce 1997 byla přidána úroveň pro začátečníky (Sun 2007, s. 130). Nové složení zahrnovalo nejnižší začátečnickou úroveň (*jíchū* 基础), původně rozdělené základní a středně pokročilé úrovně byly spojeny do jednoho testu (*chū, zhōngděng* HSK 初、中等 HSK). Nejvyšší byla pokročilá úroveň HSK (*gāojí* 高级) (Liu 1999, s. 19).

Struktura staré verze začátečnické úrovně se skládala z poslechu, čtení a gramatických struktur. Základní/středně pokročilá úroveň obsahovala čtyři části: poslechovou část, gramatické struktury, čtení a vepisování znaků. Poslech, gramatické struktury a porozumění přečtenému textu tvořily otázky s více odpověďmi, přičemž jen jedna byla správná. Psaná část vyžadovala ruční psaní znaků. Pokročilá úroveň navíc obsahovala psaní delšího textu a mluvenou část (Ni, 2000, s. 11).

Test byl rozdělen do tří úrovní, ke kterým byly podle získaných bodů přiřazeny na škále 1 až 11 certifikáty s dosaženým výsledkem (Liu 1999, s. 17). Hodnocení zkoušky bylo již od začátku vývoje rozděleno dvěma způsoby. První rovina testu hodnotila úroveň kompetence (začátečník, středně pokročilý atd.), druhá rovina potom úspěšnost zvládnutí testu (Liu 1989, s. 87). Ta byla určena číselným označením a známkami A, B a C. Pokud si student zvolil začátečnickou úroveň testu, byl v případě úspěšného absolvování zkoušky zařazen do stupně 1 se známkou C, 2 se známkou B či při výborném výsledku do stupně 3 s hodnocením A. Test základní/středně pokročilé úrovně pokrýval stupně 3 až 8. Jeho nejnižší stupeň se tedy překrýval s nejvyšším stupněm začátečnické úrovně a byl v této úrovni hodnocen známkou C v rámci základní úrovně. Základní i středně pokročilá úroveň byly opět ohodnoceny známkou A až C, stejně jako test pokročilé úrovně, který pokrýval stupně 9 až 11 (Liu 1999, s. 17).

Ke každé úrovni byl přiřazen seznam doporučené slovní zásoby. Seznam začátečnické úrovně (*jiǎjí* 甲级) obsahoval 1 033 lexikálních jednotek. Základní úroveň (*yǐjí* 乙级) měla rozsah 3 052 slov, seznam středně pokročilé úrovně (*bǐngjí* 丙级) měl 5 257 slov a seznam nejvyšší úrovně (*dīngjí* 丁级) obsahoval 8 840 slov. Součet

lexikálních jednotek byl kumulativní, seznamy vyšších úrovní již obsahovaly lexémy úrovní nižších (Unité des études Chinoises 2014).

Tabulka č. 1: Přehled úrovní staré verze zkoušky HSK

Úrovně testu	Doporučená slovní zásoba	Stupně kompetence	Známka
Začátečnická úroveň	1 033 lexikálních jednotek	1	C
		2	B
		3	A
Základní/středně pokročilá úroveň	3 052 lexikálních jednotek	3	C
		4	B
		5	A
	5 257 lexikálních jednotek	6	C
		7	B
		8	A
Pokročilá úroveň	8 840 lexikálních jednotek	9	C
		10	B
		11	A

Zdroj: Liu 1999, s. 17

Od spuštění starší formy HSK popularita této zkoušky neustále stoupala. V roce 2006 dosáhla účast na zkoušce dohromady jednoho milionu studentů (Meyer 2014, s. 22). Během této doby došlo ke konfliktu ohledně dalšího vývoje zkoušky mezi Pekingskou jazykovou univerzitou, kde byla zkouška od počátků vyvíjena a spravována, a kanceláří Hanban, která byla ministerstvem školství PRC pověřena organizováním a podporováním výuky čínštiny v zahraničí (Meng 2011).

4.1.2 Vývoj zkoušky 改进版 HSK

Na Pekingské jazykové univerzitě byla roku 2007 představena nová, upravená forma zkoušky (*gǎijìnbǎn HSK* 改进版 HSK) (BLCU HSK zhongxin 2008, s. 140). Práce na reformování původní zkoušky HSK započaly již v roce 2000, po sedmiletých přípravách měla postupně nahradit staré HSK (BLCU HSK zhongxin „HSK

gaijinban“ xiangmu xiaozu 2007, s. 131). Nová podoba však nezískala podporu Hanbanu, a tak po několika let existovaly vedle sebe dva formáty zkoušky HSK. Jeden pod správou Pekingské jazykové univerzity a jeden provozovaný kanceláří Hanban. Upravená forma zkoušky byla ovšem nabízena jen na univerzitách v Číně, zatímco formát prosazovaný organizací Hanban byl přístupný na univerzitách v Číně i v zahraničí (Meng 2011).

Upravená podoba zkoušky se na rozdíl od nové zkoušky HSK, zavedené o dva roky později, hlásí k původnímu testu, ze kterého vychází. Zatímco se 改进版 HSK považuje za vylepšenou starou formu zkoušky (BLCU HSK zhongxin „HSK gaijinban“ xiangmu xiaozu 2007, s. 127), u nových HSK má jít o zkoušku zcela přepracovanou (Hankao Guoji 2011). V mnoha ohledech má však současná nová podoba zkoušky blíže té původní, protože změny v upravené zkoušce HSK byly ještě markantnější než v dnešní verzi.

Upravená zkouška HSK omezila původní čtyři úrovně náročnosti a 11 stupňů kompetence na tři úrovně (základní, středně pokročilou a pokročilou). To byla reakce na předchozí kritiku přílišné složitosti původní škály (BLCU HSK zhongxin „HSK gaijinban“ xiangmu xiaozu 2007, s. 128). Škála známek A, B a C byla omezena na slovní ohodnocení „splnil“ a „splnil s vyznamenáním“. Bodové ohodnocení bylo naopak rozšířeno, v každé úrovni bylo možné získat dohromady 1000 bodů (BLCU HSK zhongxin „HSK gaijinban“ xiangmu xiaozu 2007, s. 133).

Nově byly také ve všech úrovních náročnosti testovány hlavní jazykové kompetence: čtení, mluvení, psaní a poslech (Wang 2008, s. 8). Část testující pasivní znalost jazyka (čtení, poslech) byla navržena ve formátu otázek s několika předem danými odpověďmi. Ta se od původní podoby příliš nelišila, hlavní změnou bylo zrušení samostatného gramatického oddílu a jeho zakomponování do ostatních testovaných kompetencí (Li 2009, s. 96). Druhá a třetí část zkoušky hodnotila schopnost studenta jazyk aktivně používat a obsahovala mluvený a psaný oddíl (BLCU HSK zhongxin „HSK gaijinban“ xiangmu xiaozu 2007, s. 133). Tyto tři části na sobě byly nezávislé, zájemce se mohl zaregistrovat a zúčastnit se všech všech tří, nebo třeba jen jedné části (Wang 2008, s. 8).

Tvůrci zkoušky kladli důraz na skutečné užití jazyka, všechny úkoly byly vytvořeny z moderních jazykových korpusů a skladovány v databázi pro případné další užití v jiném termínu (BLCU HSK zhongxin „HSK gaijinban“ xiangmu xiaozu 2007, s. 134). Obsah úkolů byl navržen tak, aby byla obsažena běžná a aktuální témata a zajištěna jejich rozmanitost. Dále byl s ohledem na studenty ve všech úrovních omezen počet úkolů a byla zkrácena doba trvání části obsahující čtení a poslech (BLCU HSK zhongxin „HSK gaijinban“ xiangmu xiaozu 2007, s. 132).

Změny ve struktuře zkoušky byly podle autorů provedeny jako reakce na podněty a potřeby zájemců o HSK (BLCU HSK zhongxin 2008, s. 140). Stejně jako v novém formátu zkoušky se tvůrci snažili přizpůsobit zkoušku poptávce a rychle rostoucímu trhu. Nakonec se ale na trhu prosadila více propagovaná nová zkouška HSK pod správou Hanbanu.

4.1.3 Vývoj nové zkoušky HSK

Díky propagaci HSK v Konfuciových institutech, které byly od roku 2005 zřizovány v zahraničí, byl původní formát zkoušky podporovaný Hanbanem mnohem úspěšnější než ten z Pekingské jazykové univerzity (Hankao Guoji, 2012b). V roce 2010 upravená forma zkoušky 改进版 HSK zanikla, protože Hanban získal výlučné právo na organizaci zkoušky (Meng 2011). Na přelomu let 2009/2010 Hanban oficiálně zavedl zkoušku HSK v nové reformované podobě (Hankao Guoji 2009b).

Nová zkouška HSK je výsledkem několikaletých snah o adaptaci formátu zkoušky na potřeby studentů (Hankao Guoji 2009b). Úrovně zkoušky byly přepracovány podle Společného evropského referenčního rámce k vyučování a hodnocení znalosti cizích jazyků (zkráceně SERR, anglický akronym CEFR) (Hankao Guoji, 2012c). SERR byl vyvinut Radou Evropy pro klasifikaci znalostí cizích jazyků v rámci Evropské Unie (Rada Evropy ©2015). Nová zkouška HSK byla rozdělena do úrovní 1 až 6, přičemž každá úroveň odpovídá jednomu stupni v klasifikaci SERR. Úrovně A1 a A2 odpovídají začátečníkům, B1 a B2 středně pokročilým a C1 a C2 pokročilým uživatelům cizího jazyka (Hankao Guoji ©2013–2016b).

Tabulka č. 2: Přehled úrovní nové verze zkoušky HSK

Úrovně testu	Doporučená slovní zásoba	Stupně kompetence SERR podle Hanbanu	Stupně kompetence podle Fachverband Chinesisch
HSK1	150 lex. jednotek	A1	–
HSK2	300 lex. jednotek	A2	A1 (bez znalosti znaků)
HSK3	600 lex. jednotek	B1	A1
HSK4	1200 lex. jednotek	B2	A2
HSK5	2500 lex. jednotek	C1	B1
HSK6	5000 lex. jednotek	C2	B2

Zdroj: Hankao Guoji ©2013–2016b; Fachverband Chinesisch 2010

Stejně jako u starší formy byla ke každé úrovni přiřazena doporučená slovní zásoba, která byla ale nyní výrazně zredukována. Toto rozdělení i doporučený rozsah slovní zásoby zůstaly stejné i po reformě slovní zásoby v roce 2012.

Struktura nové zkoušky HSK byla také přepracována. Stejně jako dříve vybírali studenti u čtení a poslechu z několika předem daných odpovědí na otázku. V testu HSK 1 a 2 je testováno jen porozumění čtenému textu a poslechu. Studenti ani nemusejí znát čínské znaky, celý test je psán ve znacích a pinyinu zároveň. V HSK 3, 4 a 5 je testováno porozumění čtenému textu, poslech a psaní znaků. Psaná část testuje skládání gramaticky správných vět z předem daných znaků a psaní znaků v rámci jednoduchých vět. Teprve u HSK 6 je vyžadováno napsání delšího koherentního textu (Hankao Guoji 2013). Další části HSK se od nižších úrovní liší jen náročností.

Nově od roku 2010 Hanban vytvořil nezávisle na písemné zkoušce i ústní zkoušku HSKK (*Hànyǔ Shuǐpíng Kǒuyǔ Kǎoshì* 汉语水平口语考试) pro všechny tři stupně kompetence. Je rozdělena do tří úrovní na základní, středně pokročilou a pokročilou úroveň. Požadavky na slovní zásobu jsou podle informací o ústní zkoušce na oficiálních stránkách organizace Hanban nižší než pro písemnou zkoušku (Hankao Guoji ©2013–2016c). Hovorový jazyk bývá formalizovanější než psaná forma jazyka, proto si lidé vystačí s menší slovní zásobou (Milton 2009, s. 55). Pro základní úroveň (*HSKK chūjí* HSKK 初级) je třeba znát pouze 200 frekventovaných slov, pro střední úroveň (*HSKK zhōngjí* HSKK 中级) stačí 900 slov a pro pokročilou úroveň (*HSKK*

gāoji HSKK 高级) 3 000 lexémů (Hankao Guoji, ©2013–2016c). O která frekventovaná slova (*chángyòngcí* 常用词) se konkrétně jedná, Hanban neuvádí.

Zkouška HSK před reformou i po reformě byla po formální stránce silně ovlivněna zahraničními standardizovanými multiple-choice testy. Zatímco starší HSK měly za vzor americký TOEFL test, současná forma se orientuje na evropské standardizované testy a formou se podobá testům, jako jsou britské Cambridge ESOL zkoušky nebo německý Goethe Zertifikat. Starší a nová forma zkoušky HSK se od sebe liší primárně rozdělením do úrovní a odlišnou slovní zásobou.

4.2 HSK v České republice

Na Univerzitě Palackého v Olomouci bylo možné složit již starší formu zkoušky HSK. Pracoviště Katedry asijských studií bylo jedno z mála míst ve Střední Evropě, kde bylo možné zkoušku absolvovat. Úspěšné složení zkoušky HSK bylo svého času dokonce podmínkou pro přípuštění ke státním závěrečným zkouškám. Studentům s vynikajícími výsledky testu bylo navíc poskytnuto stipendium Ministerstva školství ČLR pro studium v Číně (Katedra asijských studií, nedatováno).

Konfuciov institut při Univerzitě Palackého byl založen již v září roku 2007 (Konfuciova akademie ©2007–2016), ale první nová zkouška HSK byla zorganizována až dva a půl roku po její reformě. V květnu roku 2012 bylo poprvé možné opět složit zkoušku HSK v České republice, zpočátku pouze úrovně HSK2, 3 a 5 (Hankao Guoji 2012d). Dnes se v Konfuciově institutu při Univerzitě Palackého organizují zkoušky HSK úrovní dvě až šest i všechny tři úrovně HSKK (Konfuciova akademie UP v Ol. ©2007–2016a). Konfuciov institut v Olomouci dosud zůstává jediným pracovištěm organizujícím testy HSK i přes otevření pražské pobočky v roce 2014 (Embassy of the PRC in CR 2014). Úspěšné absolvování zkoušky již není podmínkou k završení studia. Na základě výsledků testu HSK je však od roku 2012 možné požádat o stipendium Konfuciova institutu (Konfuciova akademie UP v Ol. ©2007–2016b).

4.3 Formát zkoušky HSK

Tato kapitola se zabývá formátem zkoušek zkoumaných úrovní v souvislosti s deklarovanými stupni jazykové kompetence. V praktické části bude zkoumáno, zda odpovídá doporučená slovní zásoba stipulovaným schopnostem ovládnutí čínštiny. V této kapitole uvedu a porovnáám možnosti využití slovní zásoby a testování různých dovedností v rámci původního testu pro začátečníky HSK (基础) a reformované zkoušky 新 HSK4.

4.3.1 Formát zkoušky HSK (基础)

Začátečnická úroveň je nejnižším dosažitelným stupněm zkoušky s nejnižšími nároky. Testované komunikační dovednosti spadají do tří oblastí: komunikace v každodenním životě, komunikace v pracovním prostředí a porozumění literatuře (Jiang 1999, s. 31). Hlavní důraz je kladen na komunikaci v běžném životě, při níž by student měl zvládnout jednoduchou konverzaci o tématech, jako jsou jídlo, bydlení, cestování, koníčky, a zdvořilostní obraty, jako jsou pozdravy, poděkování, rozloučení atd. (Liu 1998, s. 15).

K úspěšnému složení zkoušky je uváděno absolvování 100–800 studijních jednotek a ovládnutí slovní zásoby o rozsahu 1 033 slovíček (Liu 1998, s. 16). Požadavky na rozsah slovní zásoby testovaného studenta jsou ovšem mnohem širší (600–3 000 vysoce frekventovaných slov). Samotný test by měl obsahovat 80 % doporučené základní slovní zásoby úrovně 甲级. Zbytek zahrnutých slov pochází ze slovníčku vyšších úrovní a zhruba 5 % ze slov, která nejsou součástí doporučené slovní zásoby (Jiang 1999, s. 32).

Test se skládá ze tří částí, a to z poslechu, gramatických vazeb a čtení textu. Všechny části budou popisovat na základě příkladových testů publikovaných v sylabu HSK začátečnické úrovně (Liu 2000, s. 18–132).

Poslech je rozdělen na tři části a tvoří 36 % celého testu. Tato část má formu 50 otázek s výběrem ze čtyř odpovědí, z nichž pouze jedna je správná. Každou otázku testovaný student slyší dvakrát. Poslech je dále rozdělen na tři oddíly, prvním je přiřazování odpovídajících obrázků souvisejících s obsahem slyšené věty, druhým je

vybírání odpovědi na slyšenou otázku a třetím je vybírání odpovědi na otázku týkající se krátkého dialogu. Všechny otázky i nabízené odpovědi jsou tvořeny z krátkých jednoduchých vět. Co se týče struktury a náročnosti otázek, odpovídá nárokům vytyčeným tvůrci zkoušky.

Část testující znalost gramatických struktur má dva oddíly čítající dohromady 40 otázek, což představuje 29 % obsahu testu. První oddíl testuje znalost syntaxe, student musí ze 4 možností poskládat jednoduchou větu vybrat jedinou správnou. Druhá část testuje znalost užití synsémantik. V každé větě chybí jedno slovo, ze čtyř možností je třeba vybrat např. vhodné měrové jméno, adverbium nebo předložkovou vazbu. Testované gramatické vazby souvisejí se seznamem 135 gramatických struktur (Jiang 1999, s. 35), studenti se tak mají z čeho připravovat.

Poslední částí testu je čtení, které má stejný počet otázek jako poslechová část. Zaměřuje se na pochopení smyslu kratších či delších úseků textu. V prvním oddílu je třeba do souvislého textu doplnit chybějící čtyři lexémy. Jako ve všech ostatních případech stačí vybrat z předem nabízených možností. V druhém oddílu se testuje porozumění textu o rozsahu krátkého odstavce, ke kterému je připojena otázka týkající se textu. Posledních deset otázek třetího oddílu je podobných jako ty předchozí, jedná se ale o delší texty a ke každému textu může být připojeno více otázek.

Test HSK začátečnické úrovně má přehlednou strukturu, ale příliš omezený rozsah testovaných kompetencí. Gramatická část je úzce spjatá se čtením s tím rozdílem, že v prvním oddílu gramatických úkolů je třeba doplňovat synsémantika a v porozumění čtenému textu se doplňují autosémantika. V obou částech se v podstatě jedná o správné zařazení slov s ohledem na syntax a význam. I v gramatické části je třeba nejdříve porozumět textu a až poté lze správně určit chybějící slova. Důraz na čtení je kladen také v pokynech k úkolům, které jsou svým obsahem a formulací v některých případech komplikovanější než samotné úkoly. Přečtení a pochopení úkolu zabere příliš mnoho času, pokud se student předem s formátem testu neseznámí.

Začátečnická úroveň zkoušky testuje pouze pasivní jazykové dovednosti, tedy poslech a čtení. Jak student po ovládnutí doporučené slovní zásoby dokáže mluvit a psát, je sice uvedeno v sylabu úrovní HSK, tyto důležité části jazykové kompetence však v této úrovni testu zcela chybějí.

4.3.2 Formát zkoušky 新 HSK4

Podle prohlášení na stránkách organizace Hanban z roku 2009 je nová zkouška HSK čtvrté úrovně srovnatelná s původní začátečnickou úrovní HSK (Hankao Guoji 2009b). O dva roky později je v jiném prohlášení HSK4 srovnávána s úrovní o stupeň vyšší, základní úrovní původní verze HSK (Hankao Guoji 2011). V obou prohlášeních je uvedeno, že je výsledkem porovnání obou verzí zkoušek, ale ani v jednom není uvedeno, podle jakých kritérií. Organizace Fachverband Chinesisch se přiklání ke stejnému rozdělení jako Hanban v roce 2009, tedy že je nová zkouška HSK4 na stejné úrovni jako HSK (基础) (Fachverband Chinesisch 2010).

Student, který má zájem složit zkoušku HSK4, by měl být schopen konverzace o široké škále témat a měl by se umět plynně domluvit i s rodilými mluvčími (Hankao Guoji ©2013–2016b). Této úrovně lze podle údajů na stránkách Hanbanu dosáhnout po 4 semestrech studia zahrnujících 2–4 lekce čínštiny týdně (Hankao Guoji, ©2013–2016b). Nároky na středně pokročilou úroveň HSKK jsou velmi podobné (2–3 lekce týdně po dobu dvou let), jen požadovaný rozsah slovní zásoby je nižší o 300 lexikálních jednotek (900 slov) (Hankao Guoji, ©2013–2016c). Podle Miliona (2009, s. 58), je znalost méně než 1 000 lexémů pro správné porozumění hovorovému jazyku nedostatečná, pokud se konverzace neskládá z formalizovaných frází.

Podle výzkumu Zhoua (2012, s. 2) obsahovaly zkoumané testy HSK4 z 91,75 % slovíčka z doporučené slovní zásoby. Navíc uvádí, že jsou testy navrženy tak, aby slovíčka mimo doporučenou úroveň slovní zásoby tvořily max. 15 % obsažených lexémů.

