

Semantic Polarization in U.S. Politics: A Computational Analysis of Republican and Democratic Speeches

[Author names redacted by Managing Editor]

[School names redacted by Managing Editor]

* These authors have contributed equally to this work

ABSTRACT

This research investigates the evolution of U.S. political rhetoric over the past eighty years, focusing on how thematic priorities and emotional framing have influenced partisan discourse. The study applies computational methods, including semantic similarity and sentiment analysis, to presidential nomination speeches from 1944 to 2024, with the aim of understanding patterns of ideological alignment, divergence, and rhetorical strategy. Beyond identifying these patterns in party messaging, the research examines how differences in thematic emphasis and emotional tone shape voter perception, engagement, and political polarization. By combining quantitative linguistic analysis with historical and political context, the study provides insight into the mechanisms through which political language both mirrors and reinforces broader social and ideological divisions in the United States.

INTRODUCTION

Political discourse in the United States has undergone significant transformation over the past century, affecting not only policy content but also rhetorical style, emotional tone, and approaches to persuasion. Recent empirical research indicates that political language is shifting away from evidence-based reasoning toward more emotionally charged, subjective rhetoric, with important implications for democratic governance.

A study published in *Nature Human Behaviour* (2025) analyzed over 8 million congressional speeches from 1879 to 2022 (Bhatia et al., 2025). The results show a decline since the 1970s in the use of fact-based language, such as “data,” “analysis,” “evidence,” and “investigation,” alongside a rise in opinion-based expressions, including “point of view,” “believe,” “common sense,” and “feel.” The authors argue that this move away from evidence-based rhetoric correlates with increasing political polarization, reduced legislative productivity, and growing socio-economic inequalities.

This trend has coincided with broader phenomena, including increasing political polarization, declining legislative cooperation, and rising socio-economic inequality. Empirical studies suggest that these patterns are linked to changes in political language, which has become more emotionally charged and less grounded in shared facts or evidence. Research on partisan speech over the past 130 years indicates that the phrasing and rhetorical markers associated with each major party have grown increasingly distinct and divergent. This growing divergence reflects not only differences in policy but also strategic choices in how parties communicate, reinforcing partisan identities and shaping public perception.

These shifts raise critical questions about the health of democratic discourse. When political communication moves away from shared reference points and parties adopt increasingly divergent rhetorical registers, opportunities for consensus or mutual understanding may diminish. Scholars have described this process as *Affective Polarization*, in which partisans not only disagree on issues but also increasingly distrust and dislike members of the opposing party (Gidengil & Stolle, 2022). Such polarization is part of a broader transformation in U.S. political communication, with implications for governance, civic engagement, and social cohesion.

In this context, the present study examines presidential nomination speeches in U.S. politics from 1944 to 2024. Rather than analyzing votes or policy outcomes, the research focuses on language itself, considering both thematic content and emotional tone. By applying computational linguistics methods, including semantic similarity analysis and sentiment analysis, the study traces the evolution of political rhetoric across the span of 80 years.

Through this approach, the study seeks to address deeper questions about political communication. It examines the change in the substance of party rhetoric, the evolution in the style of expression, and what these changes imply for political polarization, public discourse, and democratic engagement.

METHODOLOGY

The research presented in this paper employs a quantitative approach to analyze political polarization through the lens of language. The methodology is designed to collect, clean, and analyze a large corpus of political discourse using advanced natural language processing techniques. This section details the steps taken to construct the analytical framework.

Data Collection

The corpus for this study was sourced from the American Presidency Project (APP) at the University of California, Santa Barbara. This archive provides a comprehensive collection of presidential public documents, including speeches, letters, and conference transcripts. The dataset selected for this research consisted of convention speeches, specifically those accepting presidential nominations, from 1944 to 2024.

Once the raw data was collected, it underwent a preprocessing pipeline to prepare the texts for linguistic analysis. The speeches were stored as .docx files and thoroughly cleaned to ensure accurate results. The first step removed capitalization and punctuation. All characters were then converted to lowercase to ensure consistency within the document. Next, stop word filtering removed common words with minimal semantic value, leaving only terms with genuine meaning. The final step was lemmatization, which converted words to their base form (e.g., "running," "ran," and "runs" all become "run"). This process grouped semantically equivalent terms, which was essential for accurate embeddings and similarity analysis.

