BACHELOR'S THESIS / ENGINEERING 2025

Islamophobia and Xenophobia in Donald Trump's Rhetoric: A Comparative Study Around the 2016 Presidential Election

Uzair Ramzan Parveen

Universitat Pompeu Fabra
Barcelona

Islamophobia and Xenophobia in Donald Trump's Rhetoric: A Comparative Study Around the 2016 Presidential Election

BACHELOR'S THESIS OF

Uzair Ramzan Parveen

Supervisor: Carlos Castillo Ocaranza

Mathematical Engineering in Data Science

Academic Year 2024-2025

Universitat Pompeu Fabra Barcelona
Escola d'Enginyeria

Acknowledgements

First and foremost, I would like to thank the Almighty Allah for His countless blessings. I am truly grateful for the opportunity to work on such an engaging and meaningful topic.

I am deeply thankful to my family, especially my parents, for their unwavering support, hard work, and sacrifices, which have played a crucial role in bringing me to where I am today. Their constant encouragement, love, and belief in me have been a source of strength, especially during challenging times.

Lastly, but by no means least, sincere gratitude is expressed to the supervisor, Carlos, for the guidance and insightful feedback provided throughout the project. A great deal was learned under this mentorship, resulting in growth that enriched both academic and personal development.

Abstract

This project analyzes how Donald Trump employed rhetoric targeting immigrants, particularly Muslims and Mexicans, to gather support during his 2016 presidential campaign. Two datasets of Trump's speeches before and after the election were compiled, covering public statements, interviews, press conferences, and addresses to international audiences.

The data was thoroughly cleaned and processed, then analyzed using Natural Language Processing techniques to identify patterns. Visualizations highlighted shifts in rhetorical strategies over time, and sentence framing was examined based on Entman's Four Frame Functions.

Findings reveal a clear transition from negative portrayals of immigrants, political opponents, and foreign groups in the pre-election period to more positive, policy-focused narratives emphasizing America's economic and social progress post-election.

This study contributes to understanding the strategic use of hostile rhetoric in political discourse and offers insights into the construction and evolution of divisive narratives in contemporary politics.

Resumen

Este proyecto analiza cómo Donald Trump utilizó una retórica dirigida a inmigrantes, especialmente musulmanes y mexicanos, para ganar apoyo durante su campaña presidencial de 2016. Se recopilaron y limpiaron dos conjuntos de datos con discursos de Trump antes y después de las elecciones, incluyendo declaraciones públicas, entrevistas y ruedas de prensa.

Mediante técnicas de procesamiento del lenguaje natural, se identificaron patrones en las estrategias retóricas y se examinaron las estructuras de las oraciones según las Cuatro Funciones de Entman. Las visualizaciones mostraron un cambio claro en el discurso: del enfoque negativo hacia inmigrantes, oposición política y grupos extranjeros en el periodo preelectoral, a narrativas más positivas centradas en políticas públicas y el progreso social y económico de Estados Unidos tras las elecciones.

El estudio aporta una comprensión sobre el uso estratégico de la retórica hostil en el discurso político y cómo se construyen y evolucionan narrativas divisivas en la política actual.

Resum

Aquest projecte analitza com Donald Trump va utilitzar una retòrica dirigida als immigrants, especialment musulmans i mexicans, per guanyar suport durant la seva campanya presidencial del 2016. Es van recopilar i netejar dos conjunts de dades amb discursos de Trump abans i després de les eleccions, que inclouen declaracions públiques, entrevistes i rodes de premsa.

Mitjançant tècniques de processament del llenguatge natural, es van identificar patrons en les estratègies retòriques i es va examinar l'estructuració de frases segons les quatre funcions d'enquadrament d'Entman. Les visualitzacions van mostrar un canvi clar en el discurs: d'un enfocament negatiu envers immigrants, oposició política i grups estrangers en el període preelectoral, a narratives més positives centrades en polítiques i el progrés social i econòmic dels Estats Units després de les eleccions.

L'estudi aporta una millor comprensió de l'ús estratègic de la retòrica hostil en el discurs polític i de la construcció i evolució de narratives divisives a la política actual.

Table of Contents

List of Figures

List of Tables

Chapter 1

1. Introduction

This section provides the background of the project, presenting the context and outlining the objectives established for its execution. Additionally, it examines how the topic has been addressed within the research community, including prior work conducted.

1.1 Context

The tensions between the western world and Islam have always been high. For some reason, Islam has always been painted as the religion that goes against the western ideology [1]. Islam as well as Muslims have always been on the receiving end of a very negative attitude. To this day, in the west, Muslims are shown as the culprits for problems that are out of the hands of the western leaders and reasons that are not clear to anyone.[2]

In November 2015, on "Morning Joe," Trump said that America needs to "watch and study the mosques." Four days later, he indicated that he would "certainly implement" a database to track Muslims in the United States.[3]

In March 2016, during an interview on CNN, Trump expressed his opinion about Islam as "I think Islam hates us. There's something there that — there's tremendous hatred there. There's a tremendous hatred. We have to get to the bottom of it. There's an unbelievable hatred of us." [4]

It has become a trend for the political forces to use islamophobic agenda to gather support and rally the masses for their cause even if it means spreading misinformation and promoting hate.[5] While there are many cases where such political tactics have been used, a very prominent case is that of Donald Trump who knows how to persuade the public with his words and get exactly what he wants. His electoral campaigns rely heavily on an anti-Muslim agenda. In 2016, when elected as the president of the United States for the first time, Trump made extreme arrangements to limit Muslim immigration into the USA.[6] Moreover, his economical and social policies, be they local or foreign, targeted the Muslims in a manner never seen before. The Muslim-ban imposed by his government stunned the world since such discriminatory measures did not seem to be possible in the 21st century. Having a heavy influence in American society, being the businessman that Trump is, he managed to present all these extreme and radical measures as something necessary, even positive for the country. Sadly, that was not the end of it. After losing the elections in 2020, Trump decided to make another run for the office in 2024. This time, with the help of incompetent opponents, Trump's case appeared to be stronger.

---

[1] Islamophobia and U.S. Politics (Ali, 2024)
[2] Anti-Muslim rhetoric of rightwing politicians - The Guardian
[3] A short history of President Trump's anti-Muslim bigotry (Klaas, 2019)
[4] (86 Times Donald Trump Displayed or Promoted Islamophobia | by MPower Change | National Immigration Law Center, 2018)
[5] Politicians, media perpetuating Islamophobia in the U.S. (Xin, 2023)
[6] The Islamophobic Administration (Patel & Levinson, 2017)

1.2 Objectives

This project aims to identify patterns in the speeches delivered by Donald Trump during and after his 2016 electoral campaign with the help of a thorough comparative analysis. In these speeches, Trump frequently addresses the perceived problems and threats facing the United States, often attributing them to immigrants, particularly Muslims. While his language may appear simple and non-threatening on the surface, the underlying messages and rhetorical strategies have had significant influence, both on his audience and on the groups targeted in his discourse.

Based on this, the following preliminary hypotheses are proposed:

1. In pre-election speeches, the focus is expected to be on immigrants, Muslims, political opponents, and the negative portrayals surrounding them, appealing primarily to the emotional and subjective perceptions of the public.

2. In post-election speeches, the emphasis is likely to shift towards political strategies and policy achievements, highlighting their positive impact on the country and appealing to a more objective, fact-based perspective.

3. A recurring "us versus them" narrative is expected, positioning Americans as the in-group and portraying immigrants, Muslims, and political opponents as external threats. This narrative is likely to present Trump as a protective figure or savior in contrast to the dangers associated with these groups.

1.3 Related Work

Many researchers and analysts have studied and explored this topic with numerous approaches with the aim of discovering the patterns hidden in the rhetoric. The most prominent ones revolve around Critical Discourse Analysis, be it manual or automatic, focusing on the sociocultural analysis of Trump's narrative.

Critical Discourse Analysis (CDA) is discourse analytical research that primarily studies the way social-power abuse and inequality are enacted, reproduced, legitimated, and resisted by text and talk in the social and political context[7]. It examines how discourse reflects and reinforces social power structures, inequalities, and ideologies.

Previous work has shown that the rhetoric exploited by Donald Trump has had a detrimental impact on the creation of civilizational harmony because it has clearly contributed to the creation of hostility between Muslims and the West[8] and that Trump's anti-Muslim speeches and actions serve his political purpose but at a high cost to the Muslims seriously aggravating the Muslim – U.S. relations[9]. Additionally, it has been hypothesized that Trump uses language to create a polarity of *us versus them* between the Muslims in America and other groups. He uses negative lexical items to portray Islam and Muslims in the negative light as well as his political opponent, Hillary Clinton. The aim is to paint both Muslims and Hillary Clinton in such a way as to propagate his ideology as a true American. Trump uses a constructive strategy to build his identity as an American nationalist whose only goal is to safeguard America from Muslims[10]. Studies have concluded that Trump used Islamophobia as a

[7] CDA definition (Van Dijk, 2015)
[8] Hegemony of Trump: Manifestation of Islamophobia in an inaugural speech (Badarussyamsi et al., 2024)
[9] President Trump's Islamophobia and the Muslims: A Case Study in Crisis Communication (Nuruzzaman, 2017)
[10] Trump and Muslims: A Critical Discourse Analysis of Islamophobic Rhetoric in Donald Trump's Selected Tweets (Khan et al., 2021)

strategy for winning his campaign in the United States presidential election of 2016[11]. However, some other studies have taken a different perspective by criticizing previous studies to the extent that CDA has been accused of propagating a deterministic version of society[12].

Leaning toward automated analysis, numerous studies have employed various techniques in the field of artificial intelligence, particularly within natural language processing (NLP). One of the most commonly used approaches is sentiment analysis, often applied to datasets sourced from social media platforms, with Twitter being especially popular, to the extent that it has been referred to as the Field's Drosophila Melanogaster[13], which is the model organism for many biology/genetics studies, as Twitter is the model platform for many social media studies. Other works have focused on text classification for the detection of social biases such as racism and xenophobia. These studies frequently utilize machine learning and deep learning techniques, including Long Short-Term Memory (LSTM) networks and Recurrent Neural Networks (RNNs). The present study adopts a hybrid approach, combining both manual and automated methods. It integrates NLP techniques and text framing, supported by manual classification and labeling, to achieve the most accurate results possible.