Písemný test HSK4 je složen ze tří částí, a to z poslechu, četby a z psaní. Poslechová část má tři oddíly. V prvním oddílu student slyší krátkou promluvu a poté její shrnutí do jedné věty. Student musí posoudit, zda shrnutí odpovídá informacím v promluvě nebo ne. Druhý oddíl je stejný jako v původní zkoušce HSK (基础), po krátkém dialogu následuje otázka a je třeba na ni ze čtyř možností vybrat odpověď. Poslední část tvoří třikrát delší dialogy, po kterých opět následuje otázka. Po skončení poslechové části mají studenti 5 minut, aby vyplnili odpovědi do záznamového archu. Co se týče jazykové náročnosti a rozsahu slyšených vět, je obtížnost poslechové části srovnatelná s úkoly v starším testu HSK začátečnické úrovně.

V prvním oddílu čtené části je na výběr několik slov, která je třeba vhodně zařadit do následujících vět. Druhý oddíl obsahuje otázky tvořené třemi spolu souvisejícími krátkými větami, které je třeba v logickém sledu seřadit. Poslední část je tvořena krátkými či delšími texty s otázkou a čtyřmi možnými odpověďmi. Čtené texty mají podobný rozsah jako v HSK (基础), účastníci testu HSK4 však mohou očekávat, že v textu bude více jim známých slovíček z doporučené slovní zásoby.

Psaná část se trochu podobá původní gramatické části. V prvním oddílu je třeba ze slovíček složit syntakticky správnou větu. V druhém oddílu stačí jednou (i krátkou) větou popsat daný obrázek a použít přitom předem dané slovíčko, např. 破 (*rozbít*) u obrázku s rozbitou skořápkou vajíčka nebo 困 (*unavený*) u obrázku zívající ženy. V psané části jsou požadavky na schopnosti psaní v čínštině velmi nízké (v porovnání se standardem úrovně B2), zvláště u popisu obrázků. Vzorová odpověď obsahuje pouze pět slov. Z vlastní zkušenosti (a po zhlédnutí šesti ukázkových testů) vím, že stačí naučit se napsat několik jednoduchých frází a potom je u obrázků univerzálně použít, neboť se sice obrázky obměňují, ale tematicky zůstávají stejné. Psaná část testu proto o schopnostech studenta psát v čínštině vypoví velmi málo.

Výrazně byl zlepšen formát zadání úkolů. Zadání je nyní omezeno na jednu krátkou větu, po které následuje názorné uvedení jednoho příkladu a jeho řešení. To umožňuje studentům rychle se v textu zorientovat a správně pochopit, co se od nich očekává.

V porovnání se zkouškou HSK (基础) je obsah testu srovnatelně náročný. Mnoho úkolů má stejnou strukturu i obtížnost. Část čtení a poslechu zůstává podobná a gramatické úkoly jsou částečně stejné jako v psaném oddílu. Zcela nové je popisování obrázků, tyto úkoly ale nemají velkou vypovídací hodnotu.

4.4 Hodnocení a kritika zkoušky HSK

4.4.1 Hodnocení a kritika staré formy zkoušky HSK

Stará verze HSK testu byla kritizována v několika ohledech. Podle Penga (1998, s. 58) byla reforma starého formátu HSK nutná. Peng kritizoval absenci mluvené části či produktivní psané části, které kromě pokročilé úrovně ve starším testu chyběly. Zmiňuje

také formát testu, ve kterém princip jediné správné možnosti podporuje pouze testování pasivních znalostí, u nichž je také jednodušší bez znalostí uhádnout správnou odpověď.

Meyer (2009) kritizuje komplikovanou stupnici úrovní a hodnocení testů. V další studii o testování kompetence v čínském jazyce Mayer uvádí, že podle výzkumu, který provedl Da Jun, se doporučená slovní zásoba příliš kryla se slovní zásobou používanou v žurnalistických textech (Meyer, 2009, s. 28). Kritizován byl také přílišný důraz, který je kladen na pasivní znalost znaků, jež jsou ve staré verzi HSK používány výlučně ve všech úrovních. Tímto jsou podle výzkumu zvýhodněni korejšští a japonští studenti (Meyer, 2009, s. 32).

Kromě toho byl test komplikován tím, že v úrovni pro začátečníky sice nebyla testována znalost příliš složitých slov a znaků, ale v pokynech k úkolům někdy složitá slova použita byla a špatně pochopené zadání mohlo negativně ovlivnit výsledky celé testované části (Chen 1993, s. 133).

Andreas Guder poukazuje na to, že nároky na zvládnutí jednotlivých úrovní zkoušky byly uváděny neobjektivně ve „studijních jednotkách“ (*xuéshí* 学时), které je třeba pro dosažení kompetence v dané úrovni absolvovat. Přitom samotné studijní jednotky nejsou Hanbanem ani Pekingskou jazykovou univerzitou blíže definovány. Podle něj se jednalo o příliš obecná a navíc zavádějící kritéria (2005, s. 64).

Meyer u podobných standardizovaných testů pozitivně vnímá zejména jejich vysokou úroveň objektivity a spolehlivosti a také skutečnost, že objektivně vyhodnocují schopnost zahraničních studentů studovat na čínských univerzitách (Meyer 2015, s. 12).

4.4.2 Hodnocení a kritika zkoušky 改进版 HSK

Pekingskou jazykovou univerzitou reformovaná verze HSK 改进版 se neosvědčila a nyní se již nepoužívá. Přesto se zreformovaná podoba zkoušky snažila napravit některé nedostatky staré formy testu. Reformovaná verze kladla větší důraz na testování aktivních znalostí jazyka oproti pasivním. Mluvená a psaná část zkoušky testovala skutečné schopnosti studenta, nešlo pouze o vybírání z předem daných možností. Testování znalostí gramatiky bylo nově zakomponováno do všech tří částí (BLCU HSK zhongxin „HSK gaijinban“ xiangmu xiaozu 2007, s. 133).

Wang (2006) na reformované verzi HSK kritizuje, že forma zkoušky byla zcela přepracována, ale doporučená slovní zásoba nikoliv. Slovní zásoba zůstala rozdělena na čtyři úrovně, zatímco zkouška měla jen tři, a nebyla tak s formátem testu kompatibilní. Wang pokládá rozdělení na tři úrovně za příliš zjednodušující a navrhuje rozdělení na šest úrovní v podobě, která byla později uskutečněna v nové HSK navržené Hanbanem (Wang, 2006, s. 8).

V době, kdy vedle sebe fungovaly verze HSK Pekingské jazykové univerzity a Hanbanu zároveň, se kvůli podobě testů obě instituce navzájem kritizovaly. Tyto spory obou organizací velmi negativně hodnotila také veřejnost, jelikož například studentům nebylo jasné, který test mají pro své účely absolvovat, jaký je v testech rozdíl a která podoba je uznávanější (Meng, 2011). Za současné situace, kdy existuje pouze jedna podoba zkoušky HSK, se proto studenti čínštiny, školy i zaměstnavatelé mohou lépe orientovat a posoudit své potřeby.

4.4.3 Hodnocení a kritika nové zkoušky HSK

V nové podobě testu HSK jsou úrovně testu rozděleny na šest úrovní podle Společného evropského referenčního rámce. Toto rozdělení je považováno za značně nepřesné a nevyhovující (Meyer, 2014, s. 22). Také německá organizace Fachverband Chinesisch kritizuje nový způsob rozdělení úrovní. Podle nich tento pokus Hanbanu o informování západních studentů ohledně nároků na zvládnutí čínštiny spíše věci uškodí, neboť Hanbanem poskytované informace jsou vzdálené realitě (Fachverband Chinesisch ©2011–2016).

Kritizován je také příliš omezený rozsah doporučené slovní zásoby a příliš nízko nastavená udaná doba studia ke zvládnutí úrovní kompetence v čínštině. Fachverband Chinesisch naopak však chválí novou příležitost k certifikaci i na nízké úrovni kompetence, což slouží jako důležitý motivační faktor. Nekritizují používání SERR *per se*, navrhují jen odpovídající ohodnocení náročnosti testu. Podle nich spadají úrovně HSK2 až HSK4 do úrovně začátečníků a pokročilé úrovně HSK5 a HSK6 splňují pouze nároky úrovní pro středně pokročilé (Fachverband Chinesisch 2010).

Nesrovnalosti v požadavcích na dosažení jednotlivých úrovní jsou patrné také při srovnání s Tchajwanskou standardizovanou zkouškou z čínštiny jako cizího jazyka

TOCFL. Úrovně TOCFL byly roku 2014 přizpůsobeny Evropskému referenčnímu rámci (Education Division TECO ©2011). Požadavky na rozsah slovní zásoby testu TOCFL jsou velmi podobné nárokům staré verze zkoušky HSK.

Tabulka č. 3: Porovnání úrovní nové zkoušky HSK a TOCFL podle SERR

Udávaná úroveň podle SERR	Doporučená doba studia k dosažení úrovně		Doporučená slovní zásoba (počet ovládnutých lexikálních jednotek)	
	HSK	TOCFL	HSK	TOCFL
A1	0,5 roku	1–2 roky	150	500
A2	1 rok	1–2 roky	300	1 000
B1	1,5 roku	2–3 roky	600	2 500
B2	2 roky	3–4 roky	1 200	5 000
C1	více než 2 roky	1920–3840 hodin	2 500	8 000
C2	neuvedeno	více než 3840	5 000	8 000

Zdroj: Hankao Guoji ©2013–2016b; Education Division TECO ©2011

Novému formátu zkoušky lze také vyčítat jeho komercializaci. Již s minimálními znalostmi čínského jazyka obdrží student certifikát, podle kterého by měl na nejnižší úrovni být schopen napsat pohlednici a vyplnit jednoduché formuláře (Rada Evropy 2001, s. 26). Přitom ale k absolvování testu nemusí vůbec znát znaky. Snadná zvládnutelnost nižších úrovní testu vedla k rychlé popularizaci zkoušky a zvýšení vlivu Konfuciových institutů (Zhang 2012, s. 2). Organizování zkoušek je nyní výnosnější, neboť za písemnou a ústní část zkoušky je třeba platit zvlášť. Hanbanem zamýšlené zaměření se na potřeby studentů a zvýšení motivačního faktoru (Hanban 2012c) proběhlo na úkor snížení hodnověrnosti testu (Zhang 2012, s. 8).

4.5 Specifika čínštiny a slovní zásoba

Jak bylo již naznačeno výše, nároky na dosažení jistého stupně kompetence v čínském jazyce a nároky k získání stejných kompetencí v indoevropských jazycích se pro západní studenty velmi liší, což je patrné i na problémech při přizpůsobení zkoušky HSK Společnému evropskému referenčnímu rámci. Adaptace zkoušky HSK na SERR měla sloužit hlavně k přehlednění situace a motivaci studentů, kteří ovšem budou dříve

či později s komplikacemi při studiu čínštiny konfrontováni. He (2010, s. 3) ve své práci zmiňuje, že ovládnutí čínštiny trvá britským studentům třikrát až čtyřikrát déle než ovládnutí francouzštiny či italštiny.

I v průzkumu z roku 2003 mezi německými akademiky, kteří promovali v oboru čínských studií a minimálně 3 roky se intenzivně zabývali studiem čínštiny, nebylo sebehodnocení znalosti čínštiny nijak vysoké. Každý účastník se sám ohodnotil v různých jazykových kompetencích podle referenčního rámce vyvinutého organizací ALTE (Association of Language Testers in Europe). V dotazníku byly otázky týkající se kompetence v psaní, čtení a mluvené komunikace. Úrovně kompetence podle ALTE se skládají z pěti stupňů, od začátečníků (1) po pokročilé (5). Po vyhodnocení dotazníků bylo zjištěno, že nejvýše se účastníci ohodnotili ve čtení, a to známkou 3,42. Komunikační kompetence byla v průměru ohodnocena známkou 3 a nejnižší hodnocení (2,07) se týkalo schopnosti psaní v čínštině (Guder 2005, s. 65–66).

V souvislosti s tímto výzkumem Guder poukazuje na to, že u studentů, jejichž mateřský jazyk patří do indoevropské jazykové větve, je nutné rozlišovat mezi příbuznými a vzdálenými jazyky. Také je třeba vzít v úvahu, že tato vzdálenost je nejen lingvistická, ale navíc i kulturní. V případě velké vzdálenosti mezi jazyky (jako tomu je v případě studia čínštiny) je třeba diferencovat nároky k dosažení úrovně a také jim přizpůsobit studijní plány (Guder 2005, s. 75–76).

Čínské lexikum je problematické zejména pro studenty-záčátečníky. Výslovnost některých hlásek čínských slov je češtině velmi cizí, stejně jako tónová stupnice (Kane, 2009, s. 34). Bez znalosti znaků lze kvůli častým homofonům v čínštině postoupit jen na poměrně nízkou úroveň kompetence (Kane 2009, s. 28). Ze znaků nelze odvodit výslovnost (pokud se nejedná o fonogramy), pasivní a aktivní osvojení si grafické stránky znaků je pro studenty uvyklé latinské abecedě také velmi náročné. Porozumění čínskému textu komplikuje i absence mezer mezi slovy.

Kromě toho není pro západní studenty snadné porozumět v čínštině důležitému vztahu mezi morfémy a slovy. Kompozice je nejčastější způsob tvoření slov v čínštině, proto je při jejím osvojování užitečné rozšiřovat znalost samotných bází a nejen již složených slov (Kane 2009, s. 38). Podle McBride-Chang et al. (2006, s. 105) je kolem 80 % čínských slov tvořeno kompozicí.

Ovládnutí slovní zásoby je pro studenty cizího jazyka klíčové. Bez dostatečné slovní zásoby nelze jazyku správně porozumět ani jej aktivně používat. Zatímco textu nebo promluvě lze porozumět, i pokud obsahuje gramatické chyby, chybné použití slov či neporozumění slovům již představuje podstatnou překážku v komunikaci. Kromě toho čím bohatší znalost slovní zásoby student čínštiny má, tím lépe rozumí čtenému i mluvenému jazyku (Krames 2013, s. 50). Slovní zásoba je jedinečná i z hlediska rodilých mluvčích. Zatímco znalosti fonetiky a nejdůležitější znalosti gramatiky jazyka dítě vstřebává ve věku kolem sedmi let, slovní zásoba člověka se během života neustále rozšiřuje (Stubbs 1986, s. 60).

Na didaktiku slovní zásoby čínštiny lze podle Kanga nahlížet ve dvou rovinách (Kang 2011, s. 158). V první řadě je třeba studentům sdělit základní informace o lexému, které by se student měl dozvědět z každého slovníčku či slovníku. Jde o grafickou podobu slova, jeho výslovnost, význam a gramatické funkční charakteristiky. To však samo o sobě úspěšné ovládnutí lexému nezaručí, protože tyto informace neusnadňují osvojení si slovíčka. Do druhé roviny patří systém, podle něhož je soupis slovní zásoby vytvořen. Ten zahrnuje otázku, která slovíčka do seznamu zařadit, kdo je jeho cílovým uživatelem a jak spolu vybrané lexémy souvisejí.

V ideálním případě by podle Kanga (2011, s. 157–158) měl být student poučen o vlastnostech lexému, co se týče:

- **Formy:** již zmiňovaná grafická podoba, výslovnost atd.
- **Použití:** to může být znázorněno pomocí příkladových vět nebo informací o tom, ke kterým slovům či frázím se lexém váže. Patří sem i pragmatické a stylistické vlastnosti slovíčka.
- **Složení:** informace o významu morfémů, ze kterého je slovo složeno, popř. způsob, jakým jsou morfémy spojeny. To ulehčí studentům osvojení si dalších kompozit.

Při sestavování lexikonu cizího jazyka je nutné vzít v úvahu mnoho dalších faktorů. Níže budou nastíněny různé přístupy při navrhování seznamu slovní zásoby pro studenty cizího jazyka. Podle těchto teorií a metod budou v praktické části kvalitativně posuzovány soupisy slovní zásoby.

4.6 Výběr lexémů seznamu slovní zásoby

Základním problémem při sestavování slovníčku je otázka, které lexikální jednotky do soupisu vůbec zařadit. Jelikož zkoumám soupisy pro studenty na poměrně nízké úrovni (A2), soustředím se na dvě metody, které identifikují nejdůležitější a nejzákladnější lexémy v jazyce.

4.6.1 Role frekvence při sestavování slovníčku

Důležitost frekvence lexémů při sestavování lexikonu se projevila zejména s rozšířenou možností počítačového zpracování korpusů (Čermák 2000, s. 15). Díky údajům o frekvenci lexémů lze snadno zjistit, jak často se slova v jazyce objevují. Pokud se studenti se slovíčkem často setkávají, mají větší šanci si je zapamatovat (Milton 2009, s. 22). Vysoce frekventovaná slova tak bývají dříve osvojena (Milton 2009, s. 31).

Vysoká frekvence slov je důležitým údajem o užitečnosti lexému zvláště u studentů-začátečnicků. Naučení se nejběžnějších slovíček usnadní komunikaci a porozumění více než výběr podle jiných kritérií i tehdy, když mají studenti na nízké úrovni znalosti jazyka omezenou slovní zásobu (Alcaraz-Marmol 2015, s. 4). Za spolehlivý indikátor, která frekventovaná slova by měli studenti znát, je považováno 2 000 nejběžnějších lexémů v jazyce (Nation 2001, s. 14–15). U pokročilých studentů není frekvence spolehlivým ukazatelem, které lexémy by se měli naučit, protože s rozsáhlou slovní zásobou studenti běžná slovíčka již znají a rozdíly mezi málo frekventovanými slovy nejsou významné (Milton 2009, s. 42).

Při kompilaci lexikonu je podle Čermáka a Blatné (1995, s. 234) klíčovým faktorem frekvence užití lexémů v přirozeném jazyce. Frekvence vybraných slov má reprezentovat lexikální jednotky jak v psaném, tak v mluveném jazyce. Zastoupeny by ale měly být také reprezentativní pojmy, ačkoliv nemusejí být v jazyce příliš frekventované (například hyperonyma jako jsou *ovoce*, *rostlina* atd.).

Čermák dále zmiňuje využití potenciálu lexikonu k dalšímu rozšíření slovní zásoby. U čínštiny by se tak mohlo jednat o zařazení morfémů, které tvoří stavební jednotku mnoha dalších slov. Posledním důležitým faktorem je vybírání

bezpříznakových lexémů nad příznakovými (např. *malý* a ne *drobný*) (Čermák a Blatná 1995, s. 234).

4.6.2 Základní slovní zásoba

Definice základní slovní zásoby byla vyvíjena již od dvacátých let minulého století, kdy se lingvisté snažili zjistit, která slova jsou v jazyce klíčová. V tomto ohledu byla důležitá publikace *Basic English* z roku 1930, v níž Ogden a Richards určili 850 nejzákladnějších lexémů v angličtině (Stubbs 1986 s. 58). Dnes se zabývá moderní teorií o základních lexikálních jednotkách v jazyce (*core vocabulary/nuclear vocabulary*) americký lingvista Ronald Carter.

Definování základní slovní zásoby je užitečné zejména pro zjednodušení akvizice prvního a druhého jazyka (Carter 1982, s. 39). Přínosem této teorie je určení kritérií, podle kterých mají některé jednotky v jazyce ústřední postavení a je nutné se je naučit nejdříve nebo jsou méně užitečné a mají periferní postavení. Teorie základní slovní zásoby nabízí možnost klasifikovat lexémy bez ohledu na jejich frekvenci. Následující kritéria jsou uvedena v knize *Vocabulary: Applied Linguistic Perspective* (*Slovní zásoba: aplikovaná lingvistická perspektiva*) od Cartera (Carter 2012, s. 49–59).

Kritéria pro určení postavení slov jsou:

- Možnost tvořit nová slova: schopnost vytvářet nové významy v široké škále kontextů, např. čínská báze 电 (*elektřina*) se spojuje do kompozit jako 电脑 (*počítač*), 电梯 (*výtah*), 电话 (*telefon*) apod., a je tak díky své produktivnosti zvláště důležitá.
- Mnohovýznamová slova: slova, u kterých se v průběhu jejich vývoje rozšiřovala škála různých významů, mají v jazyce významné postavení. To platí i o potenciálu slov tvořit nové kolokace.
- Bezpříznaková slova: slova zaujímající středovou pozici na škále mezi formálním a neformálním jazykem. Také jde o slova, která jsou známá všem průměrným, dospělým rodilým mluvčím. Např. slova 会晤 – 见面 – 碰面 mají všechny význam *setkat se*, ale jen 见面 je vhodné použít univerzálně jak v hovorové, tak v psané čínštině.

- Neutrální slova: slova bez citového zabarvení. Např. slovo 瘦 (*štíhlý*) není hanlivé ani lichotné, proto má k základní slovní zásobě blíže než jeho lichotná varianta 苗条. Současně může slovo 瘦 v kontextu sloužit k vyjádření jak negativního, tak pozitivního hodnocení.
- Slova, která bývají definována sémantickými aspekty a ne synonymem: slova 害怕、惧怕、畏惧 všechna znamenají *mít strach*. V pátém vydání slovníku 现代汉语词典 (*Slovník moderní čínštiny*) je pouze u slova 害怕 popsán pocit strachu, ostatní dvě slova jsou definována synonymem.
- Základní slova jsou používána k definování synonym a ne naopak. Např. výše zmíněná slova 惧怕、畏惧 (*mít strach*) jsou obě definována slovem 害怕, to ale není definováno slovem 惧怕, přestože má stejný význam.
- Nadřazené pojmy: hyponyma mají tendenci být součástí základní slovní zásoby, protože jsou jimi definována podřazená slova. Centrální slova jsou spíše všeobecného nežli specializovaného charakteru.
- Slova, která ve svém hlavním významu nemají kulturně specifické asociace. Tyto lexikální jednotky jsou snadno přeložitelné, např. 雨 (*děšť*) 学校 (*škola*). Slova typická pro čínskou kulturu jako např. 粽子 (*sladký nebo slaný rýžový nákyp zabalený v bambusových listech do tvaru jehlanu*) součástí základní slovní zásoby nejsou.
- Snadno identifikovatelná antonyma: pokud je pro lexém snadné najít slovo opačného významu, je pravděpodobnější, že patří do základní slovní zásoby. Např. slova 瘦 – 胖 (*štíhlý a tlustý*) jsou protiklady, které snadno identifikují i děti. Určování přímého protikladu ke slovu 骨立 (*vychrtlý*) je již o něco složitější.

