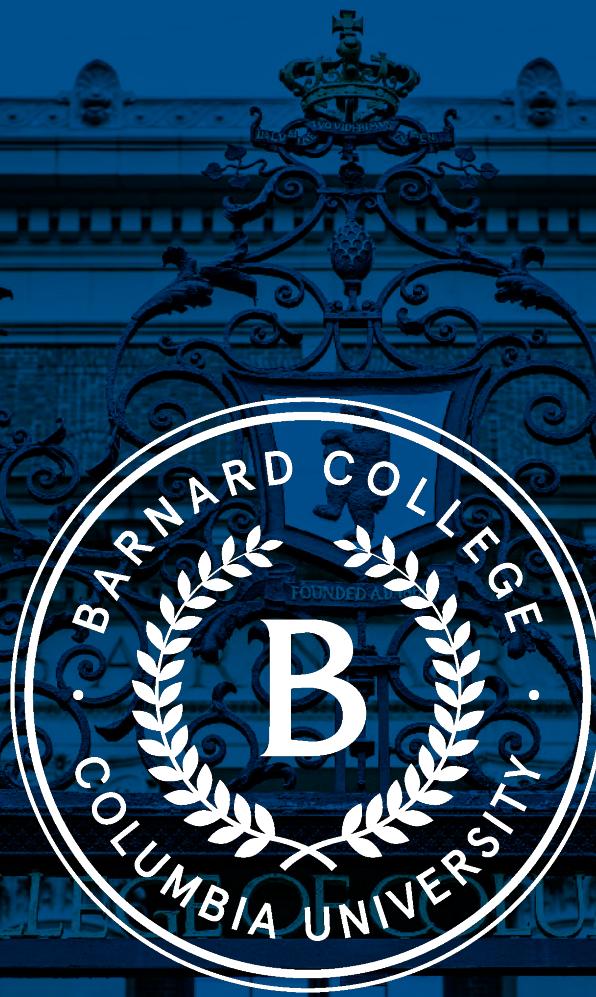


Bible, Babel, and Hermeneutics

Aaron McKeever

BC COMS 2710
Computational Text Analysis
Project Presentation





- Project Description
 - The relationship between biblical translations and hermeneutics, via computational approach to literary analysis.
- Importance
 - There is a lot of on machine translation and the ability for computer's ability to largely comprehend texts in a meaningful way, but I wanted to first be mindful of two things:
 - Variance among human-produced translation, analyzing them for specific axes of variance
 - Differences in tools ability to parse, analyze, etc. English vs. other languages
 - For machines to be able to generate and “understand” text it must first be able to parse it.

Research Question



Question:

- Do the translations vary significantly in their produced word-relation graph? If so, what are the main axes of variation among these translations of the Bible into different languages (inc. versions of same lang.)?

Possible Impacts:

- Christianity is one of the most powerful and widespread semiotic systems, and a larger scale analysis of the “gospel” may result in new analyses/interpretations.