[11] Islamophobia in an American studies approach as seen in Donald Trump's speech documentary videos (Kasiyarno & Murwantono, 2022)
[12] Critical discourse analysis and its critics (Breeze, 2011)
[13] Questions for Social Media Big data (Tufekci, 2014)

Chapter 2

Data Collection

To conduct the analysis, a comprehensive and reliable dataset was essential. This section outlines the process through which the data was acquired, as well as the various challenges encountered during its collection.

2.1 Data Scraping and Cleaning

To obtain a dataset suitable for analysis in both quality and quantity, numerous sources were explored. Given the political nature of the subject, many websites contained partial records of relevant speeches, but few offered complete and consistent collections. These sources were ultimately excluded, as their contrasting web structures would have required individual scraping scripts, increasing the complexity of the task.

After further investigation, two primary sources were selected for their thorough collections: The American Presidency Project[14] (referred to as Pre-Election collection) and The White House Archives[15] (Post-Election collection). The Pre-Election collection consists primarily of speeches delivered during the 2016 electoral campaign, while the Post-Election collection includes speeches from the post-election period during Trump's presidency. Both collections offered transcripts in formats that were accessible and suitable for further processing.

It is important to note that the final sources were not the only ones considered. The initial objective was to compare speeches across multiple campaigns (2016, 2020, and 2024). However, this idea was later revised due to the limited availability of data from the 2020 and 2024 campaigns. The selected website had ceased receiving funding, resulting in incomplete collections for the more recent election cycles. As a result, the project focus was narrowed to a comparison between the pre-election and post-election periods surrounding the 2016 campaign, where more complete and accessible data was available.

Pre-Election collection was obtained through web crawling techniques and subjected to a series of preprocessing steps. These included filtering for relevant materials, specifically speeches and interviews, while excluding unrelated content such as press briefings by Republican figures other than Donald Trump

Further cleaning involved:

- Speaker detection and attribution, to isolate only Trump's statements.
- Removal of host questions and comments in interviews.
- Exclusion of audience reactions or interjections in public events.
- Normalization of formatting inconsistencies across transcripts.

The Post-Election collection contained well-organized transcripts exclusively of Trump's public addresses, statements, and remarks. Similar cleaning processes were applied, including speaker identification, timestamp removal, and format standardization.

---

[14] The American Presidency Project URL
[15] The White House Archives URL

Finally, manual verification was conducted to ensure both the inclusion of relevant content and the exclusion of unrelated material. During this process, five speeches that were mistakenly omitted by the scraping tools but contained relevant data were identified and manually added to the dataset.

2.2 Data Processing

Once the data collections were completed and cleaned, the process of transforming this data into a format suitable for analysis began. While the final datasets were of reasonable quality and quantity, additional processing was required to ensure they could be effectively utilized to achieve accurate and meaningful results.

Initially, both datasets were loaded separately and subjected to a standard preprocessing pipeline. This process included the following steps:

- Text normalization, involving the conversion of all characters to lowercase to avoid mismatches during string comparisons.
- Extra whitespace removal, ensuring consistency and eliminating unnecessary spacing.
- Removal of stopwords, punctuation and special characters, to focus the analysis strictly on the textual content.
- Tokenization, where each document was split into individual tokens (words).

The resulting tokens from each dataset were stored in separate global lists, preserving the distinction between pre-election and post-election speech content. This structured format facilitated further analysis and comparison across both collections.

2.3 Data Overview

The following were key descriptive statistics related to the processed datasets. The Pre-Election collection contained a total of 80 speeches and 109,433 tokens, while the Post-Election collection comprised 194 speeches and 139,526 tokens. Although the Post-Election collection had more than twice the number of documents compared to the Pre-Election collection, the token count was not proportionally higher. This suggested that the documents in the Post-Election collection were generally shorter in length. This observation was justified by examining the context in which each dataset was compiled. The Pre-Election collection, which related to the electoral campaign, included public addresses and official speeches. In contrast, the Post-Election collection consisted mainly of remarks made by Trump to media personnel and the press during various encounters, which were typically shorter and less formal.

This observation is supported by the plots below, which show the distribution of document lengths for each dataset. On average, documents in the Pre-Election collection contained 1,367 tokens, compared to 719 tokens in the Post-Election collection.

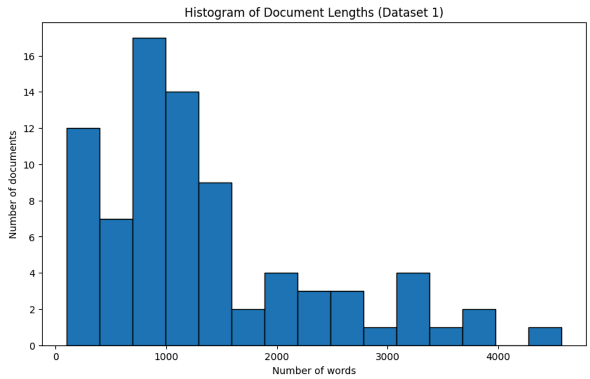


Figure 1: Document Length Histogram Pre-Election collection

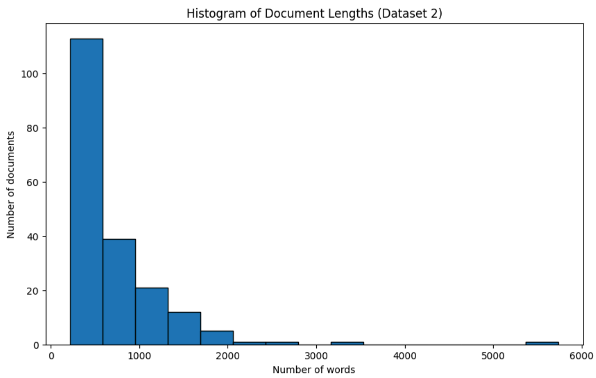


Figure 2: Document Length Histogram Post-Election collection

The vocabulary sizes were 8,736 for Pre-Election collection and 10,192 for Post-Election collection. The most frequent terms in each dataset are illustrated in the accompanying word clouds.

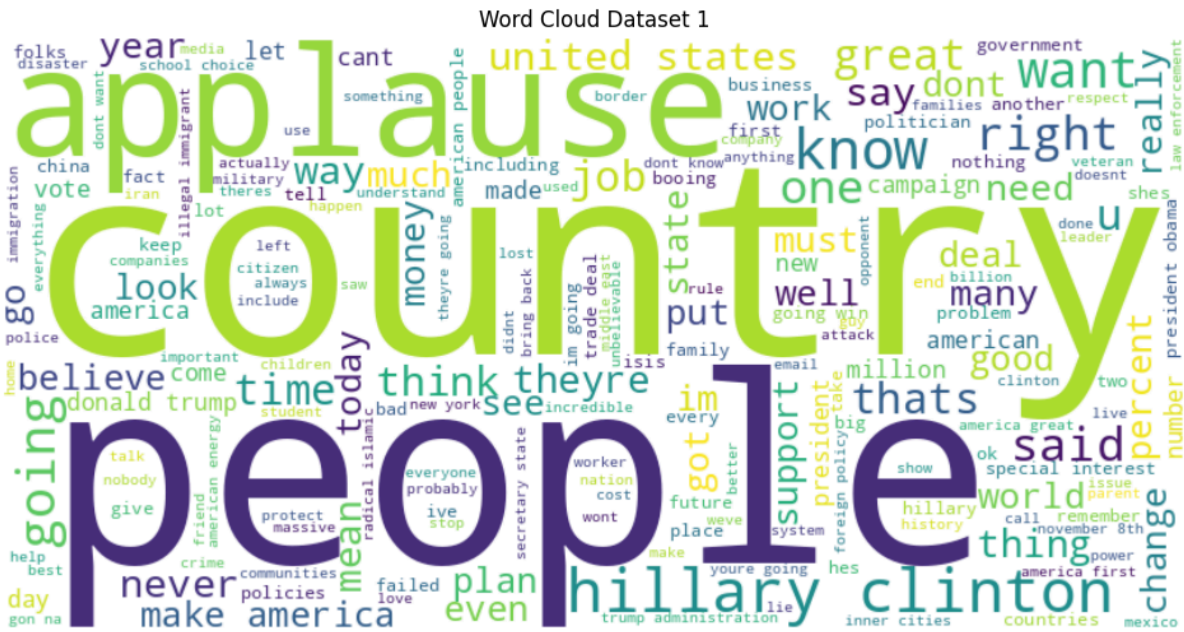


Figure 3: Wordcloud Pre-Election collection

Figure 4: Wordcloud Post-Election collection

The word clouds reveal some broad similarities, such as a focus on the American public, the people, and the country. However, on a more detailed level, distinct thematic differences emerge. The word cloud for the Pre-Election collection includes terms like "*immigration*", "*Mexico*", "*Hillary Clinton*", and "*America First*" suggesting a focus on border security, political figures, and nationalist rhetoric. In contrast, the Post-Election collection features words such as "*nation*", "*honor*", "*leader*", "*support*", "*job*", and "*healthcare*" reflecting themes related to unity, leadership, and social policy. These differences align with the overall tone and rhetorical focus observed in the respective sets of speeches.

Chapter 3

Methodology

This section outlines the methods and tools employed throughout the project. It also provides justification for key decisions, including the choice of libraries and packages, the classification model, and other technical considerations. The following diagram presents the overall structure of the work.