Carter zmiňuje, že k tomu, aby lexikální jednotka byla součástí základní slovní zásoby, nemusí splňovat všechna výše uvedená kritéria. Čím více jich ale splňuje, tím je pravděpodobnější, že zaujímá ve slovní zásobě centrální pozici (Carter 1982, s. 44). V případě pochyb o pozici dvou srovnatelných variant slov je rozhodující frekvence slova (Carter 1982, s. 45).

4.7 Obsah seznamu slovní zásoby

Po identifikování důležitých lexémů vhodných pro danou úroveň studia je třeba myslet na to, jak usnadnit studentům jejich ovládnutí. Seznam slovní zásoby má omezený rozsah a zahrnuté lexémy by spolu měly jistým způsobem souviset. Níže je uvedeno několik metod, jak studentům usnadnit akvizici nových slovíček.

4.7.1 Sémantická průhlednost lexémů

Lopez (2013, s. 2) uvádí přístup k osvojování si slovní zásoby čínského jazyka z hlediska uživatele. Podle něj by neměla být opomíjena frekvence lexému, je však třeba myslet také na úsilí, které je potřeba vynaložit na zapamatování daného lexému. Z tohoto pohledu není frekvence lexémů tolik důležitá, místo toho se student soustředí na využití pravidelností a průhlednosti lexémů (fonetik, piktogramů, symbolů apod.) a na systematizování seznamu lexikálních jednotek tak, aby bylo pro něho co nejjednodušší si je osvojit.

Stejný důraz na transparentnost nových slovíček klade ve své studii i Laufer (Laufer 1989, s. 10–18). Navíc identifikuje problematické skupiny slov, kterým je třeba při sestavování slovníčku věnovat zvýšenou pozornost a zamezit tak nedorozuměním. Laufer nejdříve cituje Ullmana, který definuje pro studenty cizího jazyka snadno zapamatovatelné kategorie průhledných lexémů (Laufer 1989, s. 11):

- Morfologicky průhledná slova: slova, která jsou při znalosti významu jednotlivých bází snadno odvoditelná. V čínštině např. 医院 (*nemocnice*) nebo 飞机 (*letadlo*). Pokud jde o výslovnost, jsou zvláště internacionalismy (fonetické výpůjčky) dobře rozeznatelné, např. 咖啡 (*kāfēi, káva*) nebo 巧克力 (*qiǎokèlì, čokoláda*).
- Onomatopoická slova: zvukomalebná slova napodobují přirozeně se vyskytující zvuky, a tak je lze snadno odvodit. V čínštině mají znaky často determinativ 口 (*ústa*). Mezi transparentní zvukomalebná slova patří např. 哈哈 (*hāhā, haha*) a 呼噜 (*hūlū, zvuk chrápání*).

- Sémanticky motivované kolokace: v těchto slovních spojeních lze z doslovného významu odvodit přenesený význam. Např. 发表意见 (*vyjádřit názor*) nebo 摆脱贫困 (*vymanit se z bídy*).

Dále Laufer poskytuje vlastní výčet problematických kategorií lexémů (Laufer 1989, s. 10–13):

- Idiomy: význam těchto slovních spojení není možné odvodit z významu jejich komponentů. Studenti idiomy nemusejí rozpoznat nebo se domnívat, že jejich význam mohou odvodit z jednotlivých slov. Příkladem neprůhledných čínských idiomatických výrazů jsou 成语 (*ustálené fráze, jejichž význam se vztahuje k všeobecně známému příběhu*). V seznamech zkoumaných úrovní HSK nebyly zařazeny žádné idiomy.
- Falešní přátelé: jde o slova nebo slovní spojení, která studentům cílového jazyka znějí stejně či podobně jako v mateřském jazyce. Studenti se pak mohou domnívat, že je význam slov v obou jazycích totožný. Vzhledem k velkým odlišnostem mezi češtinou a čínštinou nejsou falešní přátelé pro studenty problémem. V němčině může podobnost ve výslovnosti s sebou přinést komplikace. U nejčastějších společenských obrátů se např. 拜托 (*bàituō, prosím*) podobá německému „Bitte!“. Spojení si frází „Tschüss!“ a 去死 (*qùsǐ, chcípni*) podle stejné logiky však není dobrý nápad.
- Mnohovýznamová slova: pokud není ve slovníčku vyznačeno, že má lexém více významů nebo tyto významy nejsou uvedeny, snadno to studenty zmate. Mohou získat mylný dojem, že již všechny významy ovládli. K chybám může vést i nejasné vyznačení slovních druhů.
- Formální podobnost: podobná slova jsou snadno zaměnitelná, a tak je jejich ovládnutí zvláště složité. V čínštině se tento problém týká jak výslovnosti a tónové podobnosti, tak grafické podoby znaku.

V experimentu týkajícím se porozumění izraelských studentů netransparentním anglickým slovům Laufer zjistil, že výše uvedené kategorie slov studenti v porovnání s ostatními slovy častěji chybně interpretují a neumějí tato slova správně použít (Laufer 1989, s. 13). Proto by tyto kategorie měly být ve slovníčku zřetelně vyznačeny.

V případě, že slova nebo slovní spojení nejsou důležitou součástí slovní zásoby, měla by být v soupisech pro začátečníky upřednostňována raději méně komplikovaná slova.

4.7.2 Využití souvislostí mezi lexikálními jednotkami

Paradigmatické vztahy mezi slovy jsou jedním z důležitých způsobů asociování lexémů v mentálním lexikonu. Podle Chang Jingyua (Chang 2003, s. 14) je třeba při akvizici čínského jazyka využít této systematickosti. To může ulehčit ovládnutí nových slovíček z hlediska jejich významu, ale také díky jejich grafické podobě (asociací společných morfémů) (Chang 2003, s. 18). Když si studenti čínštiny uvědomí souvislosti mezi slovíčky, pomůže to jejich schopnosti ukládat lexémy do paměti a úspěšně je zase vyvolat. Chang poskytuje výčet kategorií, které díky asociativním vztahům usnadní akvizici slovní zásoby (Chang 2003, s. 16–18):

- Synonymie: využití úplně a neúplně souznačných slov k vytvoření sítě sémanticky podobných slov. V čínštině může jít často o slova se stejnou bází. Spojení si těchto slov může ulehčit studentům pochopení jejich odlišností, neboť ačkoliv mají lexémy stejný význam, mohou mít jiná pravidla použití, např. 安静 (*tichý, klidný* – popisuje hlavně atmosféru nebo prostředí) 宁静 (*tichý, klidný* – popisuje hlavně způsob života).
- Antonymie: využití souvislostí významově protikladných slov k vytvoření asociace, např. 黑 – 白 (*bílý a černý*).
- Hyperonymie a hyponymie: vztah mezi nadřazenými a podřazenými pojmy. Spojování a osvojování si slovíček ve skupinách podle jejich hierarchických vztahů. Např. rodina: členové rodiny, nebo ovoce: druhy ovoce.
- Heterosémie: vztah slov, která patří do jedné sémantické kategorie. Může se jednat o asociování dvou či více lexémů, např. 左 – 中 – 右 (*levý, střed, pravý*) nebo 春 – 夏 – 秋 – 冬 (*jaro, léto, podzim, zima*) apod.

Tvůrci slovníčku by měli dbát na koherentní uspořádání podle logicky odvoditelných vztahů, z pohledu studentů je vhodné využít i volné asociace k lepšímu zapamatování si lexémů.

Tato metoda ulehčí osvojování si slovní zásoby zejména studentům začátečnické úrovně. Studenti cizího jazyka mají zpočátku tendenci tvořit asociace na základě fonetické podobnosti, poté paradigmatické asociace a teprve jako samostatní uživatelé jazyka postupně věnují větší pozornost syntagmatickým souvislostem (Krames 2015, s. 10).

4.7.3 Lexikální přístup

Lexikální přístup (*Lexical approach*) je další významná metoda nahlízející na akvizici slovní zásoby novým způsobem. Její počátky sahají do osmdesátých let 20. století (Sun 2011, s. 14). Jedním z hlavních představitelů lexikálního přístupu je Michael Lewis, který vydal stejnojmennou knihu *Lexical approach* (Ramirez 2012, s. 239). V Číně se lexikálním přístupem (句本位、词组本位) zabývali např. Zhu Dexi (朱德熙), Lu Shuxiang (吕叔湘) či Shi Youwei (史有为) (Sun 2011, s. 14). Tato metoda výuky jazyka se soustředí mnohem více na komunikační kompetenci než na správné použití gramatiky. S gramatikou se student seznámí postupným pozorováním, jak spolu souvisejí lexikální celky, které již zná. Tento přístup se opírá o hypotézu, že se studenti jednodušeji naučí lexikální fráze než gramatická pravidla (Lewis 2005, s. 9).

Podle Lewise se velká část naší mluvené a psané komunikace skládá nikoliv ze slov, ale z několikaslovných celků (v angličtině zvaných *chunks*). Spolu užívaná slova by proto měla být osvojena jako celek rovnou (Lewis 2005, s. 8). Na morfémy a sémémy není brán žádný zřetel, celé lexikální fráze mohou být používány bez nutného porozumění jednotlivým slovům (Ramirez 2012, s. 239). Používání předem naučených frází může zlepšit plynulost projevu, protože když se mluvčí nemusí soustředit na skládání jednotlivých slov, má potom čas si lépe zorganizovat promluvu (Ramirez 2012, s. 244).

Aplikace lexikálního přístupu při čtení textů v čínštině nemusí být pro studenty na nižších úrovních přínosná. Marion Krames (2015, s. 14) ve své studii zmiňuje výzkum, v němž vyšlo najevo, že úspěšné zpracování frází při čtení textů je patrné jen u čínských rodilých mluvčích, kteří jsou schopni vnímat najednou několikaslovné fráze. Studenti čínštiny jako druhého jazyka se koncentrují na jednotlivé znaky, ze kterých následně identifikují známá slova, a až potom smysl frází. Rodilí mluvčí tak dokázali číst dvakrát rychleji než studenti čínštiny na pokročilé úrovni. Tento přístup proto není

efektivní u studentů na začátečnické a středně pokročilé úrovni. Těm nejsou lexémy dostatečně známé nebo nedokážou najít záchytné body, aby několikaslovné fráze okamžitě identifikovali (Sun 2008, s. 173).

4.7.4 Role morfémů při akvizici čínštiny

Opačnou metodou ke zlepšení komunikační kompetence, než kterou představuje lexikální přístup, je pojetí didaktiky čínštiny jako cizího jazyka, jež se soustředí v první řadě na objasnění významu a funkce čínských bází (čínsky 字本位理论, *Teorie vycházející s ústředního postavení znaků*). Byla v devadesátých letech 20. stol. definována čínským lingvistou Xu Tongqiangem (徐通锵) (Li 2004, s. 75).

Podle této teorie je při akvizici nových lexémů i při studiu čínské gramatiky třeba vycházet z porozumění bázím (Kang 2011, s. 159). Studenti jsou seznámeni s pravidly čínské gramatologie, které jim umožní získat informace ze struktury samotné báze (Xin 2013, s. 252). Báze jsou nejmenšími jednotami v jazyce, které jsou nositeli významu. Studenti se nejprve naučí význam jednotlivých bází a pravidla, podle kterých se v čínštině skládají báze do slov. Díky tomu jsou potom schopni lépe posoudit konotace známých lexémů. U neznámých slovíček mají větší šanci správně identifikovat význam a rozsah slov (Li a Yang 2004, s. 29).

Na rozdíl od ostatních teorií, inspirovaných západními lingvistickými směry, je teorie vycházející z ústředního postavení znaků založena pouze na čínské jazykovědné tradici (Li 2004, s. 75). Podle proponentů této teorie je její aplikací v praxi možné ušetřit čas strávený nad výukou slovní zásoby, usnadnit studentům akvizici dalších slovíček a zlepšit porozumění čtenému textu (Xin 2013, s. 252).

5 PRAKTICKÁ ČÁST

5.1 Obecná charakteristika seznamu slovní zásoby HSK 甲级

Seznam slovní zásoby HSK 甲级 obsahuje celkem 1 033 lexémů. Ke každé lexikální jednotce jsou připojeny informace o výslovnosti v transkripci pinyin a o gramatické funkční charakteristice daného slova. Dále ke slovní zásobě této úrovně patří soupis 800 znaků, z nichž jsou lexémy slovní zásoby složeny, a také popis 129 gramatických struktur souvisejících se seznamem.

Při zpracování vlastního korpusu slovní zásoby HSK 甲级 bylo nutné provést úpravy, aby byl porovnatelný s daty získanými z ostatních jazykových korpusů a s pozdějšími verzemi seznamů z let 2010 a 2012. Úprava spočívá ve sloučení funkčních charakteristik vázaných ke slovům. V soupisu HSK 甲级 jsou slovní druhy jednoho lexému rozděleny a definovány samostatně jako několik lexikálních jednotek. Vycházejí tedy ze zásady, že kolik má lexikální jednotka funkcí v seznamu stanovených, o tolik slov se jedná. V pozdějších soupisech toto rozdělení chybí nebo je provedeno značně nekonzistentně. Pozdější seznamy pravděpodobně předpokládají ovládnutí všech slovních druhů daného slova, pokud není uvedeno jinak. Proto bylo třeba omezit počet slov a vázat lexém na všechny gramatické funkční charakteristiky, ve kterých může vystupovat.

Tato úprava byla nutná i vzhledem k volbě korpusů. SUBTLEX-CH, LCMC a BCC(微信) používají rozdílný systém značení slovních druhů i větných členů a u korpusu TORCH2009 toto značení zcela chybí. Všechny hodnoty v mnou zpracovaném korpusu jsou uváděny pro veškeré funkční vlastnosti slova dohromady.

V seznamu HSK 甲级 jsou u některých podstatných jmen uvedeny společně rovněž varianty slova [např. 星期日 {礼拜天} (*neděle*)]. Tyto lexémy jsem rozdělila a data vyhledala pro každý jednotlivě. Po uvedených úpravách byl soupis HSK 甲级 zkrácen na 1 010 lexikálních jednotek.

Nejdůležitější vlastností seznamu slovní zásoby je jeho praktická využitelnost. Ve svém výzkumu se soustředím zejména na analýzu frekvence. Podle lingvistických studií jsou vysoce frekventovaná slova snadněji zapamatovatelná, dříve osvojovaná a

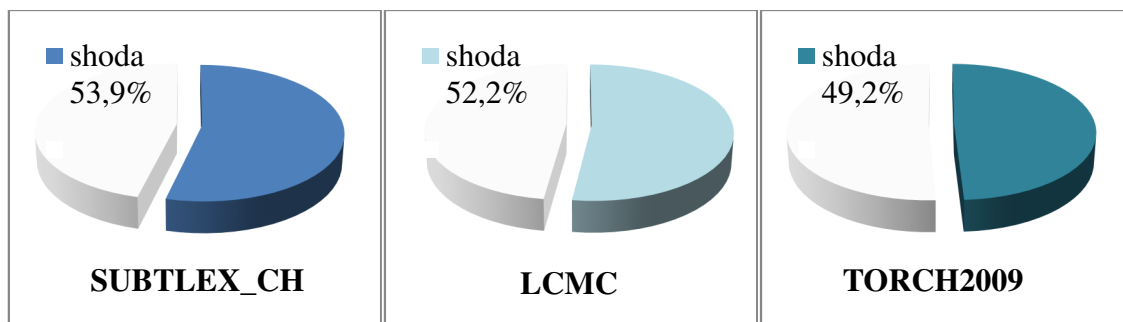
méně podléhají změnám (Bentz et al. 2016, s. 1). S častějším výskytem slov stoupá potřeba jim porozumět, což je zvláště důležité u obecných seznamů určených primárně pro dospělé studenty, jako jsou soupisy HSK. Vliv frekvence by se měl proto projevit při tvorbě a při úpravách soupisů slovíček. Zkoumám frekvenci výskytu lexikálních jednotek v hovorové čínštině (korpus SUBTLEX-CH), v textech různých žánrů publikovaných kolem roků 1991 (LCMC) a 2008 (TORCH) a výskytu lexémů v mikroblozích (BCC 微信), tj. v internetových textech nepřesahujících 140 znaků (délku jedné SMS).

Pro posouzení, který seznam slovní zásoby je pro studenty nejvhodnější, je třeba přihlídnout i k získaným komunikačním dovednostem, jež jeho tvůrci deklarují. Pro soupis HSK 甲级 jsou komunikační dovednosti po osvojení si 1 033 slov definovány velmi podrobně. Je uváděna předpokládaná rychlost čtení, psaní i výčet témat, kterým by student měl v čínštině rozumět (Duiwai hanyu jiaoxue xiaozu¹ 1995, s. 85–86). Ovládnutí slovíček a s nimi související gramatiky umožní studentovi: „... nejzákladnější dorozumění se v rámci každodenního života, vzdělávacích programů a jednoduchou společenskou komunikaci.“ (Duiwai hanyu jiaoxue xiaozu 1995, s. 85) Dále má být student schopen číst krátké a jednoduché články, nekomplikovaná sdělení, tabulky a běžné rozvrhy. Pro psaní a porozumění platí stejně nízké požadavky (Duiwai hanyu jiaoxue xiaozu 1995, s. 86).

5.1.1 Kvantitativní posouzení seznamu HSK 甲级

Pro můj výzkum je relevantní nejen komparace získaných hodnot mezi jednotlivými soupisy, ale také četnost lexémů HSK 甲级 v nejfrekventovanějších slovech vybraných korpusů. Vyhledala jsem 1 000 nejběžnějších slov v korpusech SUBTLEX-CH, LCMC a TORCH2009 a porovnávala shodu se seznamem staré slovní zásoby HSK.

¹ Celým názvem 国家对外汉语教学领导小组办公室汉语水平考试部 *Guojia duiwai hanyu jiaoxue lingdao xiaozu bangongshi Hanyu Shuiping Kaoshi bu* (Kancelář Národní vedoucí skupiny výuky čínštiny jako cizího jazyka – oddělení Zkoušky z čínštiny jako cizího jazyka)



Graf č. 1: Shoda soupisu HSK 甲级 s 1 000 nejfrekventovanějšími lexémy v korpusech

Seznam HSK 甲级 vykazuje zhruba 50% shodu ve všech třech korpusech, nejvíce v korpusu filmových titulků. V případě 1 000 nejfrekventovanějších slov v korpusech je zastoupení korpusu textů vydaných kolem roku 2009 nejnižší, což platí i v průměru pro frekvenci všech slov v seznamu. Důvodem nízkého zastoupení nejfrekventovanějších slov korpusů TORCH a LCMC je množství lexémů typických pro psaný projev, které ve slovní zásobě pro začátečníky pochopitelně chybějí. V případě korpusu SUBTLEX-CH celkové zastoupení naopak snižují slova typická pro hovorovou čínštinu, např. 500 nejfrekventovanějších slov v korpusu zahrnuje 15 citoslovců, v LCMC 5 a v TORCH jen 2 citoslovce. Většina lexémů HSK 甲级 jsou bezpříznaková slova, preference hovorového či psaného lexika se proto mezi stylově zabarvenými nejběžnějšími slovy jednotlivých korpusů příliš neprojevila.

U korpusu BCC(微信) není možné získat data pro vypočítání frekvence všech lexémů, a proto v grafu chybějí. Pro jazyk používaný v mikroblozích je typické užití hovorového jazyka, slangu, cizích slov a akronymů (Liu 2014, s. 217). V níže uvedeném textu je zvláště patrná také přemíra interpunkčních znamének. Abych více využila tento korpus, použila jsem možnost přednastaveného hledání v korpusu a z každého desátého výsledku z příspěvků poskládala text o 488 znacích, ve kterém za pomoci parsovacího programu Pangu hledám lexémy ze seznamu HSK 甲级. Modře vyznačené lexémy jsou součástí seznamu staré verze HSK a pro čtenáře srozumitelné.

