Corpus	Vocabulary Size	Word Similarity	
		Original	Our
17M-word (text8)	71,291	63.4	66.5
1B-word benchmark	1,115,011	64.0	64.1
7.2B-word collection	1,115,011	70.0	69.8