		million words in corpus	million words for training		
			Brown	Clark	Word2vec
en	Reuters RCV1	63	35	63	63
	Wikipedia (20141208)	1700	790	790	1700
	Gigaword 5th ed.	4000	-	-	4000
de	Wikipedia (20140725)	650	190	190	650
	deWac	1100	500	500	1100
es	Wikipedia (20140810)	428	246	246	428
	elperiodico (1998-2002)	60	35	60	60
	Gigaword 3rd ed.	1150	330 (afp)	330 (afp)	1150
nl	Wikipedia (20140804)	235	128	128	235
eu	Wikipedia (20141208)	60	12	60	60
	Egunkaria (1999-2003)	38	28	38	38
	Berria (2003-2014)	90	78	90	90