Tabulka č. 4: Slovíčka ze soupisu HSK 甲级 v mikroblogových příspěvcích

下午一下收到两个快递，辛苦了帮忙去取的同事!!! 还有段段这6个杯具!!!。。。之前只听说日本这样，国内也开始了?!。。。。

声音还原度好高啊。。。真服了，这几个家伙配音也太像了"居然是站票",这句话太销魂了真"原音",推荐收看! 怎么做到配得如此"原音"滴? 有没有人有防水卡片机啊借我用一下我可以先用尼康 S8000 交换啊。。。其实我觉得好欢乐啊。。。好可怜, 好愤怒。

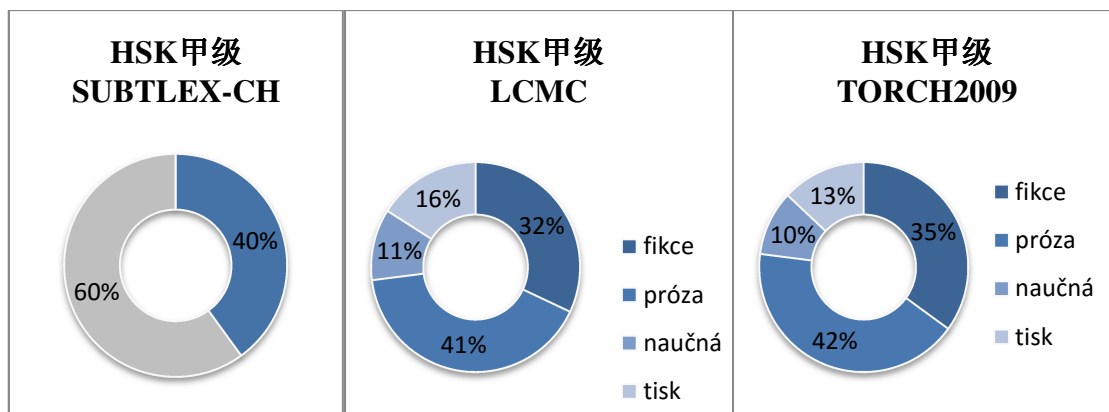
好有爱的卡布奇诺.....终于车先生下班了终于开始吃了, yo 宝已经酒足饭饱, 坐一边玩去了我的那个天.....yo 宝这个鸡翅狂人!!!! 一口气把三只鸡翅全部消灭光! 我说要留给爸爸, 她说: 爸爸来了, 再叫一份! yo 宝说: 干杯妈妈! 生日快乐! 20 这段人生的最后一个数字了, 明年就是另外一段十年人生路了。一场轰轰烈烈的爱情, 一次说走就走的旅行。

我发现我可真是赚到了, 才活了不过 20 几年, 碰上了 50 年一遇的旱灾, 100 年一遇的洪水, 60 年一遇的大雪, 300 年一遇的泥石流, 500 年一遇的地震, 600 年一遇的海啸.....人家轮回转世好几趟都碰不全呀!!

丫就不会反省一下吗?!! 你比那些他妈的好多了好不好!! 不想写了, 赶紧赶紧的下课吧。。。

Výše uvedený text obsahuje celkem 169 jedinečných slov, z nichž je 89 součástí slovníčku HSK 甲级 a dalších 6 je s jeho znalostí snadno odvoditelných. Zhruba 56 % textu je tedy pro čtenáře se znalostmi slovníčku srozumitelných, to však k porozumění smyslu textu nestačí.

Pokud jde o distribuci lexémů v jednotlivých žánrech, lze tato data sledovat v korpusech LCMC a TORCH2009. Oba korpusy jsou složeny z žánrově definovaných subkorpuseů, které je možné prohledávat zvlášť nebo jako celek. Rozlišovanými žánry jsou: fikce, próza, naučná literatura a tisk. V korpusu SUBTLEX je možné zjistit, v kolika procentech filmů ze sbírky se hledané slovo objevilo. Lexémy daného seznamu se objevují ve 40 % filmů z databáze korpusu.

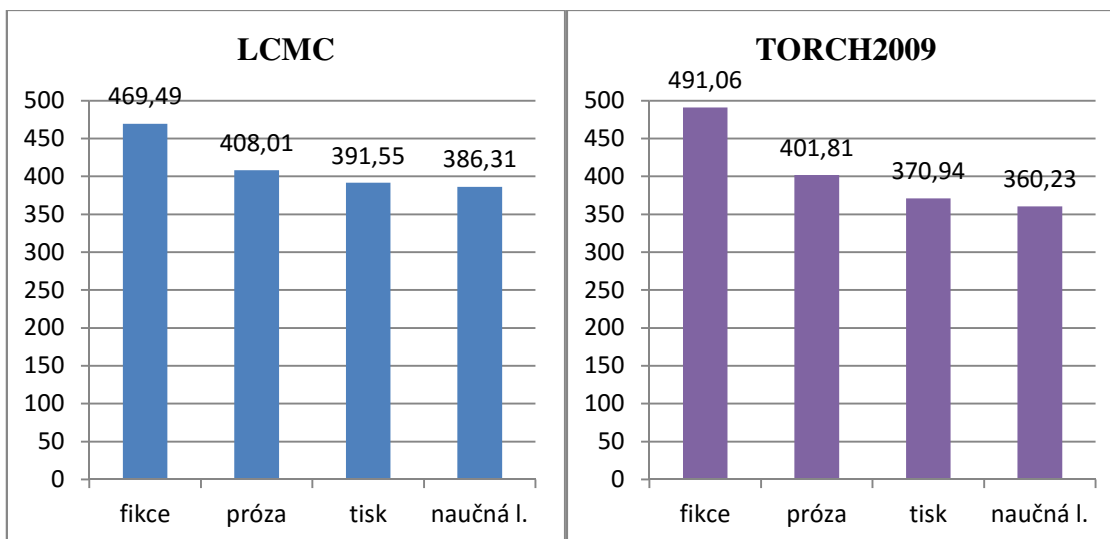


Graf č. 2: Výskyt lexémů soupisu HSK 甲級 ve filmové databázi a textových žánrech

Tematicky se slova ze soupisu nejčastěji objevovala v próze. Tato kategorie v korpusech zahrnuje texty obecného charakteru, které nespádají do zbylých tří kategorií. V této kategorii lze očekávat vyšší výskyt bezpříznakových slov, která nejsou vázána na určitou stylovou rovinu psaného jazyka. Druhá nejvíce zastoupená kategorie je fikce, kam jsou zařazeny žánry jako humorné texty, romantická literatura, dobrodružné novely, detektivky, sci-fi a obecná literární fikce, pro niž je typický umělecký styl. V této kategorii lze předpokládat vyšší zastoupení neformálního jazyka a prvky hovorové řeči. Poměrně menší část slov se objevuje v žánru tisk, tedy v publicistickém stylu. Nejmenší zastoupení má žánr naučné literatury, kde jsou běžné odborné termíny a formální úroveň jazyka.

Uvedené poměry odpovídají omezeným možnostem studenta čínštiny na deklarované úrovni. Při znalosti základní slovní zásoby staré verze HSK lze alespoň zčásti porozumět jen jednoduchým textům obecného charakteru.

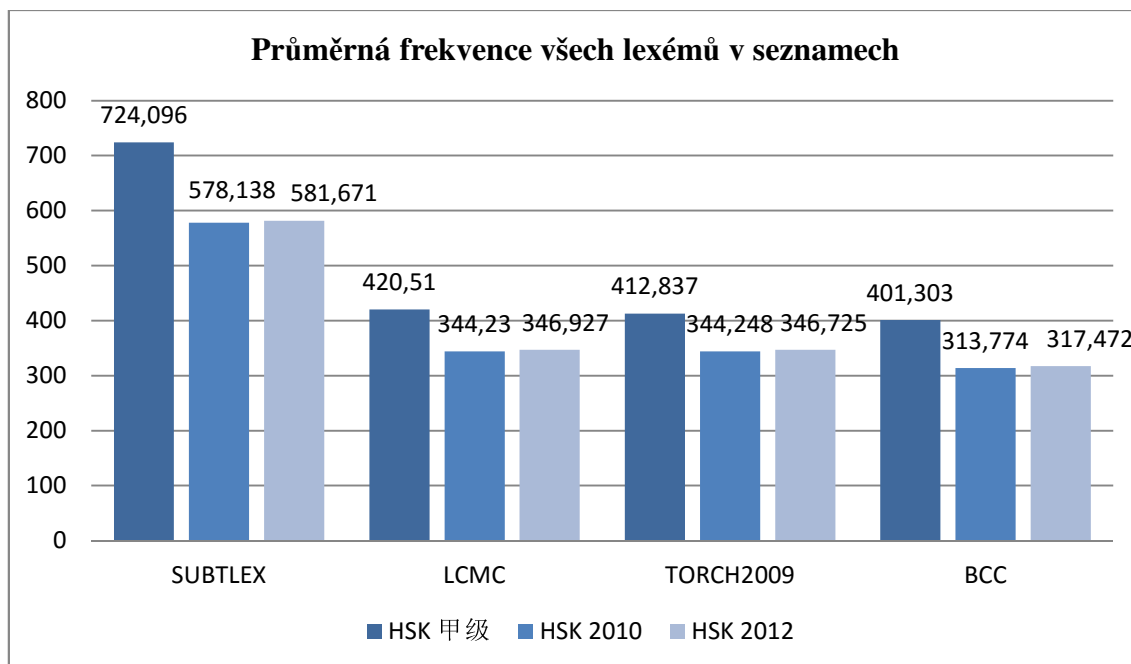
Předchozí graf uvádí distribuci slov podle žánrů bez údajů o frekvenci lexémů. Pokud vezmeme v úvahu frekvenční hodnoty v rámci žánrů (výskyt slova na 1 milion lexémů), budou data rozdělena následujícím způsobem.



Graf č. 3: Průměrná frekvence seznamu HSK 甲级 v textových žánrech

Slova ze seznamu s průměrně nejvyšší frekvencí se objevují v žánru fikce. Ačkoliv lexémy ze seznamu staré verze HSK jako celku se v průměru objevují nejvíce v obecných prozaických textech, nejběžnější slova z tohoto soupisu jsou shromážděna v beletrii, která je bližší hovorové čínštině než formálnější prozaické žánry, jako např. esej či úvaha. Žánry literární fikce a obecné prózy jsou jazykově jednodušší než specializovanější publicistický a odborný styl, které mají opět nejnižší frekvenční hodnoty. Seznam HSK 甲级 byl navržen k získání základních komunikačních dovedností. Kromě vyšších hodnot v žánru beletrie nevykazují hodnoty dalších žánrů velké rozdíly v průměrné frekvenci výskytu. V tomto ohledu svědčí rozložení dat v grafu č. 3 o tom, že slovíčka v soupisu jsou široce uplatnitelná a mají opravdu obecný charakter.

Průměrné údaje o frekvenci lexémů ze všech korpusů bez ohledu na žánr jsou zobrazeny na grafu č. 4. Vyhledané hodnoty pro všechny lexikální jednotky soupisů HSK v korpusech jsem normalizovala na frekvenci výskytu na jeden milion lexémů. Po zprůměrování výsledků mohu určit a porovnat, který seznam je v korpusech nejvíce zastoupen a také rozdíly v korpusech navzájem.



Graf č. 4: Průměrná frekvence všech lexikálních jednotek v seznamech

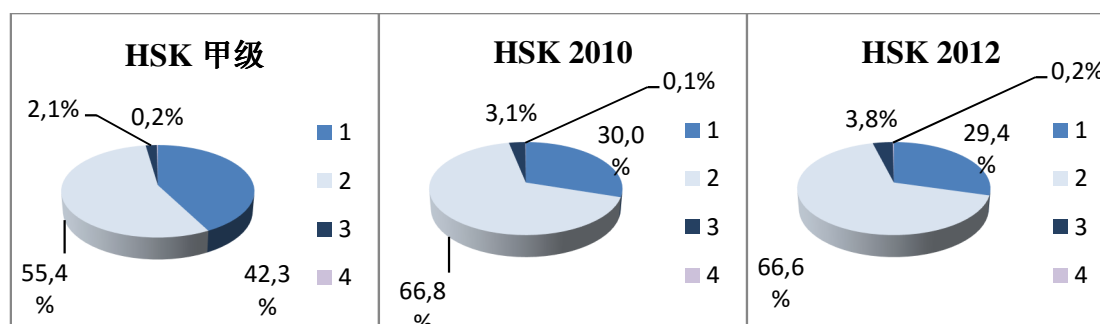
Pro všechny tři soupisy platí, že frekvence lexémů v korpusu SUBTLEX-CH jsou více zastoupeny než v ostatních korpusech. Důvodem jsou výrazně vyšší hodnoty u nejfrekventovanějších slov, zvláště osobních zájmen (např. zájmeno 我 má 9x vyšší frekvenci než v korpusu LCMC a 3x vyšší než BCC, zájmeno 你 dokonce 13x vyšší frekvenci než v korpusu LCMC a 4,5x vyšší než korpus BCC).

Velký rozdíl ve vysoce frekventovaných slovech silně ovlivnil průměrnou frekvenci mezi hodnotami v seznamu. Není pravděpodobné, že by nejpoužívanější slova v čínštině byla tolikrát frekventovanější v hovorové čínštině, navíc by se podobně velký rozdíl promítl i do korpusu mikrobloginých příspěvků. Patrné rozdíly mezi korpusem hovorového jazyka BCC(微信) a korpusy psaného jazyka LCMC s TORCH2009 sice existují, ale ne s rozdíly v řádech desetinásobku u nejběžnějších slov.

V těchto případech se zřejmě jedná o zkreslení dat. Bohužel alternativní korpus, který by dokumentoval mluvený projev, není přístupný. Ačkoliv data korpusu SUBTLEX-CH nejsou pro komparaci s ostatními korpusy spolehlivá, lze je využít k porovnávání mezi jednotlivými seznamy, neboť se problém nápadně vyšší frekvence týká jen malého počtu lexémů, které jsou součástí všech tří soupisů.

Z grafu č. 4 je patrné, že starší soupis HSK má ve všech korpusech vyšší frekvenční hodnoty. Výrazně vyšší frekvence u seznamu je způsobena vyšším

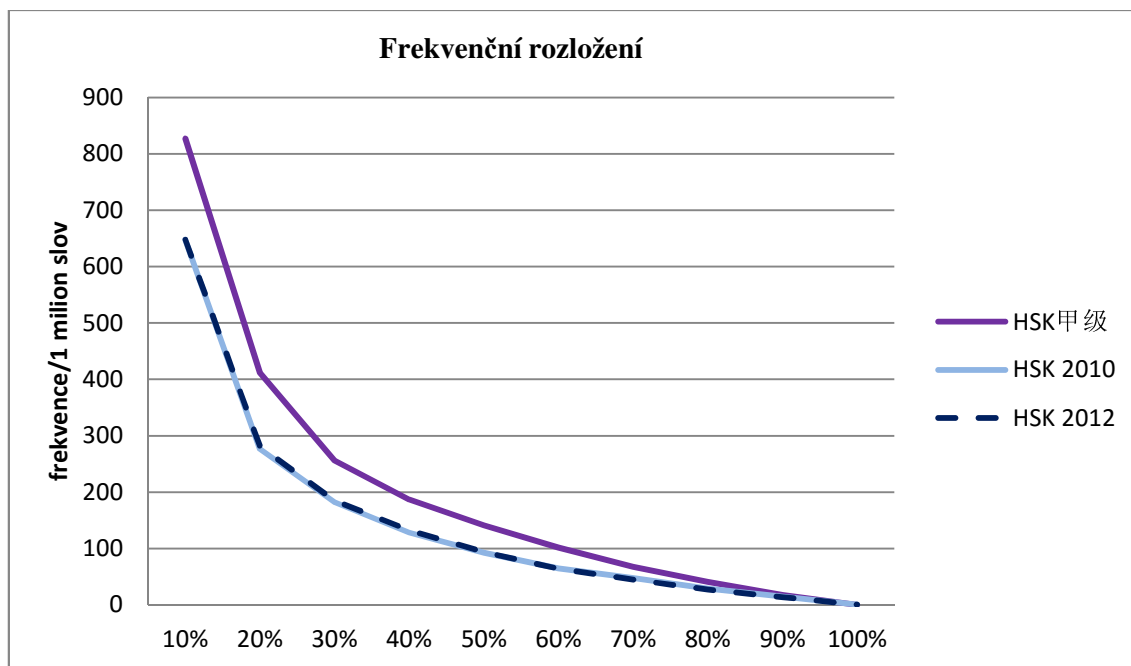
zastoupením jednoslabičných slov. Když shromáždíme lexémy všech tří seznamů a roztrídíme je podle délky, mají jednoslabičné sémémy průměrnou frekvenci 2 420 výskytů na milion slov, dvouslabičná slova 134, tříslabičná 44 a čtyřslabičná slova jen necelé 4 výskyty. To se promítlo do hodnot všech tří seznamů, v seznamu staré verze HSK nejvíce, protože je zde mnohem vyšší procentní zastoupení jednoslabičných bází. V grafu č. 5 je uvedeno rozdělení slov v soupisech podle jejich délky.



Graf č. 5: Složení seznamů slovní zásoby podle délky slov

Ve všech třech soupisech převládají dvouslabičná slova, stejně jako v moderní hovorové čínštině. Zásadní rozdíl je pouze v poměru jednoslabičných slov mezi starým soupisem HSK a dvěma pozdějšími seznamy. HSK 甲級 klade obecně větší důraz na jednotlivé morfémy, které tvoří i samostatný seznam doplňující zde zkoumaný soupis lexémů. V pozdějších verzích soupisu jsou preferována spíše kompozita, na možnosti jednoslabičných bází není brán zřetel.

Dále jsou zkoumány poměry v průměrné frekvenci všech tří soupisů. Graf č. 6 zobrazuje průměrné hodnoty u x procent slov s nejvyšší frekvencí výskytu. Rozdíl ve frekvenci je pouze u seznamu před reformou zkoušky a po reformě, úprava seznamu v roce 2010 se z hlediska průměrné frekvence lexémů téměř neprojevila.



Graf č. 6: Frekvenční rozložení lexémů v seznamech slovní zásoby

Starší soupis HSK má až do 90 % nejfrekventovanějších slov v soupisu opět vyšší průměrnou frekvenci než pozdější dva seznamy. Data reformovaných seznamů se kryjí, ať už jde o první desetinu nejběžnějších slov v seznamech nebo o nejméně frekventované lexémy.

5.1.2 Kvalitativní posouzení seznamu HSK 甲级

Starší verzi soupisu slovíček HSK byla vytýkána zejména možná zastaralost. Seznam byl navržen koncem 80. let pro potřeby studentů v 90. letech. Podle kritiků již nerefletoval dnešní jazykové potřeby (Han a Cheng 2012, s. 34). Základní slovní zásobu by tento problém příliš ovlivnit neměl, přesto jsou v seznamu některá slova podstatně aktuálnější v době, kdy seznam vznikal, která pro studenta na dané úrovni nejsou užitečná, např. 农民 (*rolník*), 同志 (*soudruh*) nebo 干部 (*kádr*). Problém zastaralosti se však týká jen minima slov, jako např. 磁带 (*magnetická páska*). Seznam spíše obsahuje příliš mnoho synonym, jejichž znalost je pro začátečníky zbytečná, např. 月亮/月球 (*měsíc*) nebo 早上/早晨 (*ráno*) aj. Většina slov je však stále běžně používaná a aktuální, což potvrdila i vysoká frekvence.

Dalším bodem kritiky starého seznamu je jeho přílišná náročnost pro studenty-začátečníky. Osvojení si minimální slovní zásoby v rozsahu 1 038 lexémů je pro studenty příliš složité a demotivující, a proto bylo doporučeno zredukování slovníčku (Zhang a Zhang 2010, s. 37). Tento problém byl zohledněn v reformě slovní zásoby z roku 2010, ukázalo se však, že omezení rozsahu slovníčku vedlo zároveň k výraznému snížení jeho užitečnosti a dalším problémům, kterým se věnuji v teoretické části práce. I po ovládnutí všech lexémů staré verze seznamu HSK mají studenti značně omezené možnosti vyjadřování, na což poukazují i tvůrci původního soupisu (Duiwai hanyu jiaoxue xiaozu 1995, s. 86).

Naprostá většina lexémů v slovníčku HSK 甲级 jsou neutrální výrazy bez citového a stylového zabarvení. Jde o slova použitelná jak v hovorové čínštině, tak v neformálním psaném stylu. Všeobecná uplatnitelnost je pro studenty-začátečníky důležitou vlastností slovní zásoby.

Autosémantika v soupisu mají obecný charakter, často jde o hyponyma. Pro sysémantika v seznamu také platí, že nejsou výlučně vázána na psaný či mluvený projev. Část sysémantik je uvedena ve formě způsobu užití, např. 除了…以外 (*kromě*), 连…都(也)… (*dokonce i*). Pokud některou vazbu nebyl korpus schopen vyhledat, byla rozdělena na jednotlivé báze. Pokud tyto báze korpus již obsahoval, byla vazba z výpočtů frekvence vyloučena (celkem 5 gramatických struktur). Tato slovíčka vyhledávání v korpusu sice komplikují, ale pro studenty je objasnění jejich souvislostí a užití velmi praktické. V pozdějším seznamu bylo od tohoto způsobu uvádění slovíček upuštěno ve prospěch uvádění sysémantik izolovaně.

Doporučená slovní zásoba staré verze HSK obsahuje základní lexémy potřebné pro komunikaci a porozumění v běžných každodenních situacích. Ve specializované slovní zásobě lze pozorovat vyšší výskyt slov úzce zaměřených na školu a studium jazyka, např. 课文 (*text*), 课本 (*učebnice*), 听写 (*diktát*), 预习 (*opakovat si dopředu*), 语法 (*gramatika*), 生词 (*slovíčka*) aj., která v pozdějších soupisech chybějí. Důvod odstranění těchto slovíček může spočívat v jejich velmi nízké frekvenci užití v běžném hovorovém jazyce. Zároveň je ale studium čínštiny a motivace složit jazykovou zkoušku jednou z mála věcí, která studenty těchto seznamů spojuje, a v těchto případech jsou tato slovíčka velmi užitečná.