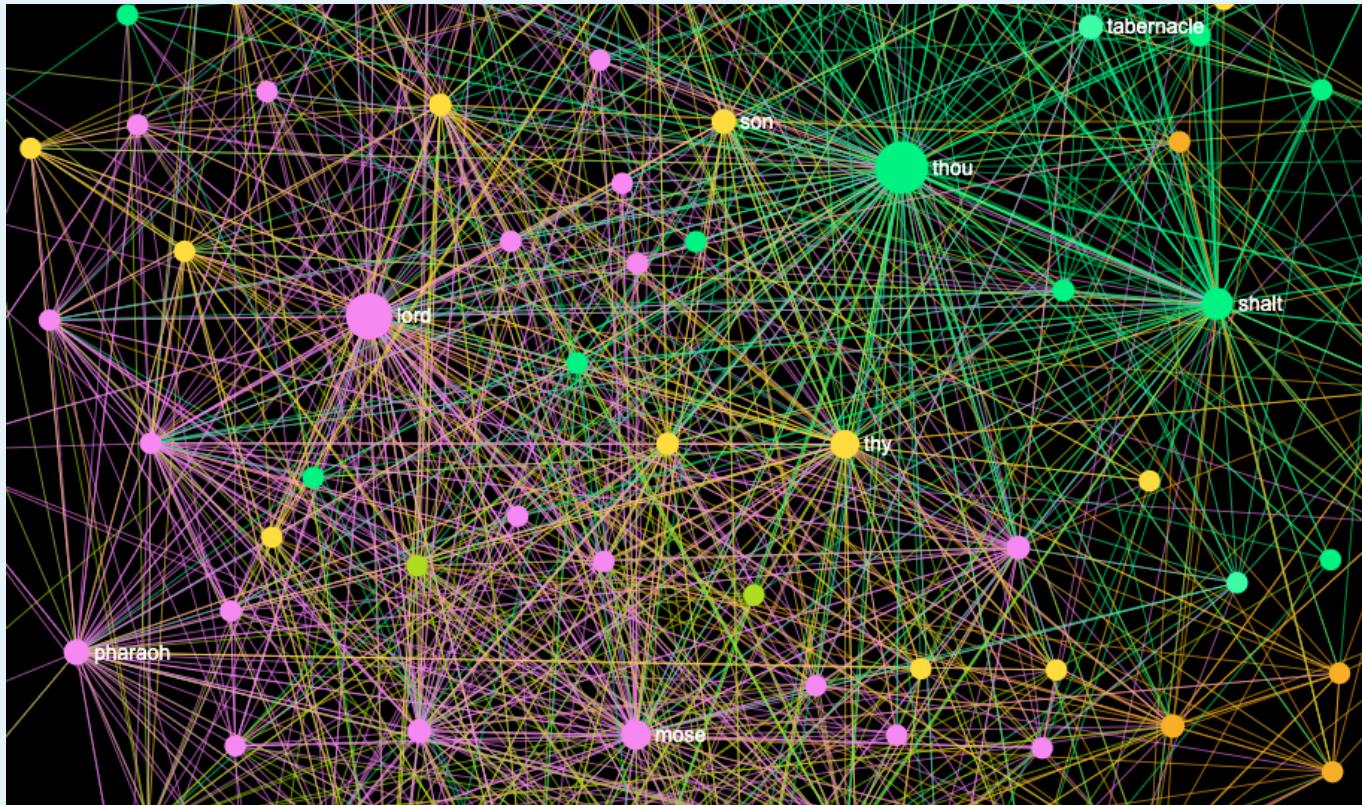
- Data Sourcing Methods
 - Web-scraping, pseudo-webscraping (manual copy-pasting into .txt docs), and pre-rendered text docs.
- Processing Methods
 - Cleaned (partially) with CoreNLP
 - Tokenizing, removing punctuations, etc.
 - Purposefully left stop words - informational value in thou vs ye, etc.
 - Lemmatization
 - CoreNLP – Text processing
 - InfraNodus (TBC in D3?)- Visualization
- Document
 - A single translation of Exodus
- Corpus statistics:
 - (So-far) 7 documents
 - ~ 30,000 words per document

