

OBSAH

I. ČÁST: TEORIE ----- 9

1	ÚVOD: KORPUS A KORPUSOVÁ LINGVISTIKA ----- 10
2	DATA A INFORMACE ----- 12
2.1	Elektronický text ----- 13
2.1.1	Obecná povaha, vlastnosti a zpracování textů ----- 13
2.2	Korpusová data ----- 16
2.2.1	Povaha a zdroje elektronických dat ----- 17
2.2.2	Recepce a produkce textu ----- 19
2.2.3	Reprezentativnost korpusových dat jako zrcadlo reality ----- 21
2.3	Korpusový kontext ----- 25
2.4	Korpus a empirie ----- 27
2.5	Technické a další aspekty přípravy korpusu ----- 28
2.5.1	Příprava dat pro korpus ----- 29
2.5.2	Dotazy a dotazovací jazyk ----- 34
3	KORPUS A KORPUSY ----- 37
3.1	Korpus, jeho povaha a vytváření ----- 37
3.1.1	Korpus a jazyk ----- 37
3.1.2	Tvorba korpusu a jeho možnosti ----- 39
3.1.3	Hledání v korpusu a jeho analýza ----- 41
3.1.3.1	Možnosti hledání a analýzy ----- 41
3.1.3.2	Souhrnný pohled na úzus slova v kontextu: Konkordance ----- 51
3.1.3.3	Kombinace textové i systémové: kolokace a koligace ----- 55
3.1.3.4	Postup od formy k významu a funkci. Typický a zvláštní ----- 62
3.1.4	Hlavní aspekty a charakteristiky korpusu ----- 67
3.2	Český národní korpus (ČNK) ----- 69
3.3	Typy korpusů ----- 74
3.3.1	Synchronní korpus ----- 76
3.3.2	Mluvený korpus ----- 78
3.3.3	Diachronní korpus ----- 80
3.3.4	InterCorp ----- 82
3.3.5	Webové korpusy ----- 83
3.4	Hlavní světové korpusy ----- 84

4	KORPUSOVÁ LINGVISTIKA ----- 90
4.1	Korpusová lingvistika jako disciplína ----- 90
4.2	Korpusová metodologie ----- 92
4.3	Korpusová statistika ----- 100
4.3.1	Základní pojmy a operace ----- 101
4.3.2	Frekvence ----- 102
4.3.3	Zipfův zákon ----- 106
5	KORPUS VE VÝZKUMU A APLIKACI ----- 109
5.1	Výzkum spojený s budováním korpusu ----- 110
5.2	Lingvistický výzkum ----- 111
5.2.1	Výzkum obecně ----- 111
5.2.2	Variabilita jazyka. Mluvená povaha jazyka ----- 112
5.2.3	Lexikon a frazeologie. Kolokace ----- 114
5.2.4	Morfologie a gramatika. Pedagogické aplikace ----- 117
5.3	Lingvistický kontrastivní výzkum více jazyků (InterCorp) ----- 118
6	ZÁVĚR: PERSPEKTIVY KORPUSOVÉ LINGVISTIKY ----- 120

II. ČÁST: SONDY, STUDIE, ANALÝZY ----- 123

A	Korpusy včera, dnes a zítra (2011) ----- 125
B	Some of Current Problems of Corpus and Computational Linguistics or Fifteen Commandments and General Truths (2007) ----- 145
C	Spoken Corpora Design: Their Constitutive Parametres (2009) ----- 156
D	InterCorp: jeho povaha a možnosti (2014) ----- 165
E	Lexical Collocability: The Case of Verbs and Adverbs (2010) ----- 182
F	Abstract Nouns Collocations: Their Nature in a Parallel English-Czech Corpus (2005) ----- 196
G	Statistical Methods for Searching Idioms in Text Corpora (2006) ----- 207
H	Povaha a úzus interjekcí: případ češtiny (2007) ----- 216
I	Konference: Staré známé nebo neznámé slovo? (2001) ----- 225

III. ČÁST: PŘÍLOHY ----- 229

Příloha A	Jazyková analýza výběrové konkordance (linout se) ----- 231
Příloha B	Frekvence slov ve Frekvenčním slovníku (prvních 100 lexémů) ----- 234
Příloha C	Jeden pohled na českou morfologii (Procentuální rozložení rodu, čísla a pádu u substantiv) ----- 237
Příloha D	Román jako korpus: Zbabělci (Josef Škvorecký) ----- 239

KORPUSOVÁ BIBLIOGRAFIE ----- 247

GLOSÁŘ ZÁKLADNÍCH POJMŮ A TERMÍNŮ ----- 261