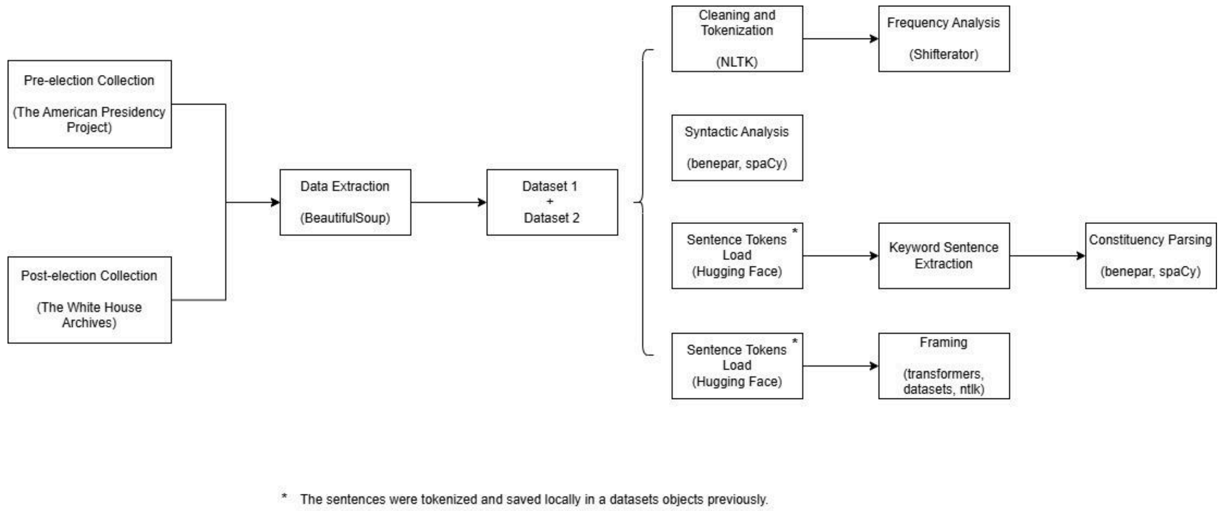


Figure 5: Methodology Structural Diagram

3.1 Corpus Cleaning and Transformation

The initial phase of the project involved web scraping, implemented using *BeautifulSoup*. Supporting modules such as *re* (for regular expressions), *time*, and *requests* were utilized to manage interactions with web servers and handle URL-based data access. As is typical in web scraping tasks, the process required working with HTML content, identifying relevant tags (such as headers), and extracting the necessary text from targeted areas within web pages.

The overall strategy for text analysis was to gradually increase the structural complexity of the process. The first stage consisted of word-level token analysis, which was later extended to grammatical relationships, including structures such as subject–verb and subject–adjective pairs among others. The final stage involved syntactic analysis at the sentence level, incorporating sentence structure interpretation and sentence framing tasks.

As discussed earlier, the collected datasets underwent a thorough preprocessing phase. This phase included:

- Removal of extra whitespaces using the *re* module
- Stopword removal, tokenization, and n-gram (e.g., unigram and bigram) formation using libraries such as *NLTK*

The following example demonstrates the result from the preprocessing of a single sentence and its conversion into bigrams.

"*I'm not knocking immigration or immigrants, but rather am very critical of the country of Mexico for sending us people that they don't want.*"

[
  ('knocking', 'immigration'),
  ('immigration', 'immigrants'),
  ('immigrants', 'critical'),
  ('critical', 'country'),
  ('country', 'mexico'),
  ('mexico', 'sending'),
  ('sending', 'people'),
  ('people', 'want')
]

3.2 Analytical Tools and Techniques

In the later stages of the project, *spaCy* was used for both syntactic analysis and dependency parsing. This involved utilizing *spaCy*'s English language model to perform part-of-speech (POS) tagging and extract syntactic relationships between words in sentences.

To support constituency-based syntactic analysis, the *benepar* package was employed in combination with *spaCy* through a plugin. This enabled the generation of constituency trees, which provided a structural representation of sentence syntax.

The example below illustrates a constituency tree constructed using the *Tree* module from the *NLTK* package, highlighting the hierarchical syntactic structure of a given sentence.

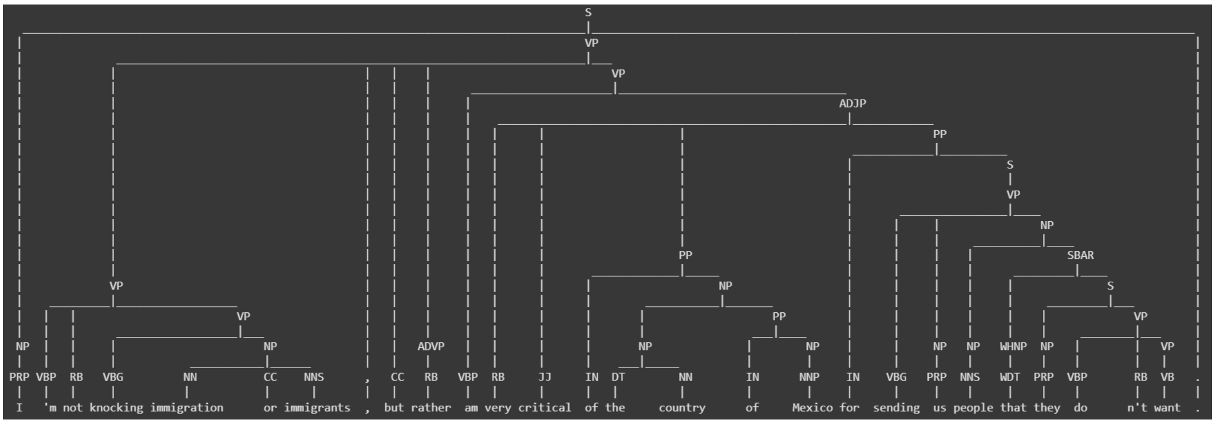


Figure 6: Constituency Parsing example

The parsing process proceeded from identifying individual part-of-speech tags, through predicate-level structures, up to the complete sentence level, forming a tree that captured all grammatical constituents within a sentence.

Furthermore, a dedicated pipeline was developed for the sentence framing task using pretrained language models. To enhance performance and minimize redundant file I/O operations, datasets were loaded into *dataset* objects using the *HuggingFace datasets* module and stored locally, enabling efficient reuse across executions. Three different models were evaluated across multiple framing categories to identify the most effective approach. The final model was selected based on a balance between classification speed and accuracy, with accuracy evaluated manually and qualitatively.

Initially, the established frames were limited to single words. These one-word frames performed reasonably well only with the smallest model, distilbert-base-uncased, a distilled version of BERT, but failed significantly with the larger models. This was primarily because DistilBERT required fine-tuning for Natural Language Inference (NLI) tasks, unlike the other two models, which were already specialized and fine-tuned for NLI. Upon further investigation, it was decided to use clauses instead of single words as frames, as clauses are known to yield better results and better align with the architectures of the larger, more capable models.

The following table presents the average classification time per 1,000 sentences for each model. These values are averaged over multiple executions to provide a more reliable comparison.

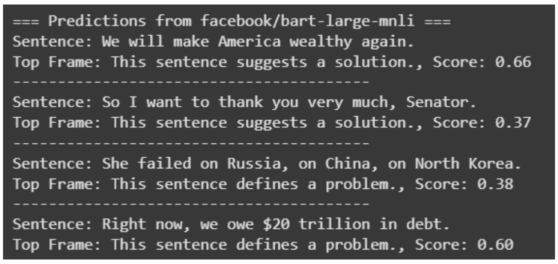
Model	**Time taken per 1000 sentences (seconds)**
DistilBERT	23.32
DeBERTa	95.39
BART-large	82.06

Table 1: Framing Model Comparison

For contextual clarity, it is important to note that the full dataset comprised approximately 40,000 sentences. While the differences in classification time shown in the table may appear minor, they scale substantially across the entire dataset, making even small efficiency gains highly impactful.

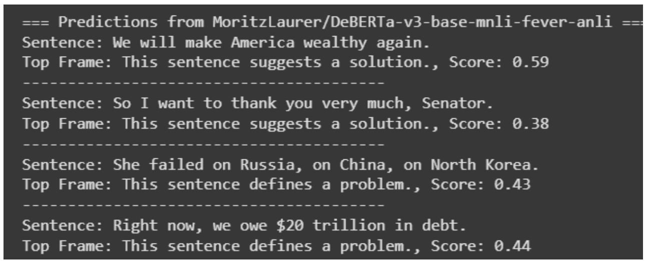
As the table illustrates, *DistilBERT* achieved the fastest inference time. However, this came at the cost of classification quality. In fact, DistilBERT consistently assigned equal probabilities across all frames, effectively failing to perform meaningful classification. This explains its unusually fast processing time as it was not truly making informed predictions.

In contrast, *DeBERTa* and *BART-large*, both fine-tuned for NLI tasks, showed slower inference but produced significantly more accurate and contextually relevant results. In most cases, both models assigned similar frames, differing slightly in the assigned scores. The following images demonstrate how the models performed on a subset of sentences, showing a high degree of alignment in their outputs, though DeBERTa required more time.



```
=== Predictions from facebook/bart-large-mnli ===
Sentence: We will make America wealthy again.
Top Frame: This sentence suggests a solution., Score: 0.66
----------------------------------------
Sentence: So I want to thank you very much, Senator.
Top Frame: This sentence suggests a solution., Score: 0.37
----------------------------------------
Sentence: She failed on Russia, on China, on North Korea.
Top Frame: This sentence defines a problem., Score: 0.38
----------------------------------------
Sentence: Right now, we owe $20 trillion in debt.
Top Frame: This sentence defines a problem., Score: 0.60
```

Figure 7: BART-large Classification example



```
=== Predictions from MoritzLaurer/DeBERTa-v3-base-mnli-fever-anli ==
Sentence: We will make America wealthy again.
Top Frame: This sentence suggests a solution., Score: 0.59
----------------------------------------
Sentence: So I want to thank you very much, Senator.
Top Frame: This sentence suggests a solution., Score: 0.38
----------------------------------------
Sentence: She failed on Russia, on China, on North Korea.
Top Frame: This sentence defines a problem., Score: 0.43
----------------------------------------
Sentence: Right now, we owe $20 trillion in debt.
Top Frame: This sentence defines a problem., Score: 0.44
```

Figure 8: DeBERTa Classification example

Ultimately, since *BART-large* offered comparable classification quality to *DeBERTa* but with lower latency, it was selected as the final model.

AI Disclosure: it is to be noted that due to some technical problems, the built-in plotting function from Shifterator library failed to provide acceptable plots which could be used to present the results. For this reason, a code was developed with the help of an AI assistant that provided very similar results to those from Shifterator. Otherwise, the input of AI in this project has been minimal and limited to making minor modifications and improvements over previously composed content.