Algorithmic Formalism Explained

1. Statistical Divergence: Jaccard Similarity

Jaccard Similarity (J) measures the overlap between the unique vocabulary of two datasets, in this case, the Democratic and Republican speeches. This metric quantifies the ratio of shared unique words relative to all unique words in both speeches. Let V_D represent the cleaned and lemmatized words in the Democrat dataset, and V_R the same for the Republican dataset.

The Jaccard Similarity is defined as the size of the intersection divided by the size of the union.

$$J(V_D, V_R) = \frac{|V_D \cap V_R|}{|V_D \cup V_R|}$$

The score ranges from 0, indicating very high vocabulary divergence, to 1.0, indicating nearly identical vocabulary with minimal divergence.

2. Cosine Similarity on TF-IDF Vectors

Semantic similarity is assessed using Cosine Similarity applied to document vectors weighted by Term Frequency-Inverse Document Frequency (TF-IDF). This method emphasizes words that are unique to one document while down-weighting common words shared between both speeches.

2a. TF-IDF Weight Calculation

The weight $W_{t,d}$ of any pre-processed term t in a speech d is the product of its Term Frequency (TF) and Inverse Document Frequency (IDF):

$$W_{t,d} = \text{TF}(t, d) \times \text{IDF}(t)$$

TF is the raw count of the term in the document, with higher values indicating centrality to that speech.

IDF measures the term's rarity across the two-speech corpus ($N = 2$) using a smoothed formula:

$$\text{IDF}(t) = 1 + \ln\left(\frac{N}{\text{DF}(t)}\right)$$

This is the formula when $\text{DF}(t)$ is the number of speeches (one or two) containing the term t . Terms that appear in only one of the two speeches receive the highest weight, amplifying their impact on semantic comparison.

2b. Cosine Similarity Formula

After vectorization, Cosine Similarity measures the cosine of the angle between the two speech vectors. The result is calculated as:

$$\text{Sim}_C(V_D, V_R) = \frac{V_D \cdot V_R}{\|V_D\| \|V_R\|}$$

The resulting score ranges from 0 to 1.0, where 1.0 indicates highly similar speech content and word usage.

3. VADER Sentiment Analysis

VADER (Valence Aware Dictionary and Sentiment Reasoner) is a rule-based sentiment analysis tool originally designed for social media text, though effective for speeches due to its handling of intensifiers. VADER provides four scores for the text as a whole:

- Positive (pos): proportion of text with a positive tone.
- Neutral (neu): proportion of text with a neutral tone.
- Negative (neg): proportion of the text with a negative tone.
- Compound (compound): a single, non-proportional score summarizing overall sentiment strength and direction, ranging from -1 (extremely negative) to $+1$ (extremely positive).

RESULTS

This section of the paper presents the empirical findings derived from the analysis of the Democratic and Republican speeches. The core analysis focuses on two comparisons:

semantic similarity and emotional valence. The semantic similarity is calculated using the cosine similarity of the speeches' aggregated word vectors. The emotional metric is quantified using VADER polarity scores, which capture the emotional valence of each speech. It is important to note that the VADER analysis is conducted on the raw text, whereas the semantic similarity analysis uses lemmatized text.

Semantic Similarity

The data from figure 1 reveals three epochs, each marked by distinct mathematical trends: an initial period of volatility, followed by a trajectory toward convergence, and concluding with steep divergence. The overall mean between 1944 and 2024 is 0.4319, with a notably high standard deviation of 0.1295, indicating significant political shifts over time.

Epoch 1: Early Volatility (1944 – ≈ 1960)

The absolute minimum occurs in 1948 (Truman v. Dewey), with a score of 0.1832—the lowest point on the graph. This reflects maximum thematic divergence, meaning the angle between the two parties' topic vectors was widest, indicating the most dissimilar issue sets in the entire eighty-year span. After 1948, the similarity rose sharply (e.g., 0.4119 in 1956) before falling again, reflecting a highly fluctuating environment as the political center of gravity continued to shift.

Epoch 2: Trend Towards Convergence (≈ 1964 – 1996)

This period spans over three decades and shows a move toward shared thematic ground, despite major events such as the Vietnam War and the end of the Cold War. The scores oscillate (e.g., 0.4014 in 1964, peaking briefly at 0.5369 in 1972, then dipping in the 1980s), but the underlying trend is upward. Mathematical estimation returns a slope of 0.00236. This positive slope confirms a steady increase in semantic similarity, culminating in a peak score of 0.4770, showing that the two parties most closely addressed similar sets of issues during this era.

