

Grace Jensen
Professor Ventura
PPOL 6801
23 Feb 2023

Replicating 'Democratization And Linguistic Complexity: The Effect of Franchise Extension on Parliamentary Discourse'

Background

In the research paper 'Democratization And Linguistic Complexity: The Effect of Franchise Extension on Parliamentary Discourse,' Arthur Spirling seeks to understand how political speeches adapt over time to new audiences by specifically evaluating how House of Commons speech's complexity changed after the Second Reform Act of 1868. The Second Reform Act granted the right to vote to lower class, and likely less educated, British citizens, and therefore greatly expanded the pool of people that British members of parliament had to campaign towards. Spirling hypothesized that this change would lead to a marked difference in how these MPs spoke, believing that more public facing cabinet members would especially reduce the linguistic complexity of their speeches to accommodate their new constituency.

Methods

In order to test this hypothesis, Spirling collected over 860,000 speeches from over 4,000 MPs from 1832 to 1915. This timeframe was chosen deliberately, seeing as it encompasses a substantial period of time both preceding and succeeding the passing of the second Reform Act. This large dataset was then filtered to only include speeches from members running for the conservative or liberal party and speeches from the first session after the 1892 general election, leaving over 670,000 speeches for analysis. In terms of preprocessing the textual data from these speeches, the researchers simply (1) lengthened honorifics, (2) split speeches into sentences using punctuation marks, (3) stripped whitespaces, and (4) dropped strictly numeric sentences.

In order to gauge the complexity of the text, Spirling calculated each speech's Flesch reading Ease statistic, or FRE, which takes into account the number of words, sentences, and syllables per speech. This allows for an insightful analysis, as documents with concise and simple language are deemed "easier" to read, while more verbose texts are considered more

challenging. The FRE score, ranging from 0 to 100, allows us to easily classify each document: with a score of 90 to 100 denoting very easy readability, 61 to 70 indicating standard ease, and 0 to 30 classified as very difficult to understand. The average score was then calculated for each session and cabinet position (cabinet and non-cabinet) in order to determine overall how these scores changed over time depending on the person's position and level of public visibility.

The results showed that both cabinet and non-cabinet MP's increased their FRE score over time, meaning their speech became less complex and easier to understand, however, cabinet members saw a much sharper increase immediately following the second Reform Act.

Differences

Recalculating FRE scores

For my replication, I utilized the data published by the researchers, which included word, sentence, and syllable count for each document as well as other descriptive variables like the session the speech originated from, the author's party affiliation, whether said author was a cabinet member, and the competitiveness of the author's district. I recalculated each document's FRE score by inputting the word count, sentence count, and syllable count variables into the formula seen below:

$$206.835 - 1.015 \left(\frac{\text{total number of words}}{\text{total number of sentences}} \right) - 84.6 \left(\frac{\text{total number of syllables}}{\text{total number of words}} \right).$$

Each document's new score can be found under the 'FRE_recal' variable. I then used this statistic to re-run the visuals and regression model code published by the researchers to see if there were any differences between our results. Ultimately, each of my outputs matched that of Spriling's. I had the same distribution of FRE scores across my corpus, with the majority of documents having a score of around 50 to 59, meaning they were moderately difficult to understand. Additionally, my corresponding figure 3 replication indicates that I also saw a significant increase in FRE scores around the 1860/1870s, especially for cabinet members, thus implying that the original hypothesis was correct.

In terms of my regression results, it is worth noting that the variable representing the combined influence of both Cabinet membership and the Reform Act also exhibited the most substantial and statistically significant impact on the re-calculated FRE score, indicating that cabinet members were likely to significantly simplify their speeches after this suffrage movement. The rest of our regression coefficients and p values using controls also matched exactly- with cabinet position, reform act, political affiliation, word count, and competitiveness each having the same effect on FRE score. The identical results indicate a high degree of consistency and reliability in the original findings. We can now safely conclude now that the effect of the Second Reform Act did in fact have a significant impact on cabinet members in terms of speech complexity.