Již zmiňovaný vysoký podíl slov tvořených jednou bází lze vnímat z hlediska studentů jako výhodu i jako nevýhodu. Jednoslabičná slova pocházejí ze starších vrstev jazyka, a tak mívají komplikovanější denotace a konotace než víceslabičná slova. Na druhou stranu tato mnohovýznamová slova patří do základní slovní zásoby čínského jazyka a ani pozdější seznamy, upřednostňující kompozita s více morfémy, se jim nemohou zcela vyhnout. Výhodou studia jednoslabičných slov je jejich vyšší užitný potenciál, neboť jejich znalost usnadňuje akvizici dalších složených slov. V každém případě jsou lexikální jednotky tvořené jednoduchými bázemi vysoce frekventované, a tak mají v seznamu slovíček jistě své místo.

5.1.3 Změny seznamu HSK 甲级 po reformě

Po reformě v roce 2010 byl sice seznam na úrovni skutečně odpovídající bývalému HSK 甲级 rozšířen na 1 200 slovíček, ale značná část lexémů ze staršího seznamu byla vyřazena, konkrétně 343 slovíček, tj. 34 % z celého původního slovníčku základní úrovně.

Vyřazena byla příliš podobná slova jako 北 (*sever*) a 北边 (*severní, na severu*), 常 a 常常 (*často*), 秋, 秋天 (*podzim*) apod., směrově modifikovaná slovesa, např. 出来, 出去 (*vyjít směrem k mluvčímu a od mluvčího*) a 进来, 进去 (*vejít směrem k mluvčímu a od mluvčího*). Zrušeny byly také varianty slov, které byly u lexémů uváděny v závorkách, např. 法语 (发文) (*francouzština*) aj. Počet jednoslabičných bází byl zredukován ze 42 % na 30 % a zrušena byla také většina informací o gramatických funkčních charakteristikách lexikálních jednotek.

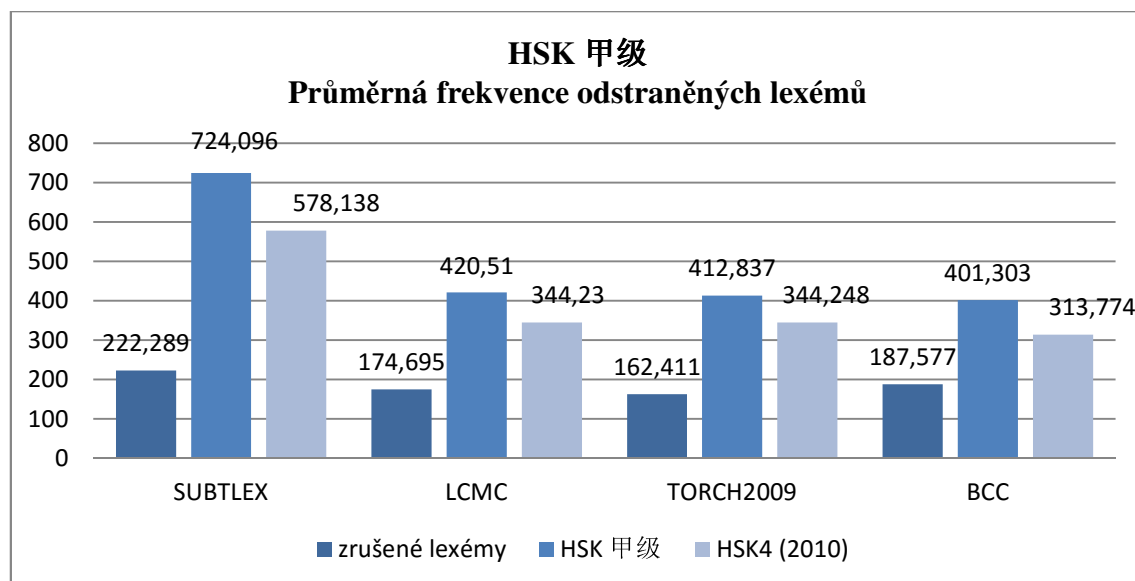
Z vysoce frekventovaných lexémů byly zrušeny vázané morfémy, jako např. ukazatel plurality jmenných výrazů 们 nebo záložky 前 (*před*), 里 (*uvnitř*), 后 (*za*). Volné morfémy, které jsou v reformovaném seznamu součástí kompozita, např. 车 (mnohovýznamové slovo, hlavní význam *vůz*) se vyskytují jen ve slovech jako 自行车 (*kolo*), 出租车 (*taxi*) apod. a 事 (hlavní význam *záležitost*) je jen součástí slov jako 事情 (*záležitost*), 故事 (*příběh*). Lze se domnívat, že vnímaví studenti mohou pochopit z textu význam slov 车 a 事 odvozením ze složených slov, která již znají. Opačný postup analýzy slov z významu jednotlivých bází je však pro studium čínštiny

praktičtější, což si uvědomovali tvůrci starého seznamu, když za cenu velkého rozsahu výsledného soupisu uváděli jako slovíčka samostatně produktivní báze a navíc i kompozita z nich složená.

Z vysoce frekventovaných bází byly zrušeny např. atributivy 男 (*mužský*), 女 (*ženský*), zcela základní adjektiva jako 早 (*časný, dříve, dávno...*) a 晚 (*pozdní, později...*). Také velmi frekventovaná měrová jména 天 (*den*) a 夜 (*noc*) nejsou samostatně uvedena. Chybějí běžné numerativy 支 (numerativ pro hůlkovité předměty) a 片 (numerativ pro ploché předměty).

Vynechána byla důležitá jednoslabičná slova každodenní potřeby, jako jsou např. 床 (*postel*), 肉 (*maso*), mezi kompozity pak 早饭 (*snídaně*), chybějí i *oběd* a *večeře*, dále 外国 (*zahraniční*), 大学 (*vysoká škola*) atd. Dokonce i velmi aktuální lexémy, které starší seznam HSK obsahoval, v novém seznamu chybějí, např. 汽车 (*auto*), 机器 (*stroj*), 拍 (*natáčet video, fotografovat*) nebo 照相 (*fotografovat*). Úkol přizpůsobit nový seznam slovní zásoby podmínkám moderního života tedy nebyl proveden důsledně.

Z pohledu průměrných hodnot byla odstraněna spíše středně a málo frekventovaná slova, jak je uvedeno na grafu č. 7, kde je jejich průměrná frekvence porovnávána s původním a reformovaným soupisem.



Graf č. 7: Průměrná frekvence lexémů odstraněných ze seznamu HSK 甲级

Provedení reformy seznamu slovní zásoby budí dojem, jako by tvůrci nevzali starší verzi soupisu vůbec v potaz, ale sestavili seznam zcela nový. Nedá se říci, že by lexémy vyřazené z nového seznamu byla výlučně slova s nízkou frekvencí výskytu nebo slova, která nemají obecný či aktuální charakter. Tvůrci nového seznamu výrazně omezili jednoslabičné báze ve prospěch kompozit, mnoho lexémů s velmi nízkou frekvencí a synonymní výrazy, kromě toho však jiné systematické změny vyzorovat nelze.

5.2 Obecná charakteristika seznamu slovní zásoby 新 HSK4 z roku 2010

Doporučená slovní zásoba nové formy zkoušky HSK4 obsahuje celkem 1 200 lexikálních jednotek. Tento seznam lze najít v oficiálních sylabech HSK čtvrté úrovně vydaných do roku 2012 a také bylo do téhož roku možné jej stáhnout z oficiálních webových stránek Hanbanu. Snadný přístup k slovníčkům a oficiálním studijním materiálům je výhodou nových HSK, starší oficiální verze soupisu takto dostupná nebyla. Slovíčka jsou abecedně řazena podle výslovnosti v pinyin. Až na výjimky nejsou ke slovům přiřazeny informace o gramatických funkčních vlastnostech, pouze grafická podoba slov a jejich výslovnost.

Nejasné označení gramatických funkčních charakteristik je jednou z největších slabin přepracovaných seznamů. V seznamu je samozřejmě mnoho slov, která mohou zastupovat různé slovní druhy, u těch však gramatické údaje zcela chybějí. Informace o slovních druzích u lexémů se stejnou grafickou formou, ale odlišnou výslovností jsou důsledně odděleny, např. u 长 (*zhǎng, cháng*), 还 (*hái, huán*) aj. V některých případech jsou funkční charakteristiky uvedeny i u bází se stejnou výslovností. I tehdy však nedůsledně, např. 对 [mnohovýznamové slovo, v seznamu chybějí významy: *stát před (něčím), být obrácen (k něčemu); spárovat, spojit*] je uvedeno jen ve funkci adjektiva a prepozičního slovesa, ale ne jako sloveso. 等 (mnohovýznamové slovo, v seznamu nejsou významy: *úroveň, stupeň; kategorie, druh*) zase pouze jako sloveso a pomocné slovo, ale ne jako podstatné jméno. Slova, u kterých byly různé funkční charakteristiky uváděny odděleně (ve 14 případech), jsem opět sloučila.

V novém seznamu jsem se setkala s další komplikací při sestavování korpusu, a sice s kolokacemi. Novější seznam uvádí kolokace jako jedinou lexikální jednotku, zatímco ve starším soupisu je tento přístup aplikován minimálně. Některé kolokace však korpusy nedokázaly společně zpracovat, např. 弹钢琴 (*hrát na klavír*), 踢足球 (*hrát fotbal*) nebo 打篮球 (*hrát basketbal*). V těchto případech jsem frekvence lexémů v korpusech zařadila jako dvě lexikální jednotky a vyhledávala odděleně. Po těchto úpravách obsahuje soupis, se kterým pracuji, 1 198 lexikálních jednotek.

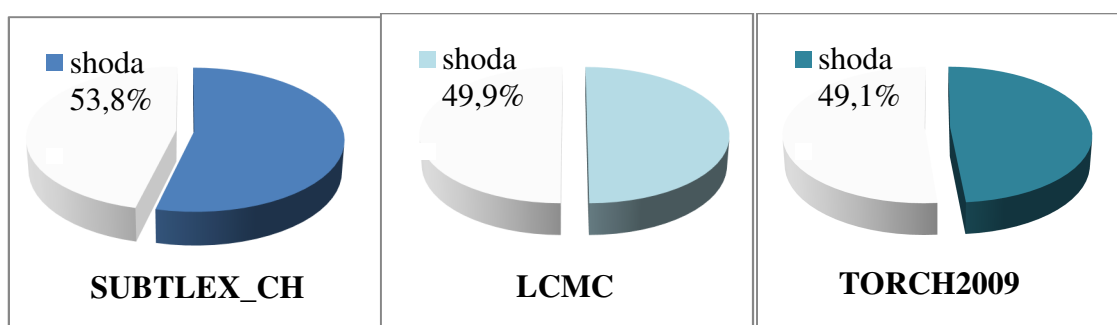
Důležitou změnou reformovaných soupisů je jejich adaptace na Společný evropský referenční rámec znalosti cizích jazyků. Pro zkoumané seznamy jde o úroveň B2, tj. středně pokročilou úroveň. V tištěných sylabech HSK4 chybějí podrobnější údaje o jazykových kompetencích studentů této úrovně, ale referenční úrovně Společného evropského rámce (SERR) jsou vypracovány velmi podrobně. V popisu úrovní zkoušky HSK na oficiálních webových stránkách jsou jazykové dovednosti charakterizovány jednou větou, u zkoumané úrovně konkrétně: „*Účastníci zkoušky, kteří jsou schopni složit HSK úroveň 4, dokáží v čínštině mluvit o široké škále témat a jsou schopni plynně konverzovat s rodilými mluvčími.*“ (Hankao Guoji ©2013–2016b)

Deklarované schopnosti uživatelů jazyka na úrovni B2 jsou přesně definovány. Následující citace jsou pouze výtahem z obecných kompetencí, podrobnější popis je přístupný v citovaném dokumentu. Student: „*Dokáže porozumět hlavním myšlenkám složitých textů týkajících se jak konkrétních, tak abstraktních témat včetně odborně zaměřených diskusí ve svém oboru.*“ (Rada Evropy 2006, s. 24) Konverzační schopnosti odpovídají výše citovaným informacím. Co se týče psaní, student: „*Umí napsat srozumitelné podrobné texty na širokou škálu témat a vysvětlit své názorové stanovisko týkající se aktuálního problému s uvedením výhod a nevýhod různých možností.*“ (Rada Evropy 2006, s. 24)

Jak již bylo výše uvedeno, vyjadřovací schopnosti studentů nového seznamu HSK čtvrté úrovně jsou podle údajů udávaných organizací Hanban očividně podstatně vyšší než v případě porovnávaného soupisu slovíček starší verze HSK. Stará verze seznamu je určena pro začátečníky, HSK4 již pro samostatné uživatele cizího jazyka, ačkoliv byla nová slovní zásoba rozšířena jen o 188 lexémů.

5.2.1 Kvantitativní posouzení seznamu 新 HSK 4 (verze 2010)

V 1 000 nejběžnějších slovech korpusů SUBTLEX-CH, LCMC a TORCH2009 je opět shoda se seznamem HSK4 (verze 2010) kolem 50 %. Výrazněji se projevil rozdíl mezi korpusem hovorové čínštiny SUBTLEX a korpusem psaného jazyka. Ve všech třech korpusech je ale shoda seznamu HSK4 nižší než ve starším soupisu. Rozdíl je nejpatrnější v korpusu LCMC s texty z roku 1991, kde podobnost se seznamem s nejméně frekventovanými lexémy klesla o 2,3 %.



Graf č. 8: Shoda soupisu HSK4 verze 2010 s 1 000 nejméně frekventovanými lexémy v korpusech

Ke zkoumání znalosti slov v textu používám stejný text jako v prvním případě, ale s odlišnými očekáváními. Studenti se znalostmi slovníčku HSK 甲级 nemají s 56 % znalostí obsažených lexémů šanci textu porozumět, což podle tvůrců slovníčku také nelze očekávat. Text obsahuje slovní hříčky a narážky vyžadující znalost kulturního kontextu, s tím by si však studenti (schopní číst složité texty) na úrovni B2 měli umět poradit (Rada Evropy 2006, s. 72).

Tabulka č. 5: Slovíčka ze soupisu HSK4 (verze 2010) v mikrobloginových příspěvcích

下午一下收到两个快递，辛苦了帮忙去取的同事!!! 还有段段这 6 个杯具!!!。。。之前只听说日本这样，国内也开始了?!。。。。

声音还原度好高啊。。。真服了，这几个家伙配音也太像了"居然是站票",这句太销魂了真“原音”，推荐收看! 怎么做到配得如此“原音”滴? 有没有人有防水卡片机啊借我用一下我可以先用尼康 S8000 交换啊。。。其实我觉得好欢乐啊。。。好可怜，好愤怒。

好有爱的卡布奇诺……终于车先生下班了终于开始吃了，yo 宝已经酒足饭饱，坐一边玩去了我的那个天……yo 宝这个鸡翅狂人！！！！一口气把三只鸡翅全部消灭光！我说要留给爸爸，她说：爸爸来了，再叫一份！yo 宝说：干杯妈妈！生日快乐！20 这段人生的最后一个数字了，明年就是另外一段十年人生路了。一场轰轰烈烈的爱情，一次说走就走的旅行。

我发现我可真是赚到了，才活了不过 20 几年，碰上了 50 年一遇的旱灾，100 年一遇的洪水，60 年一遇的大雪，300 年一遇的泥石流，500 年一遇的地震，600 年一遇的海啸……人家轮回转世好几趟都碰不全呀！！

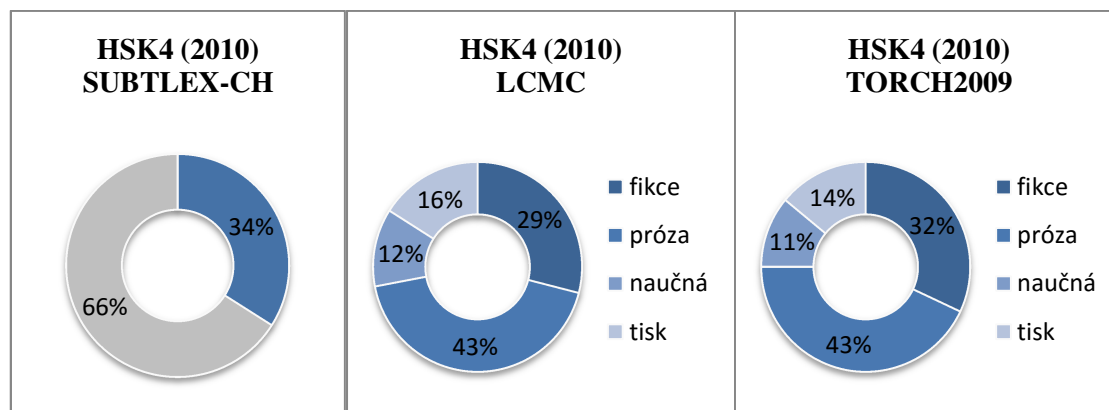
丫就不会反省一下吗？！！你比那些他妈的好多了好不好！！不想写了，赶紧赶紧的下课吧。。。

Ze 169 jedinečných lexikálních jednotek mikroblogového textu je v reformovaném seznamu obsaženo 94 slov, dalších pět až šest lexémů je snadno odvoditelných, což dohromady tvoří 59 % ze seznamu HSK4 (verze 2010). Podle údajů ze studie Andrease Gudera z organizace Fachverband Chinesisch o referenčním rámci pro hodnocení znalosti čínštiny jako cizího jazyka je třeba ke správnému pochopení textu znát minimálně 95 % slov v něm obsažených (Guder a Fachverband Chinesisch 2015, s. 18). Znalost lexémů v doporučené slovní zásobě HSK4 – co se týče výše uvedeného textu – k pochopení významu promluv nestačí, ačkoliv by podle vytyčeného standardu úrovně HSK měla.

Samozřejmě asi žádný student nespolehá na soupis slovíček HSK4 jako na jediný zdroj akvizice cizího jazyka, předpokládané studentovy schopnosti pochopit význam vyznačených slov a porozumět textu tedy i po ovládnutí slovíček ze seznamu budou vyšší než pouze místa vyznačená v tabulce. Totéž ale platí pro studenty používající starý soupis HSK, i tak tvoří rozdíl mezi lexémy nejnižší úrovně (HSK 甲级) a středně pokročilé úrovně obsaženými v textu (verze 2010) pouhá 3 procenta. Obsah lexémů v mikroblogovém textu ze slovníčku je mírně vyšší než ve staré slovní zásobě, ta byla ovšem také o 188 slovíček rozšířena, vyšší zastoupení v textu je tedy pravděpodobné a nesvědčí o lepším složení soupisu.

Ve výskytu lexémů ve filmech si reformovaný seznam výrazně pohoršil, na slovíčka obsažená v novém seznamu lze narazit v průměru pouze v 34 % zkoumaných

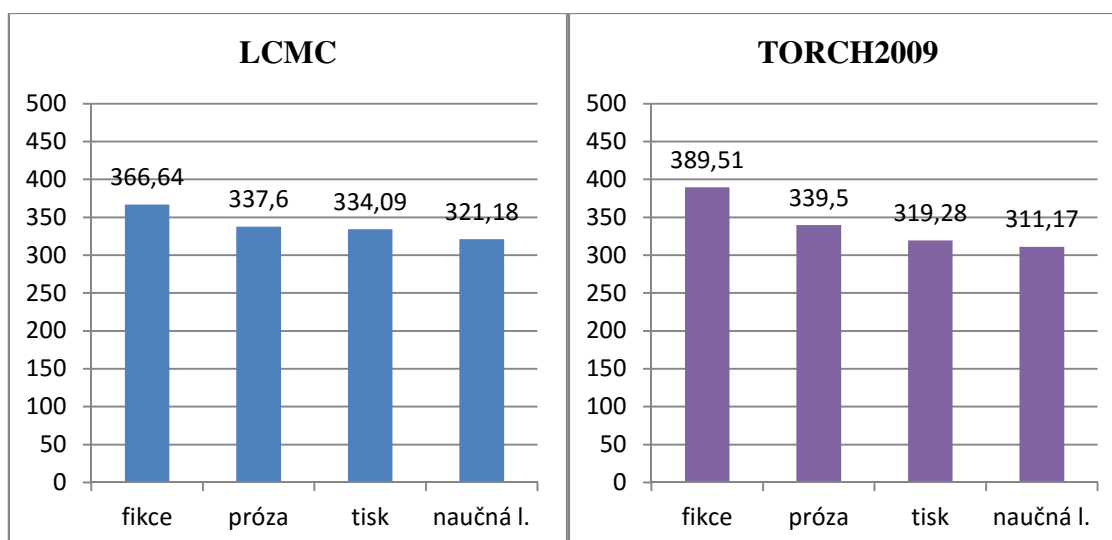
filmů, což je o 6 % méně než v seznamu HSK 甲级. Korpus filmových titulků je poměrně aktuální (z roku 2010) a filmy mají větší potenciál reflektovat způsob užití moderního hovorového jazyka než textové korpusy, přesto je v tomto ohledu seznam devadesátých let pro studenty praktičtější.