Chapter 4

Results

This section presents the results of the different analyses carried out during the project. It spans both basic and advanced levels of linguistic and structural examination. These findings provide essential insights that form the basis for the conclusions drawn later in the report.

4.1 Comparison by Shifts

4.1.1 Frequency Shifts

This section presents a comparison of term and phrase frequencies across the two datasets using shift graphs. The following graph illustrates the most frequent unigrams and bigrams in each dataset.

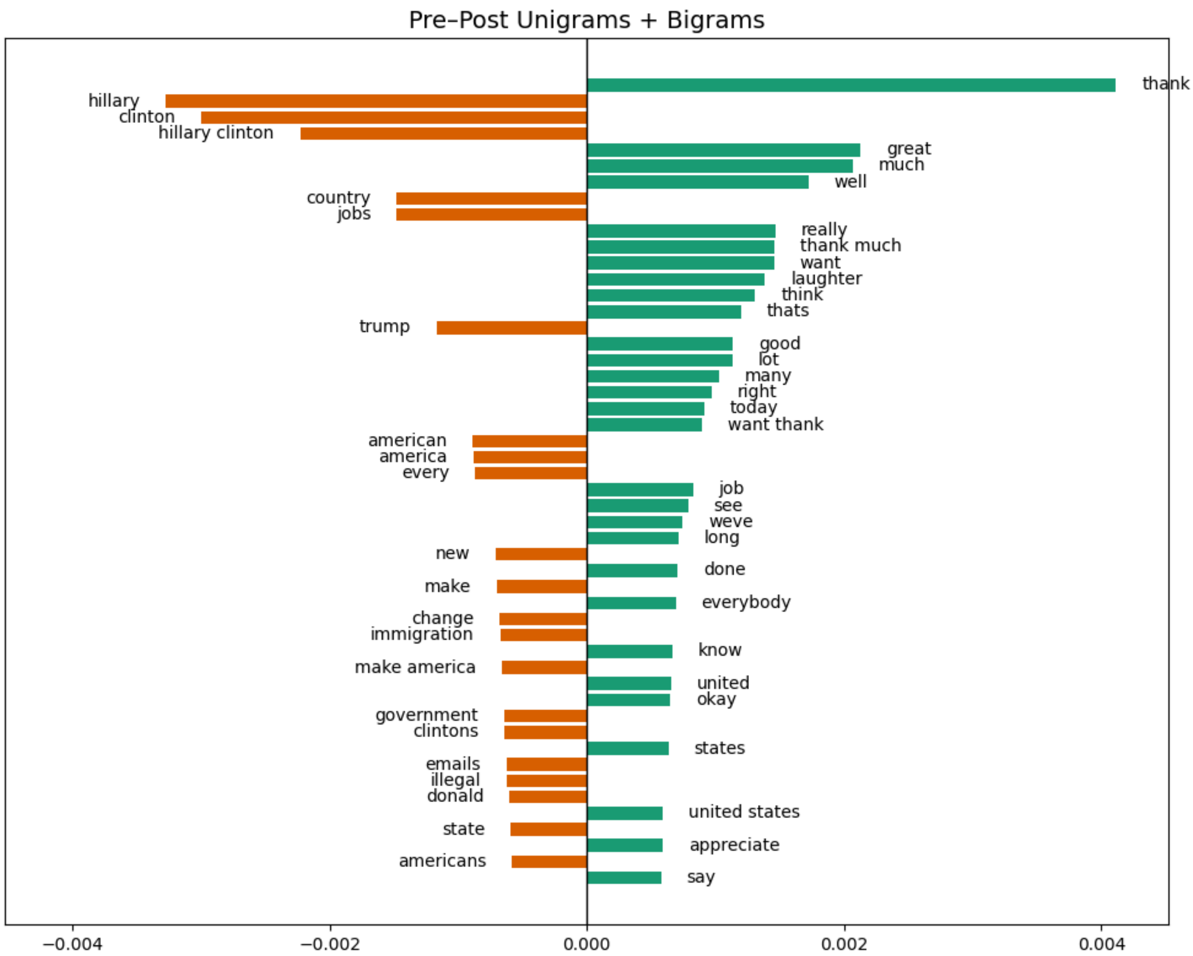


Figure 9: Unigram + Bigram Shift

The left side of the plot (orange) represents the most frequent unigrams and bigrams from the Pre-Election collection, while the right side (green) represents those from the Post-Election collection. The Pre-Election collection prominently features terms such as *Hillary Clinton*, *jobs*, *country*, *immigration*, and *illegal*, reflecting a focus on political and policy issues. In contrast, the Post-Election collection includes expressions of gratitude and references to national identity, with terms like *thank*, *appreciate*, *great*, *United States*, and *nations*.

Below is a plot that focuses only on the bigrams from both datasets.

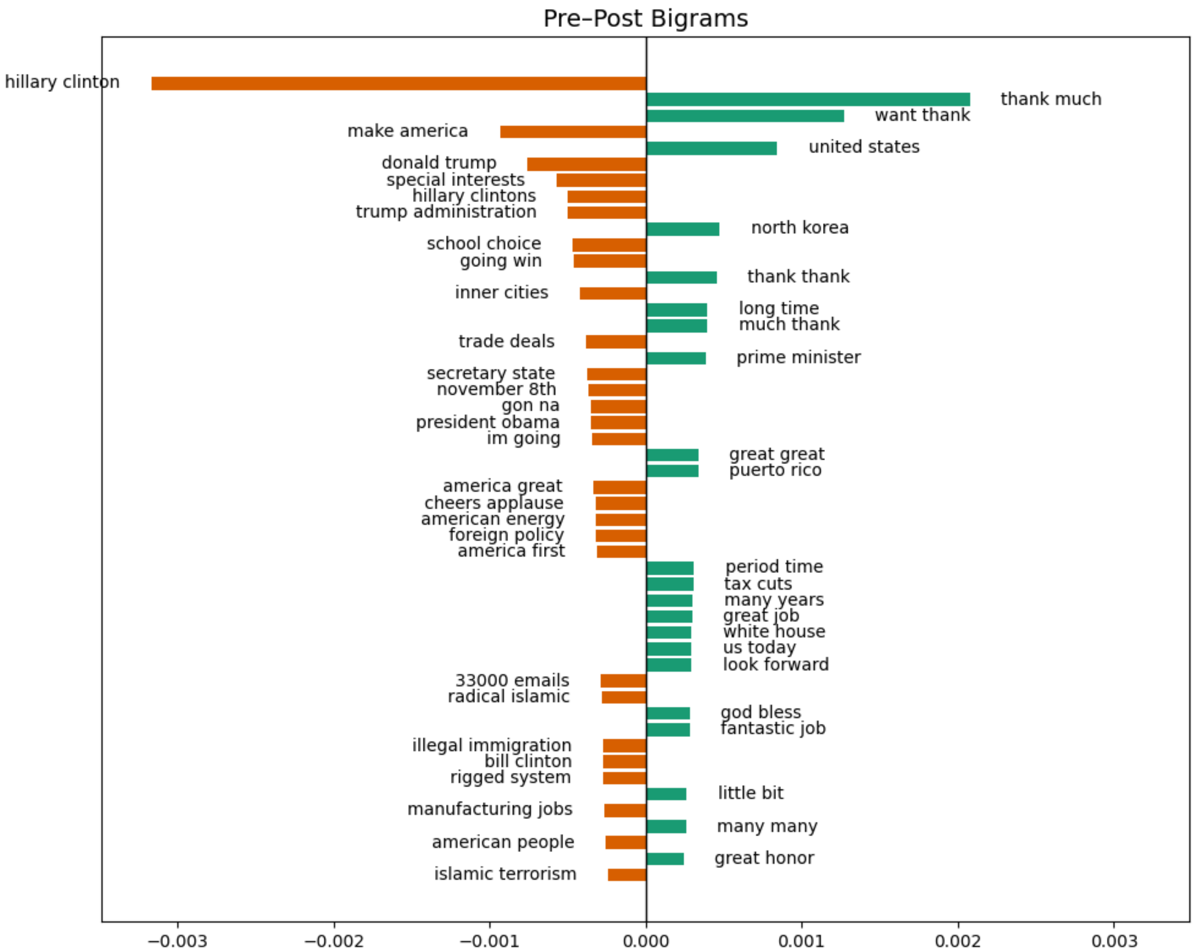


Figure 10: Bigram Shift

While the general thematic trend remains consistent, certain terms stand out, such as *radical Islam*, *rigged system*, *America first* within the Pre-Election collection and *tax cuts*, *great job* etc. in the Post-Election collection.

The graph that follows presents the most frequent unigrams from both collections.

