Language	Train	Test	WordVectors
Turkish	MC-2010 (617K)	MC-05:10 (2534)	Vectors-Gencor (361M)
English	MC-2010 (878K)	MC-05:10 (2218)	Wikipedia-Normalized (129M)
German	MC-2010 (2M)	MC-2010 (785)	Manually collected (651M)