June Summary.

* Began by following the metrics outline in Savoy paper

Didn’t anticipate significant changes, as Savoy compared between people 🡪 likely to find greater changes

MSL; Lexical diversity; lexical richness; big word usage

These metrics seem to have stayed roughly similar over the years

Not too surprising (same speaker)

Slight trend downwards of MSL (mean sentence length). Corroborates argument of language veering to less complex.

I then had to think of other analyses

* POS\_distribution
* Frequency relative to that of the general population

0-proportion / 10-proportion

Stayed similar apart from a dip in infrequent words around 70s.

This actually is when the corpus was created

Possible explanation is that the corpus represents a shot in time of language. Perhaps the queen’s language use reflected the time… therefore was more “frequent” during the time of the corpus creation…

Perhaps a corpus from a different year would show a dip too

(potential thing to look at next?)

Next steps:

* Starting to look into parsing of sentences
* Pronoun use
* Same topics / different expression
* Looking into the “prediction” of text