Language	Source	#sen	#words	#P	#PM	OOV
Bulgarian	BulTreeBank-Morph	20,558	321,538	—	564	10.07
Czech	PDT Treebank 2.5	115,844	1,957,246	—	1,570	6.09
Dutch	Lassy Small Corpus	65,200	1,096,177	—	933	7.21
French	French Treebank	21,562	587,687	17	306	5.19
German	TIGER Corpus	50,474	888,236	54	795	7.74
Hindi	Hindi Treebank	26,547	588,995	39	—	—
Italian	ISDT Treebank	10,206	190,310	70	—	11.57
Portuguese	Tycho Brahe Corpus	68,859	1,482,872	—	344	4.39
Spanish	IULA LSP Treebank	42,099	589,542	—	241	4.94
Swedish	Stockholm—Umeå Corpus 3.0	74,245	1,166,593	—	153	8.76
Thai	ORCHID Corpus	23,225	344,038	47	—	5.75
Vietnamese	(VTB) Vietnamese Treebank	10,293	220,574	22	—	3.41
	(VLSP) VLSP Evaluation Campaign 2013	28,232	631,783	31	—	2.06

(Out-of-Vocabulary): the average percentage of unknown word tokens in each test fold. For Hindi, OOV rate is 0.0