

language	dict	train	dev	devtest	test
Bengali	16k	539k	63k	61k	69k
	6k	20k	914	907	1k
Hindi	0	1,249k	67k	98k	74k
	0	37k	1k	993	1k
Malayalam	410k	664k	61k	68k	70k
	144k	29k	1k	1k	1k
Tamil	189k	747k	62k	53k	54k
	69k	35k	1k	1k	1k
Telugu	106k	951k	52k	45k	49k
	38k	43k	1k	916	1k
Urdu	253k	1,198k	67k	49k	42k
	113k	33k	736	777	605