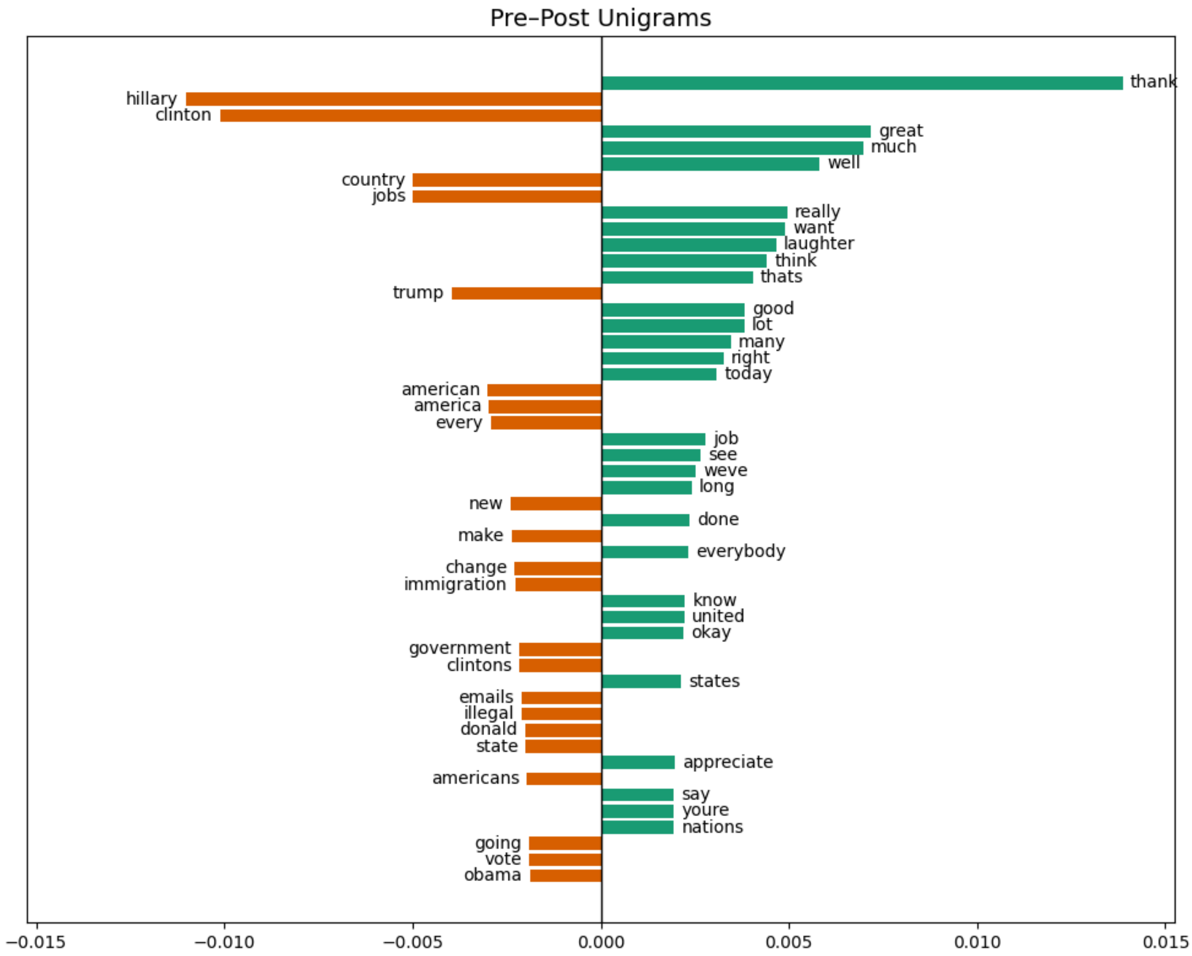


Figure 11: Unigram Shift

It can be observed that the theme persists with the Pre-Election collection containing words like *Clinton, Obama, immigration,* and *illegal* and the Post-Election collection with words such as *great, everybody, job,* and *nations*.

4.1.2 Syntactic Shifts

This section examines syntactic frequency shifts by analyzing recurring structural elements within the corpus. It includes the most frequent subjects, subject–verb pairs, whether compound or simple, and complete subject–verb–object clauses.

a) Subjects

The graphic below illustrates the most frequent sentence subjects in each dataset. In the Pre-Election collection, subjects often refer to political opponents or external entities, including terms such as *Clinton*, *Obama*, *Iran*, *China*, *Isis*, and *Mexico*. In contrast, Post-Election collection tends to center on domestic or collective identities, with frequent use of subjects like *we*, *you*, *states*, *nation*, and *regime*, among others.

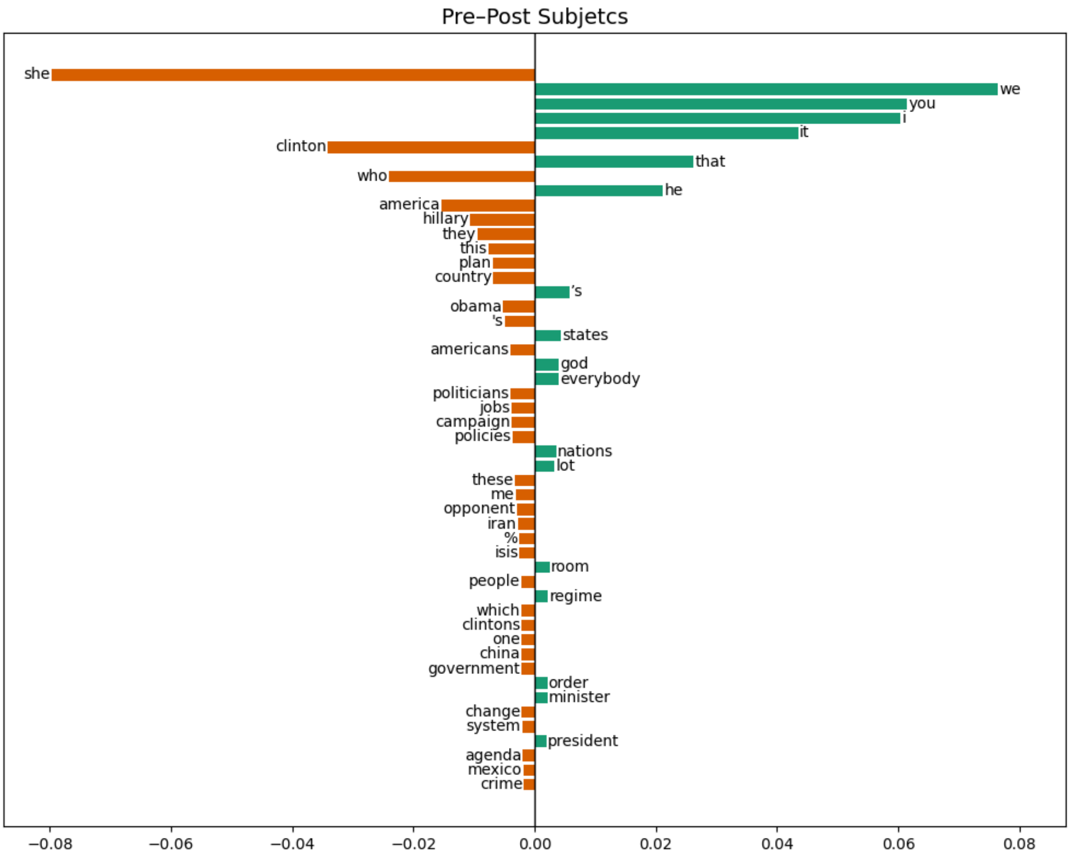


Figure 12: Subjects Shift

b) Subject-Verb pairs

The following plots demonstrate the most frequent subject-verb pairs. These graphics help clarify the actions associated with various entities across the two datasets. In the Pre-Election collection, subject–verb pairs such as *she–say, she–do, she–want,* and *she–have* are common, where *she* clearly refers to Hillary Clinton. In contrast, the Post-Election collection features clauses like *we–want, we–work, we–appreciate,* and *we–do*, emphasizing a collective or inclusive tone.

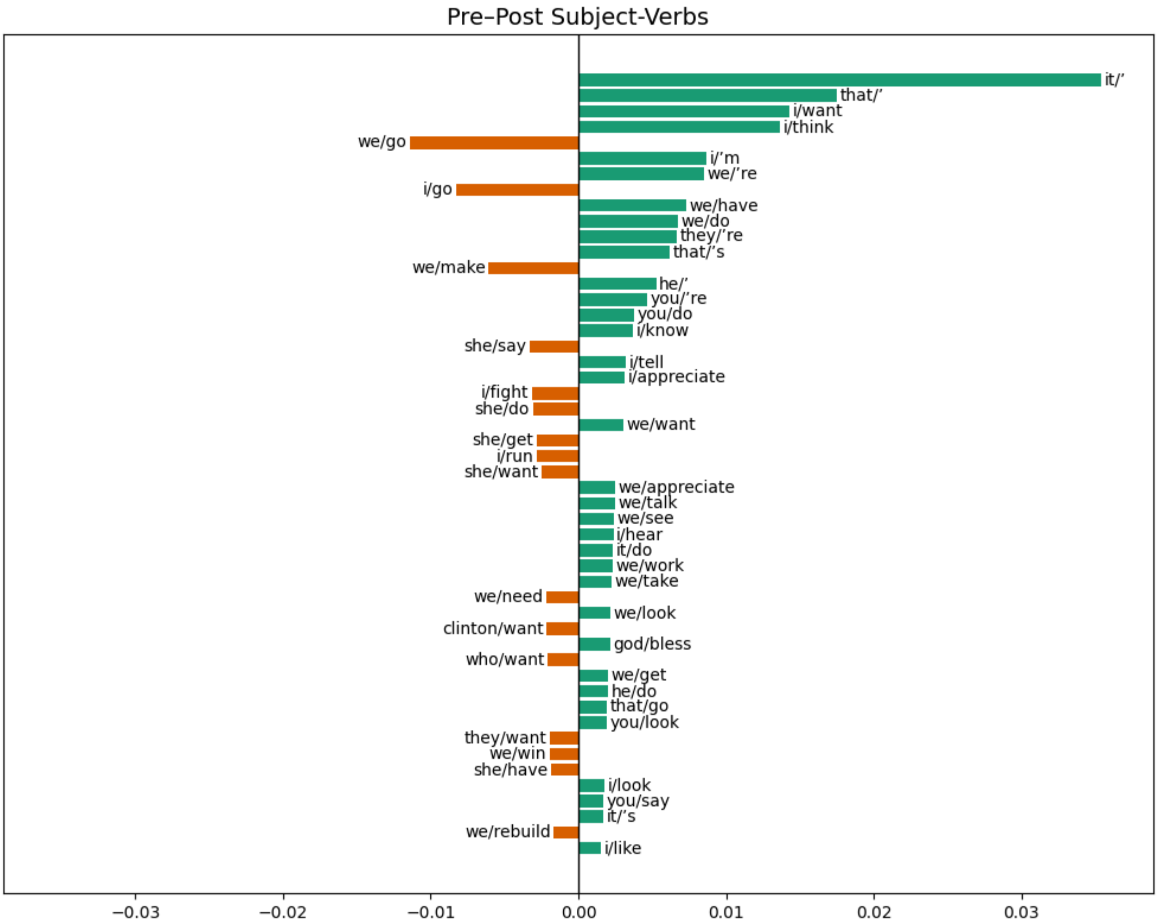


Figure 13: Subject-Verb Shift

Regarding compound subjects, the Pre-Election collection includes constructions like *your money–go, violent crime–rise,* and *your companies–leave*, which focus on loss or threat. Meanwhile, the Post-Election collection contains combinations such as *my administration–work, the people–go,* and *United States–take*, highlighting unity and national identity.