Epoch 3: Modern Divergence (2000 – 2024)

After 2000, the trend reversed sharply. Using the same slope formula, the average slope is -0.0157 , indicating divergence at a rate more than six times greater than the previous period's convergence. The score reached 0.2301, representing a substantial decline in thematic agreement in contrast with the 2000 peak.

Comparative Sentiment: Differential Use of Positive vs. Negative Language

The time-series analysis of party-specific emotional trends, quantified using VADER sentiment scores (Figures 2.1 and 2.2), establishes the parameters for differential language use by locating the centroid of emotional scores and assessing temporal volatility.

The emotional rhetoric of the Democratic and Republican Parties, as illustrated in Figure 2.1 and Figure 2.2, shows a highly stable positive differential. The key observation is the structural maintenance of this positive differential across the eighty-year period for all but one speech. The Positive Sentiment line consistently maintains a visible, non-zero distance from the Negative Sentiment line. This universal positivity of the Framing Score indicates a bounded emotional space: the emotional centroid remains anchored in the net-positive quadrant. The absence of cross-overs shows that party communication adheres to a minimum positivity threshold that prevents sustained shifts into negative valence. Although both Positive and Negative Sentiment lines trend upward—indicating growing emotional variance—the integrity of the positive differential is preserved.

However, the 2016 election cycle marks an exception to this rhetorical constraint. This year shows an ephemeral valence inversion that reduces the stability of the positive differential. The raw VADER scores reveal an atypical rise in Negative Sentiment to 0.162, coinciding with a contraction of Positive Sentiment to 0.158. The brief moment where the Negative score exceeds the Positive score produces a nominally negative Framing Score. Furthermore, the Compound VADER score for this year falls to -0.997 , reflecting extreme negative sentiment and demonstrating the party's minimum positive boundary.

DISCUSSION

This section of the paper interprets the empirical results of the semantic similarity scores and the emotional valence scores of the Democratic and Republican convention addresses. First, the two extreme points in Figure 1, representing the highest and lowest cosine similarity scores, are illustrated by outlining the political climate of each year, giving brief context on both party platforms, and assessing how closely the nominees' convention speeches align with their respective platforms, thereby explaining their points of convergence or divergence. Second, the reasons, implications, and limits behind the overall net positive tone of all speeches and the distinct net negative tone of Trump's convention address are deciphered using Figures 2.1 and 2.2.

Semantic Similarity

The Lowest Cosine Similarity Score – The 1948 Presidential Election:

Post-1945, the United States emerged from World War II as a political, economic, and military victor. While much of its cities, factories, and infrastructure remained intact—unlike Europe—it faced the complex task of transitioning from a wartime to a peacetime economy. During the war, the U.S. economy had been sustained by the production of military goods, such as missiles and tanks, which had significantly contributed to its recovery from the Great Depression. Following the war, demand for consumer goods surged while wartime production declined, leading to a rapid rise in prices and a corresponding increase in the cost of living. By late 1948, these factors had already placed the U.S. on the path to a brief recession, which fully materialized in 1949 (Gilder Lehrman Institute of American History, n.d.; TrendSpider, n.d.).

This economic transition had profound implications for labor and, by extension, politics. Between 1945 and 1946, the U.S. experienced a wave of labor strikes, as workers demanded higher wages and improved conditions amidst inflation and uncertainty. This labor unrest, coupled with broader economic difficulties, generated significant social and political pressures, as many Americans felt insecure about employment, prices, and overall stability (The National WWII Museum, n.d.).

By 1948, new sources of tension had emerged internationally. The U.S. was engaged in a global rivalry with the Soviet Union, encompassing political, economic, and ideological dimensions. In this context, the United States increasingly relied on foreign aid and diplomacy as instruments of influence, exemplified by the Marshall Plan, which aimed to rebuild war-torn Europe. The underlying objective was to foster European stability and prosperity, thereby reducing susceptibility to communism and Soviet influence. This period thus marked a shift toward a new form of U.S. global leadership, centered on economic and political engagement (U.S. Department of State, Office of the Historian, n.d.; Library of Congress, n.d.).

