Homework 1

Course: Introduction to Corpus Linguistics

Due date: November 3rd 2023

Relevant readings: Barth & Schnell 2022: Ch.1-2

Total: 12 points

1. Register for the Corpus of Contemporary American English (COCA). Search for the word *teacher's assistant*. You should receieve an error message. What does the error message tell you about the way 'clitics' are typically tagged in COCA. Discuss your answer in relation to the difference between word forms and lexemes. (length: ~paragraph) (3 points)

2. Use the COCA database to find the most common directly adjacent **collocates** for 'protestor' according to the following parts of speech: noun.ALL, verb.ALL, adj.ALL. (3 points)

3. Barth & Schnell (2022: 32-40) list and describe different types of corpora. Choose two types of corpora to compare and discuss. First provide a brief description of the corpora types (paraphrasing but not plagiarizing the descriptions in Barth and Schnell). Give two examples of research questions that could be asked with one corpus and not the other (~length: two paragraphs) (6 points)

   Send in your homework online as a pdf document.