19-07 – 22-07

* Bring the quotes back to the topical stuff

[wont do unless necessary to rerun topical]

* More detailed readme for the linguistic bit?
* Email Miguel about github
* Finish syntax – editing the correct files
* When syntax is done, run the WHOLE thing and read over the way it parses sentences. Make necessary changes.
* Character text generation: Try text generation again with higher epoch?

Now trying all text files together

* RUN TEXT GENERATION FOR ALL FILES
* BNC think of further analyses:

Instead of proportion of low freq verbs out of all text

Do low freq verbs out of verbs

🡪as have seen from POS distribution, the word-type proportions change

This could skew results?

found that there is a trend downwards, over the years, in proportion of verbs/adjectives of frequency lower than 50 [in BNC] out of total verbs/adjectives per speech.

* Also add average frequency per year

Write summary (maybe) about BNC results

* UPLOAD/UPDATE ALL EDITED FILSE TO GITHUB

Notes:

* Noticed how easy it is for results to be skewed / misinterpreted.

The importance of checking results and why they are there. Often issues with the text etc. (examples in text prep document)