Consequently, the domestic political climate in 1948 was characterized by widespread anxiety and uncertainty, regarding inflation, job security, and global threats including communism, Soviet expansion, and nuclear weapons. Divisions over labor policy, economic strategy, and foreign affairs emerged, as decisions made during this period would shape U.S. history in the coming years. This context helps explain the unpredictability of elections at the time, as citizens held contrasting views on the nation's next steps.

Amid this environment, presidential candidates Harry Truman and Thomas Dewey proposed different reform agendas to address economic and social challenges. Truman positioned himself as a populist champion, advocating for labor, farmers, full employment, fair wages,

and programs to protect workers and farmers. His foreign policy emphasized confronting the spread of communism, exemplified by the Marshall Plan and support for war-torn Europe. Dewey, by contrast, focused on moderate social reforms, including expanded Social Security, public housing assistance, civil-rights legislation, recognition of Israel, desegregation of the military, and increased federal support for health and education. In essence, Democrats sought to safeguard farmers and workers through government intervention, while Republicans emphasized individual opportunity alongside fundamental protections and civil liberties (Peters & Woolley, n.d.-a).

The candidates' campaign strategies reflected these policy approaches. Truman conducted extensive tours of small towns, speaking directly to citizens, criticizing Congress, and establishing credibility among workers and farmers through a populist message, thereby appearing relatable and committed to the common man. Dewey adopted a cautious approach, avoiding direct criticism of Truman and employing general, optimistic rhetoric, which, amid postwar uncertainty, risked appearing uninspiring or detached (Miller Center, n.d.).

Analysis of conventional speeches reveals clear stylistic and thematic differences. Truman portrayed the Democratic Party as the party of the people, citing evidence of sustained service and using economic statistics to reinforce credibility. He highlighted Republican shortcomings and invoked memories of the Great Depression to underscore the high stakes of the election, appealing broadly to the electorate. Dewey emphasized reform, good governance, and national unity through patriotic themes, but without detailing concrete policy measures. While he addressed Cold War anxieties and juxtaposed wartime trauma with future opportunities, he did not clearly articulate how this vision would be realized or why it surpassed Truman's proposals. These thematic divergences in speech content and style help explain the historically low cosine similarity observed in the 1948 election (Peters & Woolley, n.d.-b).

The Highest Cosine Similarity Score – The 2000 Presidential Election:

By the late 1990s, the United States had experienced strong economic growth driven by technology and IT investment. The previous administration had used this prosperity to reduce the national debt and produce a federal budget surplus, the largest in U.S. history. As a result, the 2000 presidential election centered on debates over how this surplus should be used. Key issues included the future of Social Security and Medicare—especially with an aging population nearing retirement—alongside taxes, healthcare, and education. Although

domestic concerns dominated, foreign-policy debates also emerged around national missile defense, American leadership abroad, and the U.S. role in international agreements.

Globally, the 1990s had been a period of significant change. Post–Cold War adjustments, expanding global trade, rapid globalization, technological innovation, and rising environmental concerns created anxiety among American workers and communities. Many feared job losses tied to globalization, even as the overall economy was growing. At the same time, environmental issues were gaining political traction as voters prioritized pollution control, environmental regulation, and sustainable development. Thus, in 2000, the United States stood economically strong, yet deliberated about inequality, social welfare, environmental protection, and the broader consequences of a globalized world (The White House, 2000; Borger, 2000).

The Democratic Party, led by Al Gore, sought to build on the economic stability of the 1990s. Their platform emphasized maintaining a balanced budget and continuing to pay down the national debt. They prioritized protecting Social Security and Medicare, and proposed tax-free savings accounts with matching government contributions to help families prepare for retirement. Democrats also focused on improving education through investment, accountability, and support for STEM fields. They advocated for workers by raising the minimum wage, strengthening overtime protections, improving workplace conditions, and safeguarding the right of unions to organize. Their environmental vision centered on addressing global warming and supporting wildlife conservation. The party framed economic growth as something that should benefit all Americans, which shaped their positions on immigrant protections, disability rights, and support for historically marginalized groups. Gore specifically emphasized reproductive rights and policies aimed at strengthening families and communities. In foreign policy, Democrats promoted international cooperation, strong alliances, the security of Israel, and the global spread of democracy, human rights, and civil society, while also supporting a robust military (Peters & Woolley, n.d.-a; OnTheIssues.org, n.d.-a).

