

Dataset	#Pages	#Words
English-1601	310	1,13,008
DLI Hindi (HS1)	1,533	4,20,100
DLI Telugu (TS1)	1,005	1,61,265