How to download the Leipzig corpus file for the paper.







Unzip the folder and put “ind\_newscrawl\_2012\_1M-sentences.txt” in the same folder with the R Markdown document and the other datasets.