Note: this contains collaboration, reflection paragraph and reference for each source

Name	Student id	Section contributed	Selected edited	Other contributions
Justin Achadinha	301438505	Corpus collection, supcorpora creation	code	Did 2 &3 subcorpora, -help each other with code
Anandjot Hothi	301499039	Corpus collection, Supcorpora creation, write up	Code and write up	Did 1 subcorpora, and write up -help each other with code

Reflection:

For this assignment, we looked at three genres: an essay, a speech, and part of a novel. All three of these had a range of lexical diversity, different longest words, and unique long sentences. Lexical diversity is viewed as a measure of how many different words appear in a text. A higher number for lexical diversity, indicates that the language used in the text has a richer vocabulary. In the first subcorpus, the essay has a low lexical diversity of 0.002, meaning there is a lack of abundant words used. When looking at the most frequent words, it is filled with filler words like "the," "and," "of," and "to". This is not mean the essay is not well done, however, the language used is much simpler, unlike reading a medical essay. The longest words in this essay were: management, environment, and development, etc, which goes with the essay topic on project management. The second subcorpus was a speech, with moderate levels of lexical diversity of 0.16, which is not very surprising because it was speech. Speeches need to be clear and concise so the listeners can understand and not be confused about the wording. The speech text's are also filled with common filler words like, "the," "and," "of," and "to". However, the most interesting part about the speech is that the longest words like, "commercialization" and "democratically" show an undertone of political discussion taking place. It is a political speech, which is expected as speeches are often done by politicians. The novel subcorpus, has a relatively high lexical diversity, of 0.2, indicating a richer vocabulary. The most frequent words include "the," "a," and "of," which are expected in fiction and an still be considered filler words. Yet, the longest words are diverse, reflecting the narrative and descriptive nature of a novel.

The longest sentence in the essay appears to be a technical and detailed statement related to project management. In the speech, the longest sentence discusses various challenges and global issues. In the novel, the longest sentence describes a physical setting and details within a room. Overall, different subcorpus suggest that essays tend to have a more focused and specialized vocabulary. Speeches may cover a broader range of topics and are usually, political. Finally, novels offer a rich and descriptive narrative. The lexical diversity reflects the different linguistic characteristics of each genre.

Reference:

Speech :

AfricaNews, & GADDAFI, M. (2017, September 19). *[Speech] Muammar Gaddafi at the 64th UN General Assembly in 2009*. Africanews. https://www.africanews.com/2017/09/19/speech-muammar-gaddafi-at-the-64th-un-general-assembly-in-2009/

Novel:

Herbert, Frank. (2019, 06). *Whipping Star*. Read Novels Online - Novel24.Com.

https://novel24.com/whipping-star-consentiency-universe-1-s11292.html

Essay:

Reference StudyCorgi. (2023, March 7). Challenges of Project Management. Retrieved from

https://studycorgi.com/challenges-of-project-management/