Recalculating FRE using Sample Text

The only major difference between my own results and that of the original researchers happened when I recalculated the FRE score of the following sentence taken from table 3 using the `textstat_readability()` function from the ‘quanteda’ package:

"Asked the Under Secretary of State for the Colonies whether Her Majesty's Government have arrived at any practical decision with regard to the establishment of emigration bureaux for the purpose of promoting as promised early in the Session a systematic system of State-directed emigration?"

While the researchers calculated an FRE score of -3.18, my recalculated score was 1.36. After some investigation, I discovered that this discrepancy was due to the `textstat` function counting the hyphenated word “State-directed” as two separate words, thus leading to a slightly higher word count and readability score. This seemingly minute difference proves the importance of consistency when it comes to preprocessing, seeing as something as insignificant as one hyphen can affect the FRE score by 4 points.

Improving Code Readability

I did make subtle changes to the provided code in order to make my replication more readable and efficient. While the original researchers utilized base R, I decided to use ‘tidyverse’ and ‘ggplot’ packages for simplified and seamless data manipulation and visualization whenever possible. The decision to adopt these packages over base R was driven by several compelling

advantages. First, tidyverse's intuitive syntax allowed for a more readable workflow when calculating average FRE, word count, syllable count, etc. across different years and cabinet positions. Additionally, the 'ggplot' package allowed me to quickly build sophisticated and visualizations that can be easily adjusted in the future by other researchers. By leveraging these packages, I aimed to enhance the reproducibility of the analysis while also ensuring that future replications could be conducted seamlessly with readily understandable and concise code.

Autopsy

The published materials for this research paper made this replication very straightforward, seeing as the the researchers published their dataset along with the R code used to create each figure and table. In order to fully replicate this paper, the user only needs access to Harvard's Dataverse and the ability to copy and paste. This level of openness in terms of the replication materials employed allows for seamless replication of the study and robust checks on their methodology. However, the available dataset that was published by the researchers does not include the raw text data from the speeches, making it especially difficult to completely replicate their process from start to finish. Because you can't actually see the text itself, the user cannot apply the same preprocessing steps or recalculate word, sentence, or syllabus counts for each of these documents. Additionally, the provided dataset allows for little flexibility in terms of applying other complexity measures that could have been used for comparative studies.

Extension

Evaluating Different Preprocessing Steps

After seeing the effect that a hyphen had on FRE scores, I believe it's important to evaluate how different preprocessing steps would affect or even improve the FRE scores and subsequent classification of these documents. It would be interesting to see how a 'wordfish' model, like the one that was implemented by Spirling in the past in his paper "Text preprocessing for unsupervised learning: why it matters, when it misleads, and what to do about it" (Denny, M.J., & Spirling, 2018) would rate the effectiveness of different preprocessing steps that were not previously utilized, like removing stopwords, really frequent/infrequent words, etc.

Applying other complexity measures

If the raw data was collected, it would also be useful to explore how other complexity measures would classify these speeches, including implementing a strategy like Benoit, MInger, and Spriling in their Measuring and explaining political sophistication through textual complexity” (2017) which was specifically tailored towards political speeches-taking into account the language used commonality and overall syntactic structure.

Expanding Regression Model

If possible to collect data on the author’s demographic information, it would be interesting to add more descriptive variables as controls for the regression model. For example, researchers could evaluate if variables like age would have an effect on FRE score. Asking the question, would a younger MP be more likely to simplify their speech than their older peers? Are there more individual factors that contribute to a political actor's adaptability?

Comparing against other suffrage movements

Other comparative applications of this study can be evaluating how speech complexity changed after different suffrage movements. Specifically in the United States, we could evaluate how legislators' speeches changed after the passing of the 19th amendment in 1919 which granted women the right to vote or after the voting rights act of 1965 which helped eliminate major barriers to black voters.

References

- Benoit, Kenneth, Kevin Munger, and Arthur Spirling. "Measuring and explaining political sophistication through textual complexity." *American Journal of Political Science* 63.2 (2019): 491-508.
- Denny, Matthew J., and Arthur Spirling. "Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it." *Political Analysis* 26.2 (2018): 168-189.
- Spirling, Arthur, 2015, "Replication Materials for: 'Democratization and Linguistic Complexity: The Effect of Franchise Extension on Parliamentary Discourse, 1832--1915'", <https://doi.org/10.7910/DVN/DDQDMS>, Harvard Dataverse, V1