Graf č. 9: Výskyt lexémů soupisu HSK4 (2010) ve filmové databázi a textových žánrech

Zastoupení lexémů v žánru naučné literatury je v obou korpusech o jedno procento nižší než ve starém seznamu slovní zásoby. Poměr v žánru publicistiky se změnil jen ve složení korpusu TORCH2009, kde zastoupení stouplo o jedno procento. Ostatní dva obecnější žánry prózy a literární fikce se poměrně v korpusu LCMC změnil vyším zastoupením prózy a o 3 % menším zastoupením fikce (v HSK 甲级 je naopak nejrozšířenější fikce). V TORCH2008 se zastoupení fikce také o 3 % zmenšilo, zatímco u tří zbylých žánrů stouplo o 1 %. Žánrová distribuce v korpusech LCMC a TORCH2009 nevykazuje podstatné změny, rozdělení je téměř stejné jako v případě soupisu HSK 甲级. Lexémy obsažené ve slovníčku HSK4 (verze 2010) mají primárně obecný charakter.

Poměry v průměrném frekvenčním zastoupení již nevykazují o tolik výrazněji vyšší frekvenci v žánru literární fikce, průměrné hodnoty žánrů jsou v případě korpusu LCMC o něco vyrovnanější. Celková průměrná frekvence je jako v ostatních případech v seznamu HSK4 (verze 2010) v porovnání se starým seznamem ve všech žánrech nižší.



Graf č. 10: Průměrná frekvence seznamu HSK4 (2010) v textových žánrech

Podle požadavků Evropského referenčního rámce by studenti úrovně B2 měli být schopni porozumět i odbornějším textům (Rada Evropy 2006, s. 72), to je však vzhledem k vyhodnoceným datům a celkovému složení slovní zásoby HSK4 nepravděpodobné. Stejně tak by měli porozumět textům v tisku, průměrná frekvence pro publicistické žánry z roku kolem 2009 (TORCH) je však nízká a navíc ještě nižší než pro žurnalistické texty publikované kolem roku 1991 (LCMC). Kromě toho mají čínské novinové texty specifickou slovní zásobu a se znalostí soupisu HSK4 by čtenář neporozuměl ani názvům některých rubrik.

5.2.2 Kvalitativní posouzení seznamu 新 HSK 4 (verze 2010)

Soupis slovní zásoby obsahuje slova obecného charakteru z mnoha oblastí každodenní komunikace. Je v něm obsažen menší počet synonym, v mnoha případech jde o varianty typické pro hovorovou a psanou čínštinu či formálnější rovinu jazyka, např. 不管/无论 (*bez ohledu na*) 确实/果然 (*opravdu, vskutku*), 差不多/大概 (*asi, zhruba*) atd. Poměr slov charakteristických pro hovorovou a psanou čínštinu je vyrovnaný.

Ze seznamu byla odebrána směrově modifikovaná slovesa, ale místo toho byla zavedena tři v hovorové čínštině často používaná potenciálně modifikovaná slovesa, a to: 来得及 (*stihnout*), 来不及 (*nestihnout*) a 受不了 (*nemoci snést*).

Již jsem se zmínila o zrušení mnoha vysoce frekventovaných bází a jejich zařazení do kompozit. Tato změna nebyla provedena příliš promyšleně, např. báze 肉 (*maso*) byla nahrazena kompozitem 羊肉 (*skopové*). Ve velkých městech vnitřní Číny je skopové však drahým a spíše sezonním druhem masa, ale všudypřítomné kuřecí či vepřové maso v seznamu zcela chybí. Stejný problém je i u báze 车 (*vůz*), kdy tvůrci seznamu zapomněli kromě *kola*, *taxi* a *autobusu* do slovníčku zařadit i *auto* a *vlak*, oba lexémy s vyšší frekvencí výskytu.

Novým prvkem v seznamu je zařazení kulturně a geograficky specifických slovíček. Nově tak bylo přidáno hlavní město Číny 北京 (*Peking*) a 长江 (*Dlouhá řeka*). Z čínských kulturních specifik to jsou pak např. 饺子 (*plněné knedlíčky*), 长城 (*Velká čínská zeď*) nebo 京剧 (*čínská opera*).

Tematické rozdělení slovíček se také změnilo. Slova typická pro studium cizího jazyka byla z nového seznamu z větší části odstraněna, zato byla přidána slova související s profesním životem jako 做生意 (*obchodovat*), 顾客 (*zákazník*) či 招聘 (*nabídka práce, nabízet práci*). Také byla doplněna slovíčka potřebná pro vycestování do Číny 签证 (*vízum*) a 护照 (*cestovní pas*) či 航班 (*letecký*). Přidána byla nová slova týkající se internetu a technologií, jako např. 上网 (*surfovat po internetu*), 网站 (*webová stránka*), 电脑 (*počítač*) a 手机 (*mobil*).

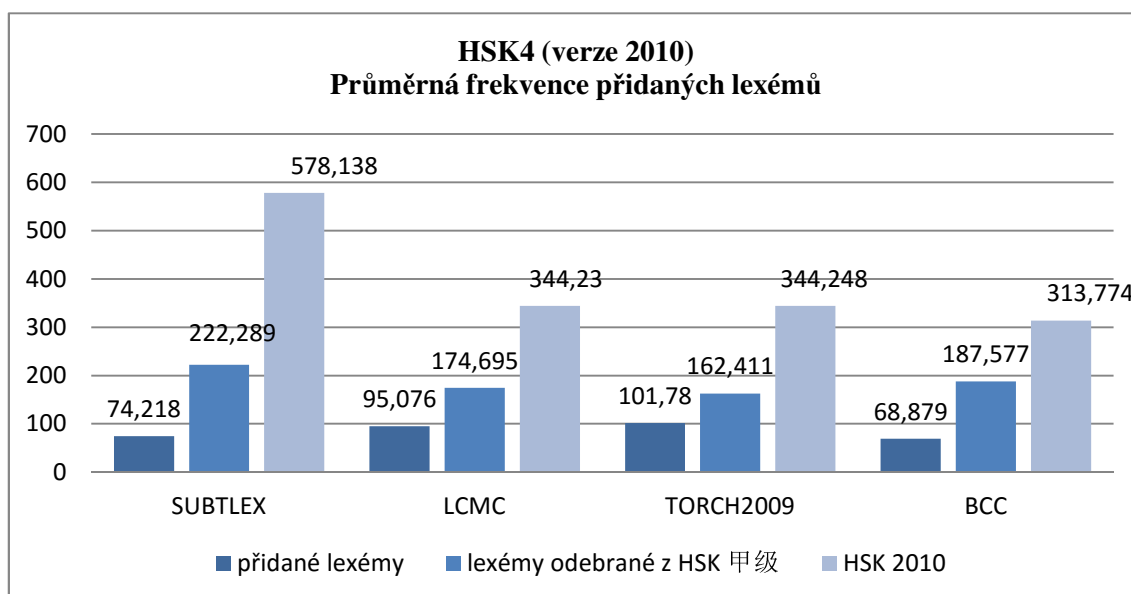
Zařazení slovíček, jako jsou 熊猫 (*panda*), 狮子 (*lev*), a jinak velkého množství různých druhů domácích i divokých zvířat, na které málokdy přijde řeč (např. opice a kůň), do nové slovní zásoby je trochu zarážející vzhledem k tomu, že cílovou skupinou účastníků zkoušek HSK jsou dospělí studenti a nikoli děti, pro které byla vytvořena zkouška YCT (Zkouška z čínštiny jako cizího jazyka pro mládež).

Reprezentativnost slovíček pro pokrytí komunikačních situací v každodenním životě je tedy sporná. Některá z výše uvedených málo užívaných slovíček patří mezi prototypní představitele jisté kategorie (čínské jídlo: *plněné knedlíčky*, čínská řeka: *Dlouhá řeka*, typické zvíře: *panda*), a tak má jejich zařazení do seznamu jisté oprávnění. Další nesrovnatelně frekventovanější či užitečnější a aktuálnější slovíčka ze staršího seznamu však chybějí, např. 不错 (*správný, dobrý, bezvadný*) nebo 服务 (*služba*).

Tvůrci seznamu spoléhají na to, že si studenti spíše odvodí významy dalších slovíček z kompozit, než že budou poznávat nová slovíčka z významů bází, což souvisí s upuštěním od jejich uvádění v soupisu. To vede k nesystematičnosti ve složení slovníčku, např. jsou součástí soupisu slova 北方 (*severní oblast*) a 南 (*jih*), ale ne jejich přímé protiklady 南方 (*jižní oblast*) 北 (*sever*). Nesystematičnost se týká i uvádění kolokací. *Hrát fotbal* je ve slovníčku jako celá kolokace 踢足球 (doslova *kopat fotbal*), což se dá pochopit. U kolokace *hrát basketbal*, kdy se ke hře používají přímo ruce, se student dozví, že se k basketbalu váže sloveso 打 (v tomto případě *hrát*). Může tedy znejistět, jaké sloveso se váže k dalším slovíčkům ze seznamu, jako jsou *tenis* a *badminton*, ke kterým je třeba raketa. Že se tato slova také pojí se slovesem 打, ale ze soupisu nevyčtou.

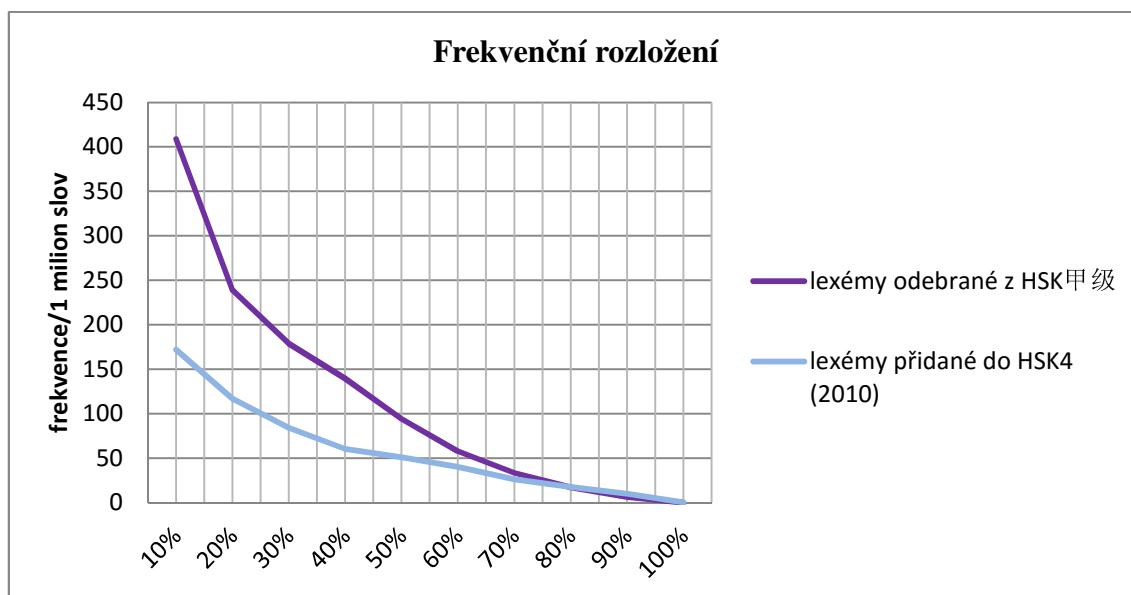
5.2.3 Změny v seznamu HSK4 (verze 2010)

Když porovnáme změny v průměrné frekvenci lexémů vyřazených ze seznamu HSK 甲级 s nově přidanými lexémy seznamu HSK4, zjistíme, že jsou přidaná slovíčka v průměru mnohem méně frekventovaná než ta, která byla ze soupisu odstraněna. Největší rozdíl je patrný v korpusech hovorové čínštiny SUBTLEX-CH a BCC(微信), kde jsou přidaná slovíčka téměř třikrát méně frekventovaná než ta zrušená.



Graf č. 11 Průměrná frekvence lexémů přidaných do seznamu HSK4 (2010)

Důležitý je ovšem i frekvenční poměr přidaných a odebraných slovíček. Již víme, že ze starého seznamu byly odebrány i velmi frekventované lexikální jednotky, které tak mohou průměrné hodnoty v korpusech v porovnání s přidanými jednotkami příliš zkreslit. Následující graf proto ukazuje frekvenční průměr slov ze všech čtyř korpusů u *x* procent slov s nejvyšší frekvencí výskytu.

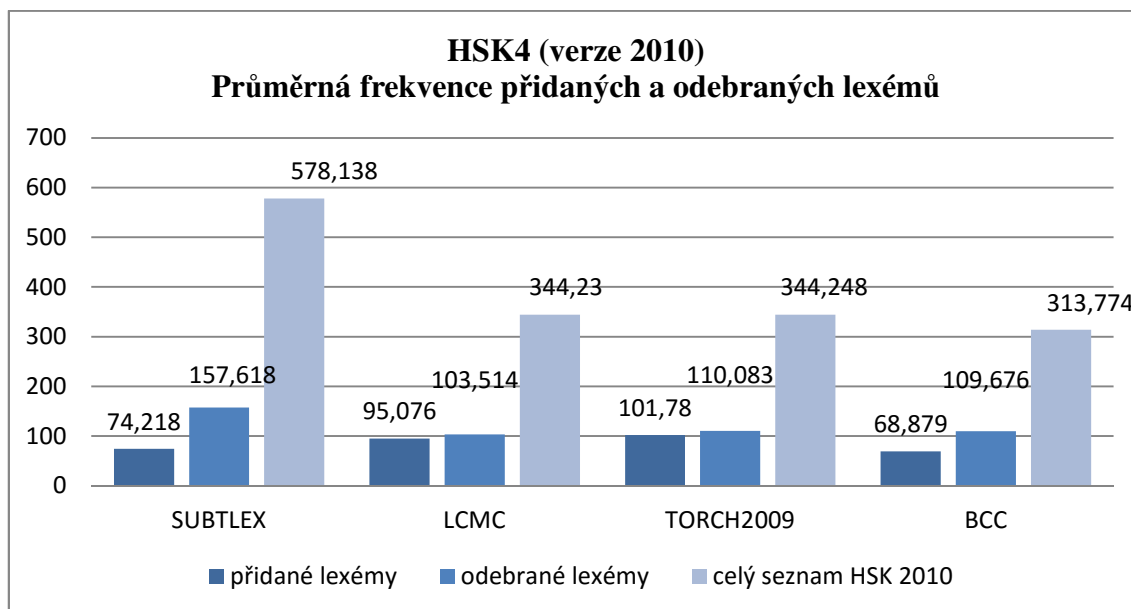


Graf č. 12 Frekvenční rozložení lexémů odebraných a přidaných do seznamu HSK4 (2010)

Nejfrekventovanější vyňaté lexémy jsou opravdu mnohem běžnější než nejfrekventovanější přidaná slova, vyšší frekvence zrušených slovíček však přetrvává až do 80 % nejfrekventovanějších slov, kde je potom jenom o málo vyšší než frekvence odebraných lexémů. To nevyovídá o lepším složení reformovaného slovníčku pro potřeby ovládnutí čínštiny. Důvodem mnohem nižší frekvence je opět odebrání jednoslabičných bází. Je s podivem, že tvůrci seznamu trvají na nahrazování bází kompozity, když jsou jednoslabičná slova tak běžnou a potřebnou součástí jazyka.

Z deseti nejfrekventovanějších slov v nově přidaných slovech v seznamu HSK4 (verze 2010) je osm jednoslabičných bází, celkově ale tvoří pouze 12 % přidaných slovíček. Většina z těchto deseti nejběžnějších jednoslabičných lexémů je typická spíše pro formální jazyk. Podstatnou část odebraných jednoslabičných bází tvoří měrová jména. Dvouslabičná kompozita tvoří celých 82 % nových slovíček a z velké části se jedná o autosémantika. Ta mají výhodu často konkrétního, snadno zapamatovatelného významu a menšího počtu denotátů. Velká část těchto autosémantik ale není kvůli nízké frekvenci v jazyce příliš užitečná.

Koncem roku 2012 byl seznam slovní zásoby HSK4 přepracován. 82 slovíček bylo zrušeno a nahrazeno stejným počtem nových slov. V grafu č. 13 jsou zobrazeny průměrné frekvence lexémů, které byly k novému soupisu HSK4 z roku 2010 přidány nebo které z něj byly odebrány.



Graf č. 13: Průměrná frekvence lexémů po změnách v seznamu HSK4 (2010)

Z 82 zrušených slovíček byla většina (65 %) součástí slov nově přidanych do seznamu HSK4 (verze 2010). Odebrané lexémy mají vyšší průměrnou frekvenci než lexémy v roce 2012 přidané, rozdíl je ale podstatný pouze u korpusů hovorového jazyka.

5.3 Obecná charakteristika seznamu slovní zásoby 新 HSK4 z roku 2012

Důvody reformy seznamu slovíček v roce 2012 nejsou na oficiálních stránkách Hanbanu uvedeny. S poznámkou „2012年修订版“ (přepracovaná verze z roku 2012) byla aktualizována pouze nová verze soupisu (Hankao Guoji, ©2013–2016a). Od té doby zůstal seznam – co se týče složení lexémů – nezměněn. Není zcela jisté, zda nyní bude soupis slovíček upravován v určitých časových intervalech nebo zda šlo na delší dobu o jediné přepracování. Rozsah zkoumaného seznamu HSK4 (1 200 lexikálních jednotek) se nezměnil. Hanban také nepřepočítal nároky na studenty a deklarované stupně kompetence pro jednotlivé úrovně.

Soupis, který je možné stáhnout přímo z oficiálního webu Hanbanu (www.chinesetest.cn), je ještě méně použitelný než verze 2010. Ten uváděl informace o grafické podobě slova a jeho transkripci v pinyinu, kdežto nynější seznam tvoří pouze abecedně seřazená slova bez informací o výslovnosti. Není ani vyznačeno, která slova byla v nové verzi přidána a která odebrána.

Důležitá změna v možnostech využití slovní zásoby HSK nastala díky zpřístupnění vzdělávacích internetových zdrojů. Roku 2014 Hanban ve spolupráci s Konfuciovým institutem spustil nenápadnou platformu pro zájemce o výuku čínštiny www.cltguides.com. Tento web s nepříliš působivým názvem 国际汉语教材编写指南 (*Průvodce sebranými mezinárodními materiály pro výuku čínštiny*) nabízí texty a cvičení navržené tak, aby odpovídaly obsahu slovní zásoby jednotlivých úrovní nové zkoušky HSK.

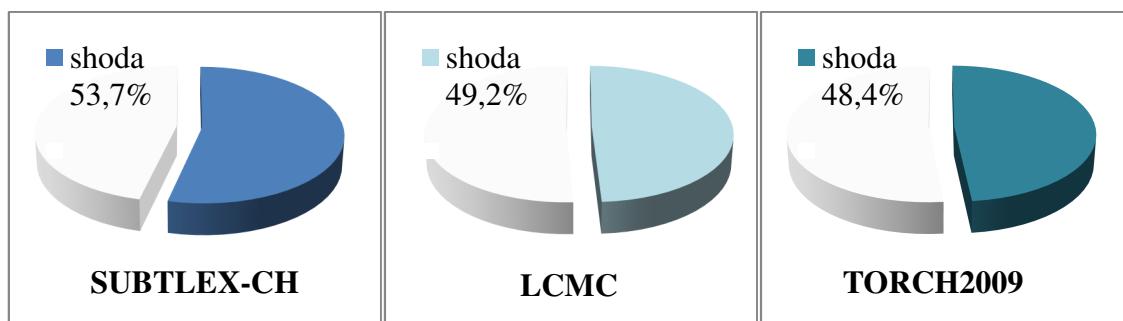
Na webu cltguides.com lze také interaktivně pracovat se seznamy slovní zásoby. Ty jsou mnohem použitelnější než oficiální verze na chinesetesting.cn. Soupisy jsou zpracovány velmi podobně jako seznamy staré verze. Kromě seznamu slov lze prohledávat seznamy jednoslabičných bází a morfémů, ze kterých jsou kompozita složena. U HSK4 (verze 2012) je to 1 000 jednotlivých znaků. Dalším doplňujícím seznamem je soupis gramatických vazeb, kde jsou ve 4. úrovni zkoušky HSK vysvětleny způsoby užití celkem 205 gramatických struktur.

Samotný seznam slovíček uvádí informaci o úrovni, do které slovíčko patří, jeho grafickou podobu, výslovnost v pinyinu, definici v čínštině, angličtině a několik příkladových vět. Všechna data kromě příkladových vět lze poměrně snadno hromadně nahrát do tabulky v Excelu a tak je zpracovat k dalšímu použití.

Mnou zkoumaný korpus slovní zásoby HSK4 (verze 2012) obsahuje celkem 1 197 lexikálních jednotek, o jednu méně než seznam starší verze, ačkoliv byl počet odebraných a přidaných slov stejný. Je tomu tak proto, že ve verzi z roku 2012 je slovo 花 (mnohovýznamové slovo, v seznamu ve významu *květina* a *utrácet; trávit*) nově rozděleno do dvou funkčních charakteristik na podstatné jméno a sloveso. Kvůli vyhledávání v korpusech jsem je musela sloučit.