Computational Text Analysis Method



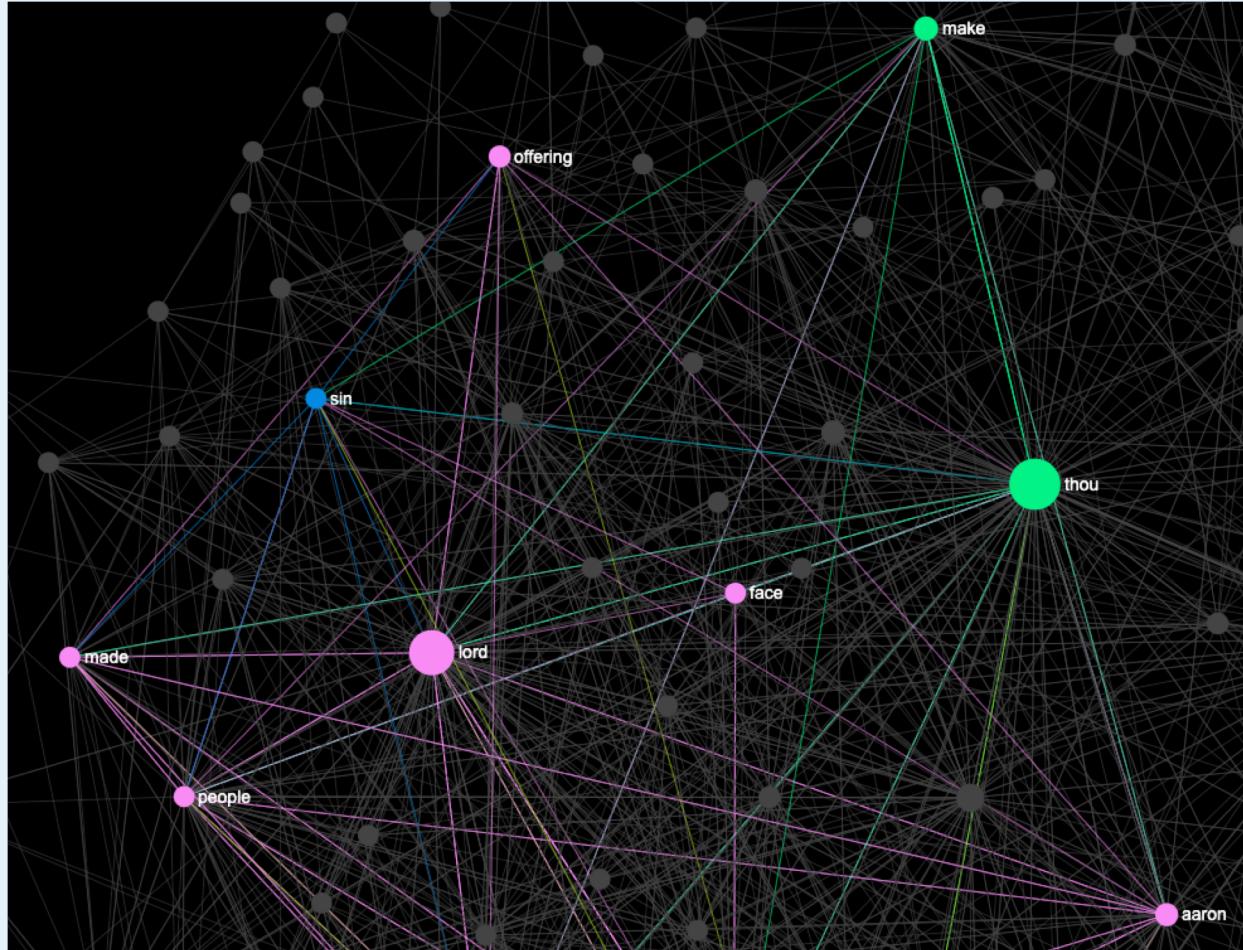
- Word Counts
 - TF-IDF won't work (due to multi-lingual nature)
 - Use of analyzing solely Exodus helps (via similar length)
 - WC, n-grams, etc.
- Visualizations/Word Graphs
 - InfraNodus - word network visualizations
 - Also, simple WC visualizations
- (Possible) Machine Translation
 - Does running a Spanish translation into English result in a text resembling anything biblical?

Results





Results





Conclusions

- Many words are “typical” translations
 - However, differ in conceptual relationship
 - Sin vs. ἀπόστασις vs Sünde
 - Word-gender relations
 - Etymologies
 - Cultural relations
- Meta-linguistic features
 - KJV vs. more contemporary translations
 - Power/ritualistic nature of archaic language
 - Relation to more immediate language vs. later translations