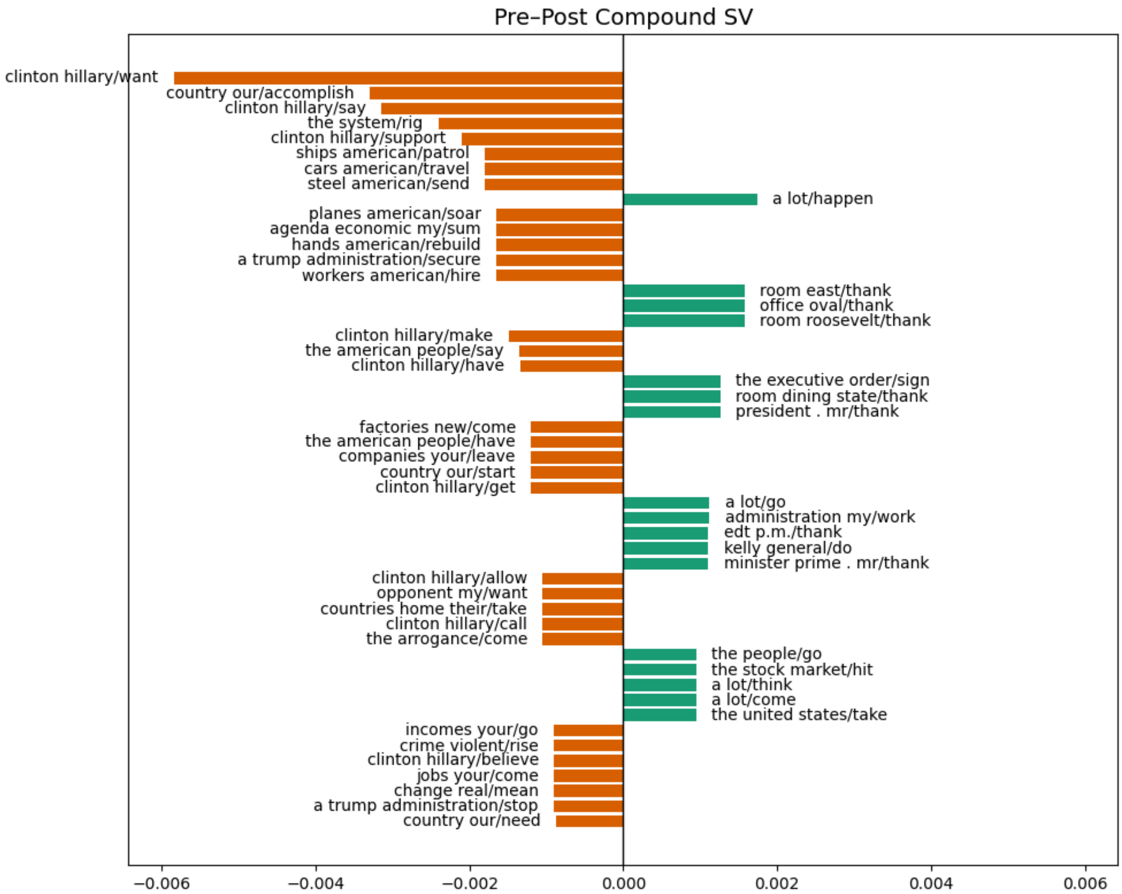


Figure 14: Compound Subject-Verb Shift

c) Subject-Verb-Object clauses

This segment presents the most frequent subject–verb–object (SVO) clauses from both datasets. These shifts illustrate not only the actions associated with recurring subjects but also the targets of those actions. This helps reveal which groups or entities are being talked about and what kind of image is being created around them.

In the graph below, the Pre-Election collection features clauses such as *I–fight–for*, *we–build–wall*, *we–rebuild–roads*, and *I–fix–it*, which emphasize problem-solving and portray the speaker as actively addressing challenges. Similarly, the Post-Election collection presents clauses like *we–appreciate–it*, *we–look–to*, *we–do–job*, and *we–thank–you*, reflecting a tone of gratitude and portraying the speaker as hard-working and solution-oriented.

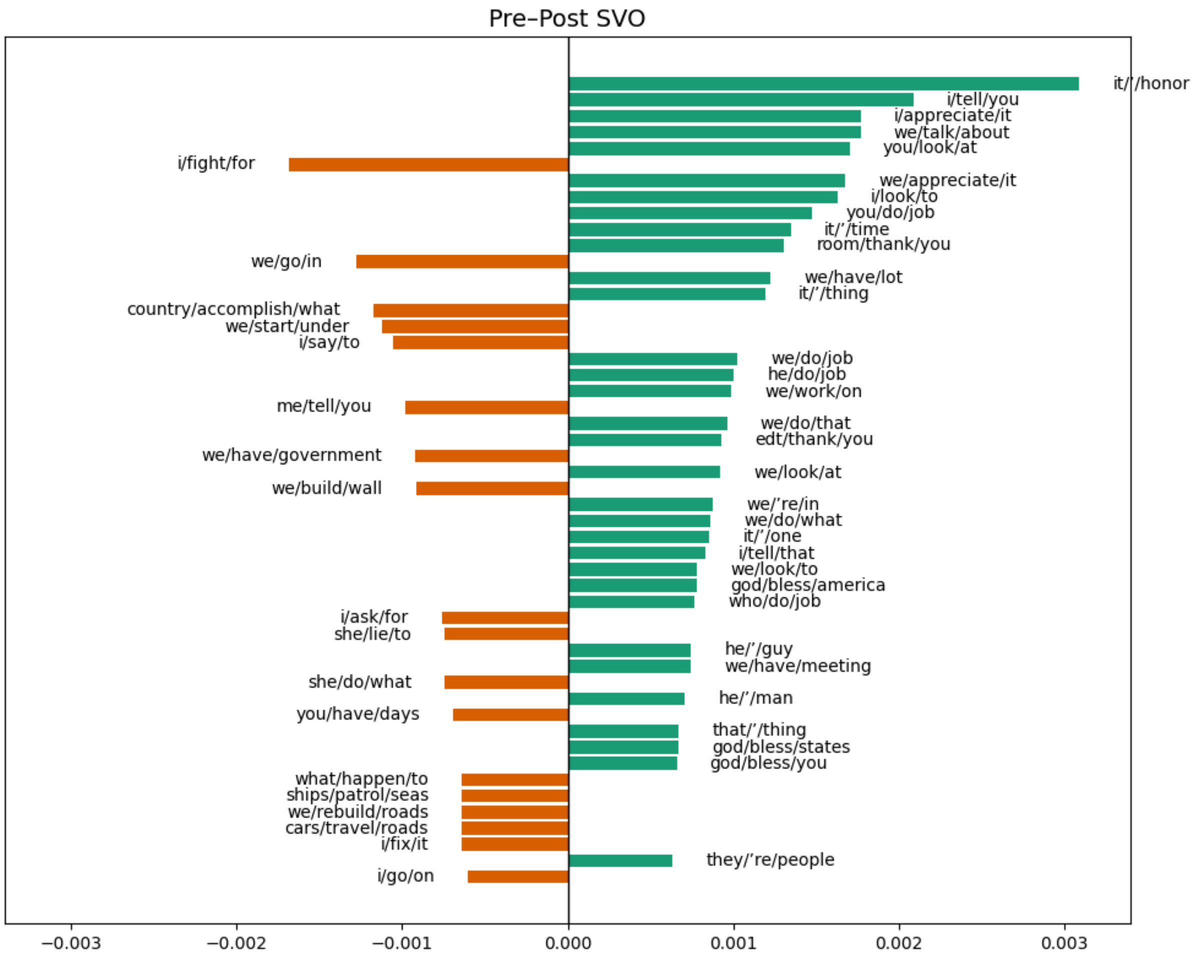


Figure 15: Subject-Verb-Object Shift

4.2 Constituency Parsing

This section presents the findings from the constituency analysis. The table below lists a set of keywords along with the adjectives associated with them, the number of unique adjectives, an example sentence, and the dataset each instance belongs to. Each keyword is analyzed across both collections, allowing for a comparative view of how the rhetoric around these terms differs between the two collections.

Keyword	# of Adjectives	Adjectives	Example Sentence	Collection
Hillary Clinton	96	suppressing, rigged, corrupt, criminal, disqualifying, crooked	“Hillary Clinton who, as most people know, is a world class liar…”	Pre-Election collection
Hillary Clinton	1	terrible	“We had Hillary Clinton give Russia 20 percent of the uranium in our country.”	Post-Election collection
My administration	4	first, major, top, overriding	“My administration, on the other hand, will work with any country that is willing to partner with us to defeat Isis and halt radical islamic terrorism.”	Pre-Election collection
My administration	18	regulatory, sovereign, committed, distinguished, pouring	“My administration is committed to working with the private sector to ensure the protection of made in America…”	Post-Election collection
Islam	86	racist, sexist, failed, disastrous, perpetrated	“Hillary Clinton can never claim to be a friend of the gay community as long as she continues to support	Pre-Election collection

			immigration policies that bring islamic extremists to our country…”	
Islam	11	imposing, tough, strong, revolutionary	“Saudi Arabia is the custodian of the two holiest sites in Islam, and it is there that we will begin to construct a new foundation of cooperation and support with our muslim allies to combat extremism…”	Post-Election collection
Immigrants	79	criminal, dangerous, illegal, violent, horrible	“We will begin removing the more than 2 million criminal illegal immigrants from the country…”	Pre-Election collection
Immigrants	6	united, tremendous, irish	“The Irish immigrants and their descendants have made it right here in the United States and throughout the world.”	Post-Election collection
Jobs	110	stolen, unfair, besieged, disappeared, lowering	“She supports the high taxes and radical regulation that forced jobs out of your community.”	Pre-Election collection
Jobs	136	creating, industrial, financed, independent, pursuing	“We are keeping the promises I made to the American people during my campaign — whether it’s cutting job-killing regulations…”	Post-Election collection

Foreign policy	73	rigged, reckless, islamic, failed, deadly	“In short, the Obama/Clinton foreign policy has unleashed ISIS, destabilized the middle east…”	Pre-Election collection
Foreign policy	7	american, key, economic	“In every foreign policy decision, we are making clear that we will always put the safety and security of our citizens first.”	Post-Election collection
Nation	126	compromised, failed, subsidized, forgotten, disadvantaged	“They're tired of a country that has horrible trade deals, that has no borders…”	Pre-Election collection
Nation	145	independent, strong, bright, unbreakable, confident	“On each trip, I have worked to advance the nation's interests and leadership in the world.”	Post-Election collection

Table 2: Constituency Analysis Results per Keyword

As previously mentioned, this table highlights the adjectives associated with selected keywords. On one level, the quantity of adjectives varies significantly between datasets, suggesting that certain keywords are more frequently discussed in one collection, depending on the message being conveyed. On another level, the adjectives themselves reflect a clear shift in tone. A particularly striking example is the keyword *Islam*, where the descriptors move from negative terms such as *disastrous* and *perpetrated* in Pre-Election collection to more assertive, and even positive, terms like *tough* and *strong* in Post-Election collection. The example sentences further reinforce this contrast. For instance, Muslims are referred to as *allies* in the Post-Election collection, which stands in clear contrast to their portrayal in the Pre-Election collection.