5.3.1 Kvantitativní posouzení seznamu 新 HSK 4 (verze 2012)

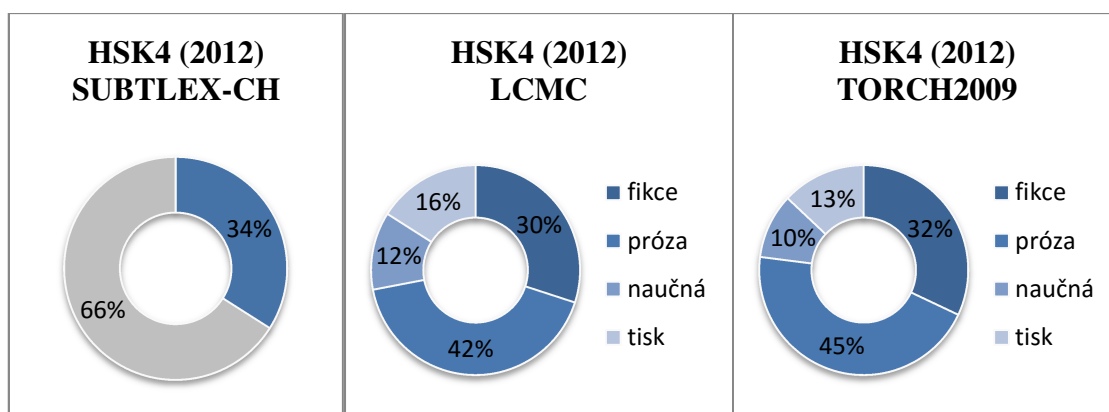
Po změnách v původním novém soupisu HSK4 je shoda s nejfrekventovanějšími slovy v korpusech o málo nižší než v nepřepracovaném seznamu. V korpusu hovorové čínštiny SUBTLEX-CH jde o pokles o jednu desetinu procenta. V korpusu LCMC a TORCH2009 je rozdíl také poměrně malý, a to 0,7 %. Změny v seznamu se z pohledu frekvence nejužívanějších slov projeví negativně.



Graf č. 14: Shoda soupisu HSK4 verze 2012 s 1 000 nejfrekventovanějšími lexémy v korpusech

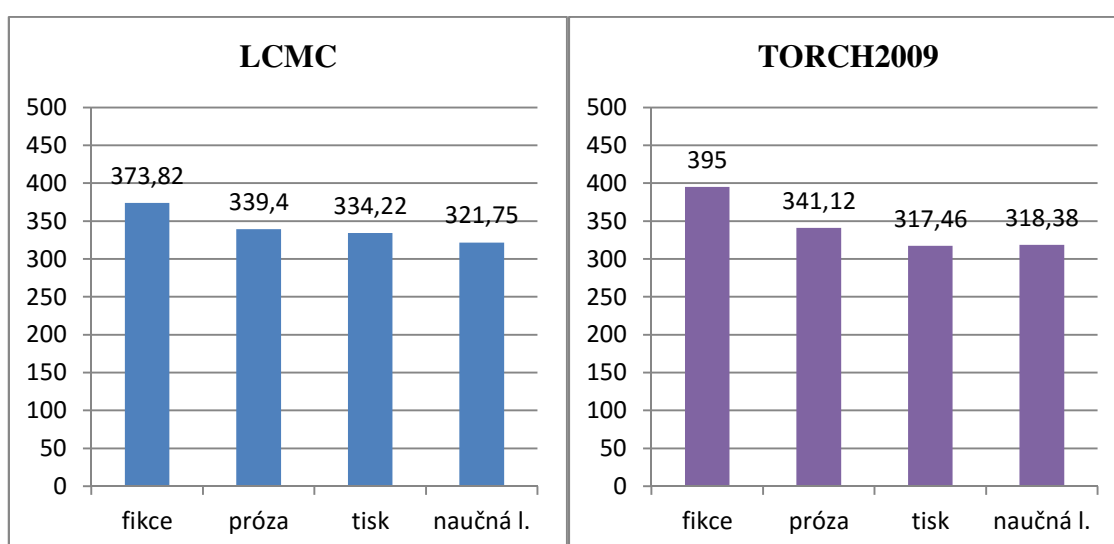
Tabulku mikroblogového textu k ilustraci změn v seznamu není třeba znovu uvádět, protože by zobrazený obsah textu s vyznačenými slovíčky byl téměř identický. V textu se změny projeví mírně pozitivně, srozumitelný poměr slov stoupl o 1 %, z 59 na 60 %. Stejně jako v předchozí verzi pro čtenáře se znalostí slovíček ze seznamu HSK4 zůstává smysl celých příspěvků nesrozumitelný.

Zastoupení lexémů ze slovníčku v procentech filmů se po přepracování seznamu projevilo pouze v desetínách procenta. Počet filmů, v nichž se objevují slovíčka ze seznamu verze z roku 2012, je v průměru o necelé dvě desetiny procenta větší než u předchozího seznamu. Žánrové rozdělení vykazuje větší rozdíl jen v korpusu TORCH, distribuce žánrů v korpusu LCMC je téměř identická. Necelé tři čtvrtiny slovíček byly užity v žánrech obecné prózy a literární fikce. V korpusu TORCH mají obecné žánry 77% zastoupení, poměr výskytu lexémů v publicistických a odborných textech klesl v obou případech v porovnání s předchozí verzí seznamu o jedno procento.



Graf č. 15: Výskyt lexémů soupisu HSK4 (2012) ve filmové databázi a textových žánrech

Co se týče průměrné frekvence užití slov, ve všech žánrech jsou lexémy nového soupisu mírně běžnější než v předchozím seznamu. Žánrové rozdělení vyrovnanější není, ale změny v průměrné frekvenci se projeví pozitivně.



Graf č. 16: Průměrná frekvence seznamu HSK4 (2012) v textových žánrech

5.3.2 Kvalitativní posouzení seznamu 新 HSK 4 (verze 2012)

Do nové verze seznamu HSK4 byly zařazeny gramatické vazby společně v podobě jedné lexikální jednotky, stejně jako byly uvedeny ve starém seznamu HSK 甲级. Dále byl rozšířen počet frázových sloves o 打招呼 (*zdravit; oznámit, informovat*) a 感兴趣 (*mít zájem o*). Společně jsou nově uvedeny jako lexikální jednotky slova 只有 (*jenom*), 还有 (*ještě, navíc*), 没有 (*nemít*). Metoda lexikálního přístupu při sestavování slovní

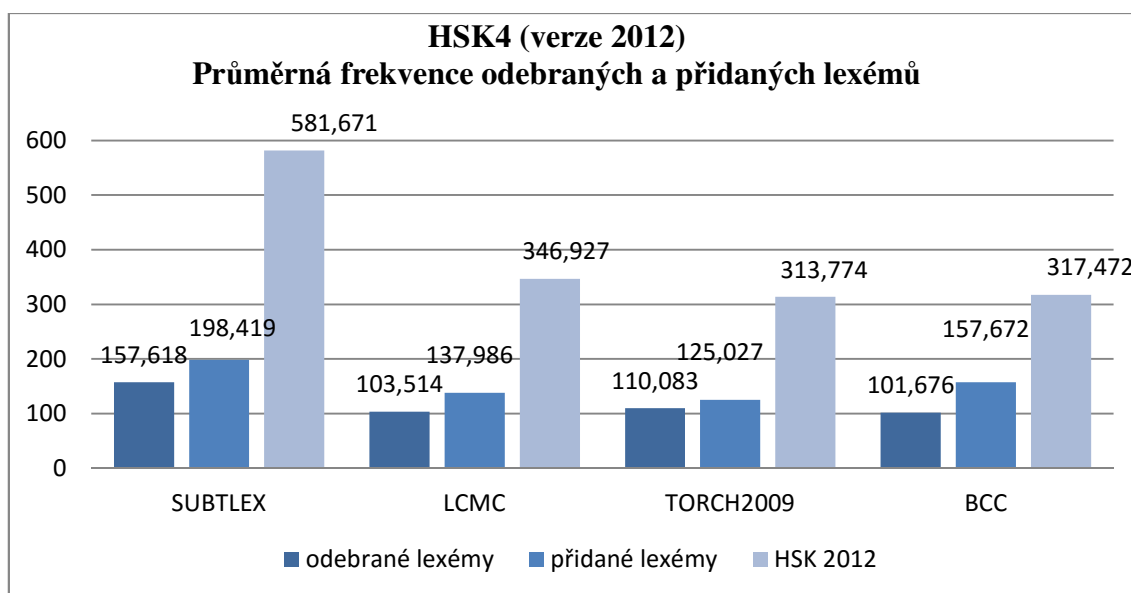
zásoby je v nové verzi seznamu aplikována mnohem zřetelněji. Také lze sledovat ještě vyšší tendenci odstraňování jednoslabičných slov a jejich nahrazování dvouslabičnými kompozity než v dřívější verzi seznamu.

V nové verzi soupisu se změnilo necelých 7 % slov. Na celkové složení seznamu tak změny neměly velký vliv, odebrané i přidané lexémy patří do různých tematických okruhů, mají různé gramatické funkce a nelze ani určit větší či menší převahu slov charakteristických pro hovorový či psaný jazyk. O zásadnějších změnách lze mluvit pouze v rámci 82 nových lexémů.

5.3.3 Změny v seznamu HSK4 (verze 2012)

Nově zavedená i zrušená slovíčka jsou velmi různorodé povahy, nelze určit přesný motiv změn v seznamu. Nahrazení málo použitelných slov běžnějšími a užitečnějšími lexémy pravděpodobně motivem změny nebylo, a pokud ano, provedení nebylo příliš zdařilé. Z informací z grafu č. 4 jsme se dozvěděli, že celková průměrná frekvence lexémů ve slovníčku stoupla pouze minimálně.

V případě změněných lexémů jsou rozdíly v průměrné frekvenci zřetelnější, ale ve skutečnosti patří většina jak odebraných, tak nově přidaných slovíček mezi málo frekventovaná slova. Pokud z obou kategorií odebereme pět nejfrekventovanějších lexémů, průměrná frekvence ze všech korpusů se vyrovná (74 výskytů/milion slov ve zrušených slovíčkách a 74,1 výskytů/milion slov v nově přidaných). Poměr složení frekvence v obou kategoriích je také velmi podobný.



Graf č. 27: Průměrná frekvence lexémů po změnách v seznamu HSK4 (2012)

Změny ve slovní zásobě se v mnoha případech týkají spolu souvisejících slovíček. Byla přidána antonyma k již dříve zařazeným slovíčkům, např. 重 k 轻 (*těžký k lehký*), 降落 k 起飞 (*přistát k vzlétnout*) nebo 应聘 k 招聘 (*ucházet se o práci k nabízet práci*). Úpravy byly provedeny i ve variantách slov, např. místo 饭馆 bylo zařazeno slovo 饭店 (*restaurace*), 刚刚 místo 刚 (*zrovna, právě*) nebo 风景 místo 景色 (*scénérie*). Ve všech případech kromě posledně jmenovaného se jedná o nahrazení více frekventovanou variantou.

Přidána byla některá dosud chybějící slova týkající se moderních technologií jako 短信 (*SMS*) nebo 互联网 (*internet*). Dále byla rozšířena slovíčka úzce související s Čínou, např. 烤鸭 (*pečená kachna*), 包子 (*plněné knedlíky*) a 黄河 (*Žlutá řeka*), přestože jde o lexémy s velmi nízkou frekvencí. Důležitým důvodem celkově nízké frekvence přidaných slov je zahrnutí velkého počtu autosémantik, která jsou užitečná jen ve velmi specifických situacích, např. 登机牌 (*palubní lístek*) či 占线 (*obsazená telefonní linka*).

Z celkového pohledu lze říci, že změny nepřinesly uživatelům slovníčku žádnou zásadní výhodu při akvizici jazyka, neboť se vesměs nejedná o běžná či v jazyce velmi důležitá slova. To však mohlo studenty zmást, protože z oficiálních zdrojů se uživatelé dozvědí pouze o úpravách v seznamu, ale ne o rozsahu, v jakém byly provedeny, nebo kterých slovíček se změny týkaly.

5.4 Shrnutí

Na základě frekvenční analýzy seznamů slovní zásoby bylo zjištěno, že starší verze slovníčku HSK 甲级 obsahuje slovíčka s nejvyšší průměrnou frekvencí. Seznam je složen ze slovíček základní slovní zásoby, s nimiž se studenti často setkají v hovorové i psané čínštině. Soupis HSK 甲级 klade důraz na uvádění bezpříznakových slov a produktivních jednoslabičných bází. Jeho důležitým nedostatkem pro studenty je z dnešního pohledu absence dnes běžných slov týkajících se moderních technologií. Slabým místem je také uvádění zbytečně velkého množství synonymních slov.

Další nevýhodou je omezená možnost využití slovníčku k účelu složení zkoušky HSK. Dnešní testy jsou navrženy tak, aby v jednotlivých úrovních korelovaly s nově navrženou slovní zásobou. Všechny testy jsou sestaveny podle nového soupisu a musejí pokrývat určité množství procent ze slovíček doporučené slovní zásoby (Zhou 2012, s. 2). Vzhledem k velkým rozdílům v obou seznamech není vhodné používat slovníček HSK 甲级 místo seznamu HSK4. Učení se starší verze slovníčku k přípravě na nižší úroveň nového testu HSK je kontraproduktivní, protože je mnohem obsáhlejší a náročnější na studium než soupis slovíček HSK3.

Uvedené slabiny slovníčku HSK 甲级 však nesouvisejí s horší použitelností slovníčku, nýbrž se změnou systému zkoušky. Pokud studentům jde o osvojení si základní slovní zásoby čínského jazyka nezávisle na učebnicích, je soupis HSK 甲级 tou nejlepší volbou. Díky své přehlednosti, vyrovnanému složení a jasně a přiměřeně definovaným kompetencím je vhodný pro samostatné studium. Navíc použití připojených seznamů znaků a gramatických struktur usnadní studentům pochopení systematičnosti čínského jazyka.

Seznam HSK4 z roku 2010 vyšel z výzkumu jako nejméně vhodná volba, slovíčka v seznamu nejsou v jazyce příliš běžná. Na zkoumané úrovni je vidět snahu diferencovat slovní zásobu zařazením lexémů ze specifických tematických oblastí za cenu vynechání mnohem frekventovanějších slov potřebných v běžném životě. Obsahuje také mnohem více slov charakteristických pro hovorový a psaný jazyk, i když by bylo užitečnější zařadit různé varianty bezpříznakových slov jako ve starším seznamu. Největším problémem seznamu je málo jednoslabičných bází a téměř chybějící vyznačení gramatických funkčních charakteristik.

Zlepšení v užitečnosti a srozumitelnosti slovní zásoby nastalo zejména díky zpřístupnění slovníčku jednotlivých znaků a gramatických struktur. V hlavním seznamu lexémů však stále přetrvává problém nízkého počtu jednoslabičných produktivních bází a celkově nízké frekvence obsažených slov. Pokud je motivací se studiu slovníčku úspěšné absolvování zkoušky HSK4, je verze soupisu z roku 2012 nejen z důvodu aktuálnosti lepší volbou než seznam z roku 2010. Pokud se studenti ovládnutím slovíček snaží dosáhnout úrovně B2 SEER, zdaleka to nestačí.

Slovníčky mají svou hodnotu v možnosti plánovaného a vědomého osvojení si důležitých slov v jazyce. Proto by měl být kladen důraz na zařazení nejběžnějších slov a bází, které mívají komplikovanější konotace a denotace, takže vyžadují časté opakování a ovládnutí široké škály různých významů.

6 ZÁVĚR

Tato práce pojednává o vývoji standardizované zkoušky z čínštiny jako cizího jazyka (HSK). Je v ní popsána historie zkoušky, přechod od starého formátu zkoušky k reformě provedené Pekingskou jazykovou univerzitou až k novému formátu spravovanému Hanbanem. Stručně uvádí, jakou formu měla zkouška HSK v kompetenci Pekingské jazykové univerzity, i popis formátu nynější nové zkoušky HSK. Kromě struktury testu je popsáno také jeho rozdělení do úrovní a rozsah slovní zásoby. V práci jsou zmíněny rozdíly mezi starou a novou formou zkoušky. Bylo zjištěno, že starší i nová forma zkoušky HSK srovnatelných úrovní mají podobný rozsah a náročnost, ačkoliv jsou cílovou skupinou začátečníci a mírně pokročilí studenti. Tyto úrovně testů se od sebe liší primárně různým rozdělením do úrovní a různou doporučenou slovní zásobou.

V další kapitole je popsáno hodnocení zkoušky. Jsou uvedeny nejčastější důvody kritického hodnocení zkoušky, ale také pozitiva v její staré, reformované i nové podobě. Nejdůležitějším předmětem kritiky je u starší formy přílišná náročnost zkoušky, u nové formy pak příliš nízká náročnost v rámci daných úrovní kompetence. Nynější podoba zkoušky HSK podceňuje nároky na naučení se čínského jazyka a poskytuje zavádějící informace o úrovních testu podle SERR.

Dále se zabývám osobitostí čínského jazyka a jeho obtížností z pozice západních studentů-začátečníků. Popisuji roli slovní zásoby při akvizici cizího jazyka. Poté se věnuji vlastnostem seznamu slovní zásoby. Udávám, jaké informace by měl ideálně sestavený slovníček pro studenty cizího jazyka obsahovat. Výběr lexémů ve slovníčku pro studenty na zkoumaných úrovních posuzuji z hlediska příslušnosti do základní slovní zásoby a vysoké frekvence lexémů. Dále se stručně zabývám atributy, které by měl mít vyvážený lexikon, a popisuji několik metod a teorií, podle nichž lze z kvalitativního hlediska posoudit způsob stavby slovníčku.

V praktické části jsem na základě analýzy frekvence lexémů zkoumala soupisy slovní zásoby starší verze HSK pro začátečníky (HSK 甲级), seznam HSK4 z roku 2010 a upravený seznam HSK4 z roku 2012. K analýze jsem použila korpus filmových titulků SUBTLEX-CH, korpus mikroblových příspěvků BCC(微博) a korpusy

různých textových žánrů LCMC a TORCH2009. Následně jsem ve čtyřech zmiňovaných korpusech tyto tři seznamy kvantitativně a kvalitativně posoudila.

Bylo zjištěno, že starší verze slovníčku HSK 甲级 obsahuje slovíčka s nejvyšší průměrnou frekvencí, o 10 % více jednoslabičných bází než seznamy HSK4, a sdílí více nejfrekventovanějších lexémů v jednotlivých korpusech než soupisy HSK4. Ve slovníčku HSK 甲级 je kladen větší důraz na roli frekvence lexémů a na zařazení slovíček ze základní slovní zásoby. Nedostatky soupisu HSK 甲级 se projevily zejména v příliš velkém obsahu úplných synonym a také již není pro účely složení zkoušky HSK aktuální. Výhodou seznamu slovní zásoby HSK 甲级 je i dnes jeho všeobecná uplatnitelnost, neboť obsahuje zvláště lexémy obecného, základního charakteru a množství bází, které usnadní studentům porozumění novým slovíčkům.

Oba soupisy HSK4 z hlediska zkoumaných hodnot srovnatelné. Důvodem nižších hodnot než ve starém seznamu, co se týče frekvence, průměrného výskytu slov v různých žánrech a filmech, je zařazení mnoha kompozit a slovíček potřebných pro specifické a ne běžné použití a hlavně odstranění mnoha bází. Největším pozitivem současného seznamu je možnost jeho upotřebení v souvislosti s online zdroji nabízenými Hanbanem. Největším nedostatkem tohoto seznamu je, že v případě nevyužití dalších zdrojů zůstává jen výčet slov bez dalších informací.

Výchozí hypotéza, že v současnosti platný seznam je z hlediska akvizice a orientování se v moderním čínském jazyce pro studenty usilující o dosažení daných úrovní nejvhodnější, se potvrdila jen částečně. Současný seznam je nejvhodnější pro dosažení úrovně HSK4, protože by testy HSK4 měly být složeny alespoň z 85 % z těchto slovíček. Z tohoto důvodu je vhodnější než seznam z roku 2010, ve kterém se 7 % slovíček změnilo. K dosažení deklarované úrovně B2 osvojení si soupisu HSK4 nestačí, protože nemá dostatečný rozsah a nepokrývá lexémy, které by uživateli umožnily komunikaci a porozumění podle požadavků SERR na úrovni B2.

Předpokládám, že moje diplomová práce bude přínosem zejména pro studenty a zájemce o zkoušku HSK, kteří budou moci lépe posoudit objektivnost zkoušky HSK a užitečnost seznamů slovní zásoby při akvizici čínského jazyka.

Summary

In this thesis the author examines the development of the Chinese Proficiency Test and analyses selected levels of recommended vocabulary of the HSK test. The intention is to compare the list of vocabulary for beginners of the older HSK test with the new HSK4 vocabulary and a list of the same level redesigned in 2012. The theoretical part focuses on the development of the tests, on controversy about the change in the test structure and the challenges regarding the creation of glossary for students of foreign languages. In the practical part of the thesis, the author uses the LCMC, TORCH2009, SUBTLEX-CH and BCC (微信) corpora to analyze the frequency of occurrence of lexical items in all three lists and their distribution in texts of various genres. By means of frequency-based analysis of the vocabulary lists, it was found that the older form of the HSK vocabulary better reflects the use of language in written and spoken standard Chinese than the new HSK4 lists.