The Republican Party, led by George W. Bush, focused on continued economic expansion through large tax cuts and support for innovation. They argued that lower taxes would boost savings, investment, entrepreneurship, and business activity. Their platform included expanding affordable housing through programs that would revitalize government property and help low-income families purchase homes. Republicans aimed to increase access to higher education and stimulate job creation by supporting small businesses. They supported globalization and open markets by reducing trade barriers. Their values emphasized tradition, community involvement, and moral leadership. In environmental policy, they argued

that economic growth and environmental protection should advance together. In foreign affairs, Republicans prioritized strengthening the U.S. military, preparing for new national-security threats, and protecting allies (Peters & Woolley, n.d.-a; OnTheIssues.org, n.d.-b).

Despite these policy differences, the candidates' speeches shared a notable degree of similarity in tone, structure, and thematic emphasis. Both Gore and Bush opened by thanking party leaders, supporters, and family members. Both discussed the need for educational reform, particularly around accountability and school safety. Their foreign-policy sections focused on national security and military readiness. Both speeches incorporated personal stories and encounters with Americans affected by national problems, using these narratives to explain their motivations and character. Each candidate also critiqued the opposition in an effort to frame their own party as the one prepared to take action.

Both candidates drew on American values of responsibility and community empowerment, though the specific framing differed. Bush emphasized "compassionate conservatism," rooted in tradition and community action. Gore highlighted government support for families, parents, and the middle class. Both acknowledged the country's economic prosperity and used it to outline their visions for policy—though their proposed solutions diverged sharply. Bush advocated for large, broad-based tax cuts, while Gore favored targeted tax cuts for those who needed them most. Gore supported an active government role in expanding social protections and guiding welfare reforms; Bush promoted limited government, deregulation, and greater reliance on community and private-sector initiatives.

Gore's speech included a wider range of policy areas—campaign-finance reform, national-debt reduction, childcare expansion, health and medical research, mental-health awareness, affirmative action, abortion rights, victim and disability rights, child safety, and community policing. This breadth was intended to portray him as a candidate focused on clear, actionable solutions. In contrast, Bush emphasized a set of core values—integrity, dignity, unity, compassion, tolerance, optimism, readiness, transparency, and patriotism—to present himself as the moral and steady choice for leadership. Democrats positioned themselves as the party of the common man and the middle class, while Republicans framed themselves as the party of innovation and courage—traits they claimed Democrats often misinterpreted as risk-taking.

A closer examination of the speeches reveals that the two parties were addressing many of the same national challenges but differed in their strategies for solving them. This underlying similarity—combined with contrasting solutions—helps explain the extremely close vote in

the 2000 election, which was ultimately decided in George W. Bush's favor through the Electoral College (Peters & Woolley, n.d.-b).

Comparative Sentiment: Differential Use of Positive vs. Negative Language

An analysis of presidential nomination acceptance speeches over the past eighty years shows that nearly all candidates delivered addresses with a net positive emotional tone, with one exception: Donald Trump's 2016 Republican National Convention speech. Four main reasons explain why every other speech conveyed net positive sentiment.

First, nomination acceptance speeches are moments of celebration, gratitude, and party unity. Candidates traditionally begin with expressions of thanks—using phrases such as “I accept your nomination” or “I would like to thank”—which contain lexicon items tagged as positive in VADER. Second, these speeches are designed to be forward-looking. Candidates outline promises, opportunities, and hope, as the primary goal is to inspire and mobilize voters. Third, nominees address the entire nation instead of narrow voter blocs. They emphasize shared values, the common good, and the American Dream, themes that naturally produce positive sentiment in lexicon-based methods. Finally, sustained use of inclusive pronouns such as “we,” “our,” and “us,” alongside words like “proud,” “honoured,” and “grateful,” helps maintain an overall positive emotional profile (Peters & Woolley, n.d.-b).

Positive emotion in acceptance speeches has significant public effects. Inspiring language strengthens enthusiasm among existing supporters and offers undecided voters reasons to align with the candidate or party. Such speeches can increase turnout, volunteering, donations, political discussion, and overall engagement. They also help set the agenda for the election by communicating a candidate's priorities, values, and policy direction, shaping how voters interpret the campaign. When used consistently, positive emotional messaging can strengthen party reputation and build trust, ultimately supporting electoral success (Tavits & Potter, 2022).