4.3 Framing Analysis

This section presents the results of the sentence framing analysis. As previously described, the sentences were categorized into four framing types: defining a problem, diagnosing a cause, making a moral judgment, and suggesting a solution. The following images display selected examples from each framing category. Each sentence is accompanied by the collection it belongs to. It is important to note that the examples shown here were manually selected from the classifier's outputs in order to condense the otherwise extensive results.

sentence	frame	collection
We have all the cards, but we don't know how to use them.	defines a problem	Pre-Election
Unfortunately when there is action it's always the wrong decision.	defines a problem	Pre-Election
They don't know what to do.	defines a problem	Pre-Election
They're not working for us.	defines a problem	Pre-Election
People don't realize I inherited a mess.	defines a problem	Post-Election

Figure 16: Framing Case 1

This image presents examples of sentences framed as *defining a problem*. In these cases, the identified problem centers on the perceived incompetence and poor judgment of the opposition, as reflected in phrases like *"They don't know what to do"* and descriptions of the country as *"a mess."* Notably, all the examples belong to the Pre-Election collection, which aligns with the critical tone directed toward the opposition during that period.

sentence	frame	collection
For those who feel like they don't have a voice, I have a simple reply: I Am Your Voice.	suggests a solution	Pre-Election
I'm going to get things done for you.	suggests a solution	Pre-Election
Hillary should spend more time producing her illegally hidden emails and less time trying to obfuscate a statement by me that is totally clear and obviously very much accepted by the public as true.	suggests a solution	Pre-Election
I will work with Congress to require that for every 1 new regulation, 2 old regulations must be eliminated.	suggests a solution	Pre-Election
I have some more thoughts on Clinton's remarks I will be sharing momentarily, but first I want to tell you what I am going to do to make your life better.	suggests a solution	Pre-Election

Figure 17: Framing Case 2

This graphic displays sentences framed as *suggesting a solution*. These examples highlight how Trump positions himself as the savior, the voice of the people, and the one who *"gets things done."* As can be seen, all of these sentences originate from the Pre-Election collection, a pattern that is not coincidental. This reflects a strategic narrative in which Trump presents himself as the solution to the nation's problems, which he attributes to the actions and failures of his opposition.

sentence	frame	collection
But the one common thread behind all of these problems was a failure to protect and promote the interests of the American people and American workers.	diagnoses a cause	Post-Election
This reality — this wonderful place — your success is the greatest cause of anxiety, alarm, and even panic to the North Korean regime.	diagnoses a cause	Post-Election
The migration, caused to a large extent -- ISIS was caused by them.	diagnoses a cause	Pre-Election
Her policies in Iraq, Libya and Syria are responsible for the rise of ISIS.	diagnoses a cause	Pre-Election
In the Wall Street Journal, Heather Mac Donald of the Manhattan Institute just wrote an important article about the soaring crime in our inner cities, and the responsibility of Democratic politicians for creating this problem.	diagnoses a cause	Pre-Election

Figure 18: Framing Case 3

The image above presents examples of sentences framed as *diagnosing a cause*. A review of these sentences reveals that the identified causes are largely external, referencing entities such as the *“North Korean regime,” “Iraq,” “Libya,”* and *“ISIS.”* These foreign threats are then linked to the opposition, reinforcing a binary narrative that positions the opposition as responsible for or complicit in these issues.

sentence	frame	collection
It is not compassionate, but reckless to allow uncontrolled entry from places where proper vetting cannot occur.	makes a moral judgment	Post-Election
Ask yourself, who is really the friend of women and the LGBT community, Donald Trump with his actions, or Hillary Clinton with her words?	makes a moral judgment	Pre-Election
The hardworking people she calls deplorable are the most admirable people I know: they are cops and soldiers, teachers and firefighters, young and old, moms and dads, blacks, whites and Latinos – but above everything else, they are all American.	makes a moral judgment	Pre-Election
Her comments are not only reckless, but beneath the dignity of the office she seeks.	makes a moral judgment	Pre-Election
Like how she'll afford to give lifetime welfare and entitlements to illegal immigrants, or how many people will be victimized because of the illegal immigrants that will be released from federal custody – or at the border.	makes a moral judgment	Pre-Election

Figure 19: Framing Case 4

This graphic presents examples of sentences framed as *making a moral judgment*. The judgments expressed here primarily concern immigration policies, which are criticized using terms such as *“reckless”* and *“victimizing.”* The opponent, Hillary Clinton, is portrayed as unfit for the presidency, while a tone of national unity emerges in phrases like *“but above anything else, they are all American.”*

Although these examples may not appear directly related to the core topic of the study, their inclusion is significant. Analyzing them helps uncover broader rhetorical patterns and the linguistic strategies used to communicate ideological messages.

The following graphics show more of the outputs from the model in framing analysis.

There is no financial security.	defines a problem	Pre-Election
We have to do so many things.	defines a problem	Pre-Election
We have regulations on top of regulations.	defines a problem	Post-Election
I have a problem with our incompetent leadership allowing that to happen.	defines a problem	Pre-Election
But they're not getting approved.	defines a problem	Post-Election
And they have some great doctors in the VA, some really talented doctors —\n— the problem is, you just can't — oh, somebody said thank you.	defines a problem	Post-Election
It is under siege in so many ways.	defines a problem	Pre-Election
For example, I understand that once or twice, First Class Cadet Bruce Kim — where's Bruce?	defines a problem	Post-Election
We have no growth.	defines a problem	Pre-Election
Infrastructure question.	defines a problem	Post-Election
Remember, we are competing in a rigged election – the media is trying to rig the election by giving credence to false stories that have no validity and making it front page news, only to poison the minds of the American voters.	defines a problem	Pre-Election
You know, usually, if I do something it's like bedlam, right?	defines a problem	Post-Election
Like we want to start to negotiate with Mexico immediately, and we have these provisions where you have to wait long periods of time, you have to notify Congress, and after you notify Congress, you have to get certified, and then you can't speak to them for 100 days.	defines a problem	Post-Election

Figure 20: Case 1 extended

We are going to lower taxes on American business from 35 percent to 15 percent.	suggests a solution	Pre-Election
We have a great plan.	suggests a solution	Post-Election
We will stop apologizing for America, and we will start celebrating America.	suggests a solution	Pre-Election
Here's your chance.	suggests a solution	Post-Election
I've proposed a moratorium on new federal regulations that are not compelled by Congress or public safety, and I will eliminate all needless and job-killing regulations now on the books.	suggests a solution	Pre-Election
We will build a Marine Corps based on 36 battalions, which the Heritage Foundation notes is the minimum needed to deal with major contingencies – we have 23 now.	suggests a solution	Pre-Election
We're going to do more with less.	suggests a solution	Post-Election
We will establish 'centers of excellence' in places like Philadelphia and Portsmouth New Hampshire and Hampton Roads in Virginia to produce the master craftsmen we need to rebuild our Fleet.	suggests a solution	Pre-Election
Very quickly, I think, we'll have something solved.	suggests a solution	Post-Election
This maternity leave will be paid straight out of the unemployment insurance fund and, again, this safety net will be completely paid-for through savings within the program.	suggests a solution	Pre-Election
The way to make health insurance available to everyone is to lower the cost of health insurance, and that is what we are going do.	suggests a solution	Post-Election
Let's go.	suggests a solution	Pre-Election
Here's what I want.	suggests a solution	Post-Election
We're going to also put a lot of people that have not been able to find jobs, we're going to put them back to work — because they're not registered right now; they'll be registered in a positive sense.	suggests a solution	Post-Election

Figure 21: Case 2 extended

The victory and the win were something that really was dedicated to a country and people that believe in freedom, security, and the rule of law.	diagnoses a cause	Post-Election
We have Rick Perry — is going before.	diagnoses a cause	Post-Election
We all see what happens.	diagnoses a cause	Post-Election
By now they're analyzing.	diagnoses a cause	Pre-Election
No, the mentality.	diagnoses a cause	Pre-Election
[applause] To accomplish a goal, you must state a mission.	diagnoses a cause	Pre-Election
Lee's anthem is the perfect description of the renewed spirit sweeping across our country.	diagnoses a cause	Post-Election
This is part of the ongoing 13-year commemoration of their sacrifice for freedom.	diagnoses a cause	Post-Election
They are driven by a love of country and by a conviction that every American, no matter who they are or where they come from, should have a community that is safe where their families are secure and where their needs will always come first.	diagnoses a cause	Post-Election
You see what's going on.	diagnoses a cause	Post-Election
We are on a mission of change.	diagnoses a cause	Pre-Election
We are fighting for every citizen who believes that government should serve the people – not the donors and special interests.	diagnoses a cause	Pre-Election
And that the economy is a much bigger problem."	diagnoses a cause	Pre-Election

Figure 22: Case 3 extended

I also learned from them, firsthand, how much they despise the Clintons for what they did to the people of Haiti.	makes a moral judgment	Pre-Election
They're not supposed to be doing it.	makes a moral judgment	Pre-Election
So I just want to let the world know: I am 100 percent in favor.	makes a moral judgment	Post-Election
They are the ones she will hurt the most.	makes a moral judgment	Pre-Election
These are individuals encountered or identified by ICE, but who were not detained or processed for deportation because it wouldn't have been politically correct.	makes a moral judgment	Pre-Election
So I've been clear for a long time that we should not have gone in.	makes a moral judgment	Pre-Election
Hillary Clinton is the most corrupt person ever to seek the office of the Presidency.	makes a moral judgment	Pre-Election
These are the most dishonest people anywhere.	makes a moral judgment	Pre-Election
[applause]\nAnd these are among the most dishonest people.	makes a moral judgment	Pre-Election
But to this Administration, their amazing daughter was just one more American life that wasn't worth protecting.	makes a moral judgment	Pre-Election
But there is no right to engage in violent disruption, or to threaten the public safety and peace.	makes a moral judgment	Pre-Election
Our commitment to them is absolute, and I mean absolute.	makes a moral judgment	Pre-Election
When I say it they say that's terrible the way you talk.	makes a moral judgment	Pre-Election
Hillary Clinton is the most corrupt person ever to seek the office of the presidency of the United States.	makes a moral judgment	Pre-Election

Figure 23: Case 4 extended

These examples are shown to illustrate the boundaries of the classification model. It can be seen that while the classified classes are mostly accurate, there are cases where the classified frame does not provide any information about the sentence and the interpretation of results is difficult without proper context. Some evident examples are “We have to do so many things.” being framed as a definition of a problem or “Let’s go” framed as a solution suggestion.