Keywords:

HSK, Hanyu Shuiping Kaoshi, frequency-based analysis, corpus, frequency, Hanban, learning Chinese as a foreign language

Seznam literatury a dalších zdrojů

- ALCARAZ-MARMOL, Gema (2015): „Dispersion and frequency: Is there any difference as regards their relation to L2 vocabulary gains?“ *IJES*. roč. 15, č. 2, s. 1–16. ISSN: 1578-7044.
- BENTZ, Christian et al. (2016): „Variation in word frequency distributions: Definitions, measures and implications for a corpus-based language typology.“ In *Journal of Quantitative Linguistics* (přijato). s. 1–40.
- BLCU HSK zhongxin „HSK gaijinban“ xiangmu xiaozu 北京语言大学汉语水平考试中心 “HSK 改进工作” 项目组 (2007): „Hanyu Shuiping Kaoshi (HSK) gaijin fang'an 汉语水平考试(HSK)改进方案.“ *Shijie hanyu jiaoxue* 世界汉语教学, č. 2, s. 126–136.
- BLCU HSK zhongxin 北京语言大学汉语水平考试中心 (2008): „Zhongguo Hanyu Shuiping Kaoshi (HSK) (gaijinban) tongguo zhuanjia shending 中国汉语水平考试(HSK)[改进版]通过专家审定.“ *Shijie hanyu jiaoxue*, č. 1, s. 140.
- CAI Qing a BRYSSBAERT, Mark (2010): „SUBTLEX-CH: Chinese Word and Character Frequencies Based on Film Subtitles“. *PLoS ONE* [online]. roč. 5, č. 6, s. 1–8 [cit. 2015-10-25]. ISSN: 1932-6203. Dostupné z: <http://www.plosone.org/article/fetchObject.action?uri=info:doi/10.1371/journal.pone.0010729&representation=PDF>
- CARTER, Ronald (1982): „A Note on Core Vocabulary.“ *Nottingham Linguistic Circular*, roč. 11, č. 2, s. 39–50.
- CARTER, Ronald (2012): *Vocabulary: Applied Linguistic Perspectives*. 2. vyd. New York: Routledge, 305 s. ISBN 978-0-415-699334-1.
- ČERMÁK, František a BLATNÁ, Renata (1995): *Manuál lexikografie*. 1. vyd. Jinočany: H & H, 283 s. ISBN 80-857-8723-7.
- ČERMÁK, František et al. (2000): *Studie z korpusové lingvistiky*. 1. vyd. Praha: Karolinum. 531 s. ISBN 80-7184-893-X.
- Education Division of Taipei Economic and Cultural Office (T.E.C.O.) (©2011): T.O.C.F.L. Test Of Chinese as a Foreign Language. *Education Division TECO* [online] [cit. 2016-03-10]. Dostupné z: <http://www.tw.org/tocfl/>

- Embassy of the People's Republic of China in the Czech Republic (2014): J.E. Ma Keqing se zúčastnila slavnostního založení pražské pobočky Konfuciova institutu. *China Embassy* [online] [cit. 2016-04-08]. Dostupné z: <http://www.chinaembassy.cz/cze/rdzt/t1215063.htm>
- Fachverband Chinesisch (©2011–2016): Sprachprüfungen. *Fachverband Chinesisch e.V. (FaCh)* [online] [cit. 2015-12-10]. Dostupné z: <http://www.fachverband-chinesisch.de/chinesischindeutschland/pruefungen/index.html>
- Fachverband Chinesisch (2010): Statement of the Fachverband Chinesisch e.V. (Association of Chinese Teachers in German Speaking Countries) on the new HSK Chinese Proficiency Test. *Fachverband Chinesisch e.V.* [online]. 2010-6-1 [cit. 2016-02-16]. Dostupné z: http://www.fachverband-chinesisch.de/sites/default/files/FaCh2010_ErklaerungHSK.pdf
- FU King-wa a CHAU, Michael (2013): „Reality Check for the Chinese Microblog Space: A Random Sampling Approach“. *PLoS ONE* [online]. roč. 8, č. 3, s. 1–7 [cit. 2016-02-16]. ISSN: 1932-6203. Dostupné z: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0058356>
- GUDER, Andreas (2005): „Chinesisch und der Europäische Referenzrahmen“. *Chun* [online]. roč 20, č. 1, s. 63–78. [cit. 2016-02-16]. ISSN 0933-8381. Dostupné z: <http://www.fachverband-chinesisch.de/fachverbandchinesishev/fachzeitschriftchun/index.html>
- GUDER, Andreas a Fachverband Chinesisch (2015): „Europäischer Referenzrahmen für Chinesisch als Fremdsprache.“ *Chun* [online]. Zvláštní vydání, 207 s. [cit. 2016-04-16]. ISSN 0933-8381. Dostupné z: http://edocs.fu-berlin.de/docs/receive/FUDOCS_document_000000022875
- Guojia duiwai hanyu jiaoxue lingdao xiaozu bangongshi Hanyu Shuiping Kaoshi bu 国家对外汉语教学领导小组办公室汉语水平考试部 (1995): „Hanyu shuiping dengji biao zhun 汉语水平等级标准.“ *Hanyu wenzi yingyong* 语言文字应用, č. 4, s. 85–91.
- HAN Songtao 韩松涛 a CHENG Runshu 成润淑 (2012): „Jiantan xin jiu HSK zhi bijiao 浅谈新旧 HSK 之比较.“ *Jiaoyu chuangxin* 教育创新, č. 8, s. 31–34.
- Hankao Guoji 汉考国际 (2009a): „Xin Hanyu Shuiping Kaoshi (HSK) dagang“ qianyan 《新汉语水平考试 (HSK) 大纲》前言. *Chinesetest* [online].

2009-11-12 [cit. 2016-02-16]. Dostupné z:

<http://www.chinesetest.cn/gonewcontent.do?id=5589337>

- Hankao Guoji 汉考国际 (2009b): Guanyu shishi xin Hanyu Kaoshi de tongzhi 关于实施新汉语考试的通知. *Chinesetest* [online]. 2009-10-14 [cit. 2016-02-16].
Dostupné z: <http://www.chinesetest.cn/gonewcontent.do?id=677486>
- Hankao Guoji 汉考国际 (2011): Guanyu xin, jiu HSK fenshu duiying guanxi de shuoming 关于新、旧 HSK 分数对应关系的说明. *Chinesetest* [online]. 2009-10-14 [cit. 2016-02-16]. Dostupné z:
<http://www.chinesetest.cn/gonewcontent.do?id=2317021>
- Hankao Guoji 汉考国际 (2012b): Xin Hanyu Shuiping Kaoshi (HSK) haiwai shishi baogao 新汉语水平考试 (HSK) 海外实施报告. *Chinesetest* [online]. 2012-11-12 [cit. 2016-02-16]. Dostupné z:
<http://www.chinesetest.cn/gonewcontent.do?id=5589387>
- Hankao Guoji 汉考国际 (2012c): Hanyu Shuiping Kaoshi (HSK) gaige shexiang 汉语水平考试 (HSK) 改革设想. *Chinesetest* [online]. 2012-11-12 [cit. 2016-02-16].
Dostupné z: <http://www.chinesetest.cn/gonewcontent.do?id=5588368>
- Hankao Guoji 汉考国际 (2012d): Jieke chengong juban shouci xin HSK kaoshi 捷克成功举办首次新 HSK 考试. *Chinesetest* [online]. 2012-05-24 [cit. 2016-02-24].
Dostupné z: <http://www.chinesetest.cn/gonewcontent.do?id=4555990>
- Hankao Guoji 汉考国际 (2013): Dui di er yuyan kaoshi fenshu jieshi wenti de biyao yanjiu 对第二语言考试分数解释问题的比较研究. *Chinesetest.cn* [online]. 2013-12-26 [cit. 2016-03-16]. Dostupné z:
<http://www.chinesetest.cn/gonewcontent.do?id=8794924>
- Hankao Guoji 汉考国际 (©2013–2016a): *Chinesetest* [online]. Xin HSK cihui 2012 nian xiudingban 新 HSK 词汇 2012 年修订版 2012-11-12 [cit. 2016-03-06].
Dostupné z: <http://www.chinesetest.cn/userfiles/file/HSK/HSK-2012.xls>
- Hankao Guoji 汉考国际 (©2013–2016b): Xin Hanyu Shuiping Kaoshi HSK 新汉语水平考试 HSK. *Chinesetest* [online]. 2012-11-12 [cit. 2016-02-06]. Dostupné z:
<http://www.chinesetest.cn/gosign.do?id=1&lid=0#>

- Hankao Guoji 汉考国际 (©2013–2016c): Hanyu Shuiping Kouyu Kaoshi HSKK 汉语水平口语考试 HSKK. *Chinesetest* [online]. 2012-11-12 [cit. 2016-03-08].
Dostupné z: <http://www.chinesetest.cn/gosign.do?id=1&lid=1#>
- He Ran (2010): *Textbook Analysis: Vocabulary Instruction of Chinese as a Foreign Language*. San Diego: San Diego State University, 40 s.
- CHANG Jingyu 常敬宇 (2003): „Hanyu cihui de wangluoxing yu duiwai hanyu cihui jiaoxue 汉语词汇的网络性与对外汉语词汇教学.“ *Jinan daxue huawen xueyuan xuebao* 暨南大学华文学院学报, č. 3, s. 13–18.
- CHEN Manhua 陈满华 (1993): „Youguan Hanyu Shuiping Kaoshi de ji dian yijian 有关汉语水平考试的几点意见.“ *Yuyan jiaoxue yu yanjiu* 语言教学与研究, č. 2, s. 133–142.
- JIANG Dewu 姜德梧 (1999): „Zhongguo Hanyu Shuiping Kaoshi (jichu) de sheji yuanze he shijuan jigou 中国汉语水平考试(基础)的设计原则和试卷结构.“ *Shijie hanyu jiaoxue* 世界汉语教学, č. 3, s. 30–36.
- KANE, Daniel (2009): *Knížka o čínštině*. 1. vyd. Mirošovice: DesertRose, 207 s. ISBN 978-80-903296-1-4.
- KANG Yan 康燕 (2011): „Tan duiwai hanyu jiaoxue zhong de cihui jiaoxue 谈对外汉语教学中的词汇教学.“ *Yuwen xuekan* 话文学刊, č. 6, s. 157–159.
- Katedra asijských studií (nedatováno): Testy HSK konané na KAS [online]. *Katedra asijských studií* [cit. 2016-04-08]. Dostupné z: http://oldkas.upol.cz/testy_hsk.php
- Konfuciova akademie (©2007–2016): O akademii. *Konfuciova akademie 孔子学院* [cit. 2016-04-08]. Dostupné z: <http://konfucius.upol.cz/old//o-akademii/>
- Konfuciova akademie Univerzity Palackého v Olomouci (©2007–2016a). Zkouška HSK a HSKK. *Konfuciův institut Univerzita Palackého v Olomouci* [online] [cit. 2016-04-08]. Dostupné z: <http://konfucius.upol.cz/cinstina/zkouska-hsk-a-hskk/>
- Konfuciova akademie Univerzity Palackého v Olomouci (©2007–2016b). Stipendia. *Konfuciův institut Univerzita Palackého v Olomouci* [online] [cit. 2016-04-08]. Dostupné z: <http://konfucius.upol.cz/cinstina/stipendia/>
- KRAMES, Marion (2013): *Leseverstehen und lexikalische Kompetenz in Chinesisch als Fremdsprache (ChaF): Auswertung psycholinguistischer Untersuchungen im Hinblick auf die Lese - und Wortschatzdidaktik in ChaF*. Trier: Universität Trier. 288 s.

- KRAMES, Marion (2015): „Wortschatz, Worterkennung und Leseverstehen in Chinesisch als Fremdsprache (ChaF)“. *Chun* [online]. roč 30, č. 1, s. 5–29. [cit. 2016-02-16]. ISSN 0933-8381. Dostupné z: <http://www.fachverband-chinesisch.de/fachverbandchinesischev/fachzeitschriftchun/index.html>
- LAUFER, Batia (1989): „A Factor of Difficulty in Vocabulary Learning: Deceptive Transparency.“ In NATION, Paul a CARTER, Ron (ed.) *Vocabulary Acquisition*. Basel: AILA, s. 10–20.
- LEWIS, Michael (2005): „Towards a Lexical View of Language – a Challenge for Teachers.“ *Babylonia*, č. 3, s. 7–10. ISSN: 1420-1658.
- LI Guimei 李桂梅 (2009): „HSK (gaijinban) (zhongji) de gouxiang xiaodu yanjiu HSK(改进版)[中级]的构想效度研究.“ *Kaoshi yanjiu* 考试研究, č. 1, s. 94–102.
- LI Rulong 李如龙 a YANG Jichun 杨吉春 (2004): „Duiwai hanyu jiaoxuezhong de cihui jiaoxue wei zhongxin 对外汉语教学中的词汇教学为中心.“ *Jinan daxue huawen xueyuan xuebao* 暨南大学华文学院学报, č. 4, s. 27–49.
- LI Xiangping 李香平 (2004): „Zitiwei lilun de zhongyao fazhan 字本位理论的重要发展.“ *Hanyu xuexi* 汉语学习, č. 2, s. 75–77.
- LIU Jie 刘杰 (ed.) (2000): *HSK Zhongguo Hanyu Shuiping Kaoshi yingshi zhinan (jichu) HSK 中国汉语水平考试应试指南 (基础)*. Beijing: Culture and Language University Press. 136 s. ISBN 7-5619-0887-3.
- LIU Lianli 刘镰力 (1999): „Zhongguo Hanyu Shuiping Kaoshi (HSK) de dengji tizhi 中国汉语水平考试(HSK)的等级体制.“ *Shijie hanyu jiaoxue* 世界汉语教学, č. 3, s. 17–23.
- LIU Yinglin 刘英林 (1989): „Hanyu Shuiping Kaoshi (HSK) de jiben moshi 汉语水平考试(HSK)的基本模式.“ *Yuyan jiaoxue yu yanjiu* 语言教学与研究, č. 1, s. 84–94.
- LIU Yinglin 刘英林 (ed.) (1998): *Hanyu shuiping dengji biaoqun yu yufa dengji dagang 汉语水平等级标准与语法等级大纲*. Beijing: Gaodeng jiaoyu chubanshe. 180 s. ISBN: 7-04-005845-6.
- LIU Yuting 刘瑀婷 (2014): „Jiantan wei shidai weibo wenxue fazhan 浅谈微时代微博文学发展.“ *Beifang wenxue* 北方文学, č. 1, s. 217–218.

- LOPEZ, Hugo (2013): *A mobile application for a cost/benefice approach of CFL vocabulary acquisition*. Paris: National Institute of Oriental Languages and Civilizations. s. 1–5.
- MCBRIDE-CHANG, Catherine et al (2006): „Morphological Structure Awareness, Vocabulary, and Reading“ In WAGNER, Richard (ed.) *Vocabulary Acquisition: Implications for Reading Comprehension*. 1. vyd. New York: The Guilford Press, s. 104–123. ISBN: 978-1593853389.
- McENERY, Tony a XIAO, Richard (2004): „The Lancaster corpus of Mandarin Chinese: a corpus for monolingual and contrastive language study“. In *Proceedings of LREC*. Lisbon: ELRA, s. 1175–1178. ISBN: 2-95117508-0.
- MENG Xiangchao 孟祥超 (2011): „Hanyu Tuofu“ kaoshi jiang tongyi wei xin HSK “汉语托福”考试将统一为新. *CNTV* [online]. 2011-1-20 [cit. 2016-02-16]. Dostupné z: <http://news.cntv.cn/china/20110120/100403.shtml>
- MEYER, Florian (2009): „Was ist Chinesischkompetenz? Standardisierte Chinesischprüfungen und Untersuchungen zur Sprachstandmessung“. *Chun* [online]. roč 24, č. 1, s. 21–45. [cit. 2016-02-16]. ISSN 0933-8381. Dostupné z: http://www.fachverband-chinesisch.de/sites/default/files/chun/CHUN_2009_Was_ist_Chinesischkompetenz.pdf
- MEYER, Florian (2014): *Language Proficiency Testing for Chinese as a Foreign Language: an Argument-based Approach for Validating the Hanyu Shuiping Kaoshi (HSK)*. 1. vyd. Frankfurt am Main: Peter Lang. 349 s. ISBN 36-316-4891-X.
- MILTON, James (2009): *Measuring Second Language Vocabulary Acquisition*. 1. vyd. Bristol: Multilingual Matters. 328 s. ISBN 978-1847692078.
- NATION, I. S. P. (2001): *Learning Vocabulary in Another Language*. 1. vyd. Cambridge: Cambridge University Press. 447 s. ISBN: 0-521-800927.
- NI Mingliang 倪明亮 (ed.) (2000): *Zhongguo Hanyu Shuiping Kaoshi yingshi zhinan* 中国汉语水平考试影视指南. Beijing: Beijing Culture and Language University Press. 424 s. ISBN 7-5619-0560-2.
- PENG Fengju 彭凤菊 (1998): „HSK cunzai de wenti ji zai wo qu hanyu jiaoxue zhong de yingyong celue HSK 存在的问题及在我区汉语教学中的应用策略.“ *Yuyan yu fanyi* 语言与翻译, č. 1 s. 58–59.

- Rada Evropy (2006): *Společný evropský referenční rámec pro jazyky: jak se učíme jazykům, jak je vyučujeme a jak v jazycích hodnotíme*. 2. vyd. Olomouc: Univerzita Palackého, 2006. 267 s. ISBN 80-244-1425-2.
- Rada Evropy (©2015): Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). *Council of Europe* [online]. [cit. 2016-01-26]. Dostupné z:
http://www.coe.int/t/dg4/linguistic/Source/Framework_en.pdf#page=4
- RAMIREZ, Amelia (2012): „The lexical approach: collocability, fluency and implications for teaching.“ *Revista de Lenguas para Fines Específicos*, č. 18, s. 237–254. ISSN: 1133-1127.
- REN Xiaomeng 任筱萌 (2001): „Zhongguo Hanyu Shuiping Kaoshi HSK de huigu xianzai yu fazhan 中国汉语水平考试 HSK 的回顾现状与展望.“ *Chinese Language Learning*, č. 2, s. 63–66.
- SONG Fuqian 宋赴前 (1997): „Hanyu Shuiping Kaoshi HSK de qishi 汉语水平考试 H S K 的启示.“ *Yuyan yu fanyi 语言与翻译*, č. 1, s. 69–70.
- STUBBS, Michael (1986): „Language Development, Lexical Competence and Nuclear Vocabulary“ In DURKIN, Kevin (ed.) *Language Development in the School Years*. London: Croom Helm, s. 57–76.
- SUN Dejin 孙德金 (2007): „Hanyu Shuiping Kaoshi (HSK) de kexue benzhi 汉语水平考试 (HSK) 的科学本质.“ *Shijie hanyu jiaoxue 世界汉语教学*, č. 4, s. 129–138.
- SUN Ying 孙莹 (2008): „Jiantan Cihui xuexi delüe 浅谈词汇学习策略.“ *Wanfang Shuju 万方数据*, č. 1, s. 172–176.
- SUN Zhongqiang 孙中强 (2011): „Guanyu “zitiwei” lilun de ji dian sikao 关于“字本位”理论的几点思考.“ *Yuyan Wenzhi Xiuci 语言文字修辞*, č. 4, s. 14–17.
- TIAN Shuhui 田书慧 (2014): *Dui Weibo xin cihui de yanjiu 对微博新词汇的研究*. *Renmin Wang* [online]. s. 1–2, 2014-08-26 [cit. 2016-02-16]. Dostupné z:
<http://media.people.com.cn/n/2014/0826/c388272-25541239.html>
- Unité des études chinoises (2014): Listes HSK. *Université de Genève* [online]. [cit. 2016-03-16]. Dostupné z:
<http://www.unige.ch/lettres/estas/chinois/general/HSK.html>

- WANG Chaozhong 王朝中 (1993): „Hanyu Shuiping Kaoshi (HSK) zonghe pingjia 汉语水平考试(HSK) 综合评介.“ *Xinjiang shifan daxue xuebao – zhe, she* 新疆师范大学学报 (哲社), č. 2, s. 74–77.
- WANG Qi 王琦 (2008): „Dui gaijinban Zhongguo Hanyu Shuiping Kaoshi (HSK) ji dian xiangfa 对改进版中国汉语水平考试(HSK) 几点想法.“ *Kaoshi yanjiu* 考试研究, č. 4, s. 8.
- XIN Juan 辛娟 (2013): „Lüetan ziti benwei yu duiwai hanyu cihui jiaoxue 略谈字本位与对外汉语词汇教学.“ *Beifang wenxue* 北方文学, č. 9, s. 252.
- XU Lin 徐琳 (ed.) (2009): *Xin Hanyu Shuiping Kaoshi dagang HSK si ji* 新汉语水平考试大纲 HSK 四级. Beijing: The Commercial Press. 54 s. ISBN 978-7-100-06887-1.
- ZHANG Jie 张洁 a ZHANG Jinjun 张晋军 (2010): „Hanyu Shuiping Kaoshi yongci tongji fenxi baogao 汉语水平考试用词统计分析报告.“ *Zhongguo kaoshi* 中国考试, č. 3, s. 34–38.
- ZHANG, George (2012): „Concept of Partial Competence and its Implications for Chinese Language Teaching“ In XING Minjie a LI Mingfang (ed.) *Applied Chinese Language Studies*. London: Sinolingua, s. 1–10. ISBN: 978-1-907838-36-1.
- ZHOU Ziheng 周子衡 (2012): „Xin Hanyu Shuiping Kaoshi dagang cihui yu kaoshi zhenti cihui de duibi fenxi 新汉语水平考试大纲词汇与考试真题词汇的对比分析.“ *Kaoshi zhoukan* 考试周刊, č. 57, s. 1–3.

Seznam příloh

- CD-ROM s korpusem lexémů slovní zásoby HSK a jejich frekvenčními údaji