

# Obsah

---

<b>1 Úvod</b>	8
<b>2 Korpus a jazyková realita</b>	11
2.1 Korpus jako zdroj dat	11
2.2 Reprezentativnost korpusů	12
2.3 Jazyk a jeho variety	14
2.4 Kritéria pro kategorizaci textů v korpusech	16
2.5 Druhy korpusů a homogenita	18
2.6 Velikost korpusu a volba vzorku	20
2.7 Tradiční jazykový korpus a internet	21
<b>3 Srovnávání korpusů</b>	23
3.1 Synchronní a diachronní srovnání	23
3.2 Reprezentativnost diachronních korpusů	24
3.3 Diachronní srovnávací studie	26
3.3.1 Druhy diachronních srovnávacích studií	26
3.3.2 Studie založené na diachronních korpusech	27
3.3.3 Studie založené na synchronních korpusech	28
3.3.4 Studie založené na brownovských korpusech	30
3.3.5 Studie založené na monitorovacích korpusech	31
3.4 Vliv složení korpusu na interpretaci výsledků	34
<b>4 Popis zdrojových dat</b>	36
4.1 Korpusy řady SYN	36
4.2 Technické aspekty výstavby korpusů řady SYN	37
4.2.1 Terminologický úvod	37
4.2.2 Akvizice textů	38
4.2.3 Konverze do meziformátu	39
4.2.4 Anotace	39
4.2.5 Dočištění a zařazení do banky	41
4.2.6 Výběr textů z banky	42
4.2.7 Lemmatizace a morfologické značkování	44
4.3 Reprezentativnost korpusů řady SYN	46

4.3.1	Kategorizace textů	46
4.3.2	Vymezení synchronie	51
4.3.3	Shrnutí	53
4.4	Hlavní rysy korpusu SYN	54
4.5	Subkorpusey a způsob jejich výběru	55
<b>5</b>	<b>Popis použité metody</b>	<b>65</b>
5.1	Úvod	65
5.2	Předchozí práce	66
5.2.1	Povaha jazyka a frekvenčních rozdílů	66
5.2.2	Rovnoměrnost rozložení výskytů	70
5.3	Vylepšení původních metod	72
5.3.1	Srovnávání po hlavních typech textu	72
5.3.2	Srovnávání uvnitř publicistiky po jednotlivých letech	76
5.3.3	Shrnutí	79
5.4	Použitá metoda	80
5.4.1	Úroveň lexikální	80
5.4.2	Úroveň lexikálních kombinací	82
<b>6</b>	<b>Výsledky a diskuse</b>	<b>85</b>
6.1	Úvod	85
6.2	Iterativní <i>cbf</i> , <i>chi</i> , <i>ll</i> na reprezentativních subkorpusech	87
6.3	Iterativní <i>cbf</i> , <i>chi</i> , <i>ll</i> na publicistických subkorpusech	100
6.4	<i>tau</i> na publicistických subkorpusech	120
6.5	<i>taumed</i> na publicistických subkorpusech	139
6.6	Shrnutí	162
<b>7</b>	<b>Závěr</b>	<b>170</b>
	<b>Použité korpusy</b>	<b>173</b>
	<b>Literatura</b>	<b>174</b>