Chapter 5

Conclusions, Limitations and Future Work

This section presents the conclusions drawn from the previously discussed results. It evaluates whether the initial hypotheses have been confirmed and addresses the limitations encountered during the study. Additionally, the section discusses the study's contributions to the field and outlines potential directions for future research it may inspire or support.

5.1 Conclusions

In order to check whether the hypotheses were fulfilled, it is necessary that they be recalled:

1. In pre-election speeches, one could expect an emphasis on immigrants, Muslims, political opponents, and the negative portrayals surrounding them, appealing primarily to the emotional and subjective perceptions of the public.

2. In post-election speeches, the emphasis is likely to shift towards political strategies and policy achievements, highlighting their positive impact on the country and appealing to a more objective, fact-based perspective.

3. A recurring "us versus them" narrative is expected, positioning Americans as the in-group and portraying immigrants, Muslims, and political opponents as external threats. This narrative is likely to present Trump as a protective figure or savior in contrast to the dangers associated with these groups.

Based on the premises and the previously presented results, it can be concluded that all three of the initial hypotheses have been supported. In the Pre-Election collection, there is a pronounced focus on figures and entities perceived as foreign or alien to the nation. Throughout this dataset, the groups identified in the hypotheses, namely Muslims, immigrants, and political opponents, are consistently portrayed in a negative light. These groups are framed as the root causes of the problems faced by the common American and as external threats to the identity and integrity of the United States and its citizens. This framing appears to be strategically designed to persuade the public to turn against these groups and align themselves with a figure who promises safety, security, and protection from such threats.

In contrast, the Post-Election collection reflects a significant shift in focus toward political and social policy. The discourse emphasizes the failures of previous administrations and portrays the country as being in disarray. Trump is presented as the individual capable of restoring order and addressing the needs of the working-class American through policy measures aimed at economic and social reform. Particularly in the frames classified as *solution suggestions*, Trump positions himself as the voice of the voiceless, a capable and selfless leader who can bring about real change and return the nation to greatness.

In summary, the rhetorical strategies observed across both collections function to construct a binary opposition between, on one side, Trump, his administration, and the American people, and on the other, immigrants, Muslims, and political opponents. Trump aligns himself with the common American, emphasizing shared values and loyalty to national identity, while casting external groups as threats. This *us vs. them* narrative effectively establishes a clear division and has served as a powerful tool in garnering public support, ultimately contributing to his rise to the presidency.

5.2 Limitations

One of the main challenges encountered during this project concerned the quality and consistency of the data. While the sources from which the data was collected generally provided reasonably clean content, there were notable inconsistencies in data formatting. For instance, the two main sources had different structures for speech transcripts, which required manual inspection and cleaning of the scraped files. In some cases, even within the same source, structural inconsistencies were found. A particularly significant issue arose with interview transcripts, which contained varied speaker tags for hosts and interviewees. Additionally, inconsistent use of colons and dashes to separate the speaker's name from the spoken text complicated the preprocessing efforts. As a result, transcripts, especially those from the Pre-Election collection, had to be manually reviewed to ensure that such tags did not interfere with the analysis.

Another major challenge was the use of the Shifterator library, which was central to the comparison process through its shift-based analysis. However, the library offers limited plotting capabilities. The built-in plot function includes a distribution subplot that takes up excessive space and overlaps with the main visualization. Moreover, this subplot cannot be adjusted or removed via parameters. Since the alternative modules provided by the library do not support plotting, a custom implementation of the plot function had to be developed to produce usable and clear visual outputs.

A further complication involved the sentence framing models. Despite significant advancements in Natural Language Inference (NLI), the models occasionally classified vague references as full actions. For example, a sentence like *"This is the reason for our poor performance"* could be labeled as *diagnoses a cause*, even though the actual cause, referred to as "this", is not explicitly stated. This limitation motivated manual inspection of framed sentences to ensure accuracy and meaningful interpretation. Consequently, a filtering process was conducted to retain only the most relevant and clearly framed examples.

5.3 Future Work

While this project has provided valuable insights into rhetorical strategies across political discourse with the use of notable techniques like CDA and the modern tools of NLP, several opportunities remain for further development and refinement.

Firstly, future research could benefit from the use of more robust and standardized datasets. Considering the nature of the study, the development of pipelines tailored to handle political speech could reduce manual effort and increase the reproducibility of the analysis. Moreover, the limitations of current Natural Language Inference (NLI) models in detecting contextually complete frames highlight a need for more advanced or fine-tuned models specifically adapted to political language.

Finally, the binary opposition narrative observed in this study presents an opportunity for deeper rhetorical and ideological analysis. Further studies could investigate how such narratives evolve over time, extending across different speakers or political parties. Expanding the scope beyond a single political figure could also provide comparative insights into how rhetoric strategies differ across ideological lines.

Bibliography

[1] Ali, M. (2024). *Islamophobia & U.S. Politics*. AMERICAN CULTURE STUDIES. Retrieved May 19, 2025, from https://amcs.wustl.edu/islamophobia-us-politics

[2] Badarussyamsi, B., Syam, N., Ermawati, E., & As'ad, A. (2024, May 31). Hegemony of Trump: Manifestation of Islamophobia in an inaugural speech. *Hegemony of Trump: Manifestation of Islamophobia in an inaugural speech*. https://ejournal.upi.edu/index.php/IJAL/article/view/70355

[3] Breeze, R. (2011, January 4). Critical discourse analysis and its critics. https://www.jbe-platform.com/content/journals/10.1075/prag.21.4.01bre

[4] *86 Times Donald Trump Displayed or Promoted Islamophobia | by MPower Change | National Immigration Law Center*. (2018, April 19). Medium. Retrieved May 20, 2025, from https://medium.com/nilc/86-times-donald-trump-displayed-or-promoted-islamophobia-49e67584ac10

[5] Kasiyarno, K., & Murwantono, D. (2022, February 28). Islamophobia in an American studies approach as seen in Donald Trump's speech documentary videos. https://jurnal.unissula.ac.id/index.php/edulite/article/view/20159

[6] Khan, M. H., Qazalbash, F., Adnan, H., Yaqin, L. N., & Khuhro, R. A. (2021, March 31). Trump and Muslims: A Critical Discourse Analysis of Islamophobic Rhetoric in Donald Trump's Selected Tweets. https://journals.sagepub.com/doi/10.1177/21582440211004172

[7] Klaas, B. (2019, March 15). *A short history of President Trump's anti-Muslim bigotry*. The Washington Post. Retrieved May 20, 2025, from https://www.washingtonpost.com/opinions/2019/03/15/short-history-president-trumps-anti-muslim-bigotry/

[8] Nazeer, T. (2024, February 26). *The anti-Muslim rhetoric of rightwing politicians is fuelling hate crime – I've experienced it myself | Tasnim Nazeer*. The Guardian. Retrieved May 19, 2025, from

https://www.theguardian.com/commentisfree/2024/feb/26/anti-muslim-rightwing-politicians-hate-crime-lee-anderson

[9] Nuruzzaman, M. (2017, August 3). President Trump's Islamophobia and the Muslims: A Case Study in Crisis Communication. https://greenpublishers.org/jms/index.php/ijcc/article/view/66

[10] Patel, F., & Levinson, R. (2017, April 19). *The Islamophobic Administration*. Brennan Center for Justice. Retrieved May 15, 2025, from https://www.brennancenter.org/our-work/research-reports/islamophobic-administration

[11] *2016 Presidential Election Speeches & Remarks | The American Presidency Project*. (n.d.). The American Presidency Project. Retrieved May 26, 2025, from https://www.presidency.ucsb.edu/documents/presidential-documents-archive-guidebook/documents-related-to-presidential-elections/2016/report/200301/all/63

[12] Van Dijk, T. A. (2015, January 16). *Critical Discourse Analysis*. Discourses.org. Retrieved June 2, 2025, from https://discourses.org/wp-content/uploads/2022/07/Teun-A.-van-Dijk-2015-Critical-discourse-Analysis.pdf

[13] *The White House Archives*. (n.d.). Remarks by Donald Trump as President. https://trumpwhitehouse.archives.gov/remarks/page/199/

[14] Xin, Y. (2023, November 27). *Politicians, media perpetuating Islamophobia in the US*. China Daily. Retrieved May 19, 2025, from https://www.chinadaily.com.cn/a/202311/27/WS6563ff43a31090682a5f02d1.html

[15] Tufekci, Z. (2014, March). Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. https://www.researchgate.net/publication/261217624_Big_Questions_for_Social_Media_Big_Data_Representativeness_Validity_and_Other_Methodological_Pitfalls

Appendix

A.1 Python Code Notebooks

All the code, datasets and other related files can be found in the following Github repository.

https://github.com/uzair-uni/TFG_Comparative_Study.