However, this strategy has limitations. Positive rhetoric can appear unrealistic or disconnected when the country is facing a crisis, war, or economic instability. In such times, voters may perceive overly optimistic messaging as incompetent or out of touch. Furthermore, positive emotion does not guarantee political advantage. Research shows that negative emotions spread faster and evoke stronger reactions, which is why fear-based appeals dominate political advertising and social media. Given the current climate of intense

polarization and public frustration in the United States, a purely positive speech may not effectively mobilize voters who feel exhausted, angry, and disappointed with the government (Iyengar et al., 2019; Pew Research Center, 2023).

By contrast, Donald Trump's 2016 acceptance address showed a clear net negative sentiment, the only instance of its kind in the history of presidential nomination speeches. The dominant tone of his speech was rooted in fear. He framed the United States as being in a state of crisis, repeatedly referencing terrorism, violence, and attacks on police. His use of intense descriptors—such as “brutally murdered,” “barbarians,” “savage killers,” “horrible,” “unfair,” “disaster,” “ruins,” and “chaos”—created vivid, violent imagery that contributed to the overall negative emotional score. More importantly, Trump devoted substantial time to detailing the failures of existing policies and explicitly blamed Democratic leaders. He directly criticized Hillary Clinton, Bill Clinton, and Barack Obama for rising crime, terrorism, immigration problems, violence involving undocumented immigrants, corruption, economic stagnation, unemployment, failing infrastructure, weakened military strength, national-security threats, poor trade deals, foreign-policy shortcomings, and political elitism. Although presidents often point to systemic problems as a way of introducing their own plans, Trump's emphasis on failure was unusually detailed and accusatory (Peters & Woolley, n.d.-b).

The public response to the speech was mixed in both the short and long term. Immediately after the conventions, Trump experienced a larger polling “bump” than Hillary Clinton—his support rose by 3 percent, compared to her 2 percent increase. However, because Clinton was already ahead before the conventions, this translated into only a 1-point net gain for Trump, leaving her still in the lead (Peters & Woolley, n.d.-c). A Gallup poll found that after the conventions, voters were 15 percent less likely to support Trump and 4 percent more likely to support Clinton. It was the first time since Gallup began asking the question in 1984 that a candidate left a convention with a net negative effect (Edwards-Lee, 2016).

Despite these shifts, both candidates ended the 2016 campaign with historically low favorability ratings. The year produced two of the least liked nominees in modern election history, suggesting that convention speeches did little to reshape voters' deeper opinions. Scholars note that conventions can influence short-term enthusiasm, media narratives, and undecided voters, but long-term turnout and final vote decisions depend on broader factors such as economic conditions, campaign activity, and later events. In 2016, the small convention effects did not meaningfully predict the election outcome. Experts therefore view the speech's direct influence on turnout as limited and overshadowed by subsequent developments (Norman, 2016).

CONCLUSION

This analysis of U.S. presidential election speeches from 1944 to 2024 identifies clear patterns in both thematic alignment and emotional framing across time. The semantic similarity results indicate that political rhetoric has moved through distinct phases: an early period of volatility (1944–1960) marked by sharp fluctuations in issue emphasis, a prolonged phase of convergence (≈1964–1996) during which parties addressed substantially overlapping national concerns, and a modern period of divergence (1996–2024) characterized by rapidly increasing thematic distance. The 1948 election reflects an extreme point of divergence shaped by post-war political, economic, and social uncertainty, whereas the 2000 election illustrates high semantic alignment despite policy differences, demonstrating how shared national priorities can persist during times of economic stability and globalization.

The comparative sentiment analysis shows that presidential nomination speeches have historically maintained a net-positive emotional tone aimed at reinforcing party unity and encouraging public mobilization. Deviations from this pattern, such as Donald Trump's 2016 address, reveal how highly negative, fear-based rhetoric can shift short-term public perceptions without necessarily determining final electoral outcomes. Together, these trends show how political communication balances substantive issue framing with the strategic use of emotion.

Overall, the findings highlight the evolving nature of U.S. political discourse. While parties may converge around core national issues during certain historical moments, the methods through which they frame and address these concerns—both semantically and emotionally—can differ substantially. Understanding these dynamics offers insight into historical electoral outcomes and broader patterns of political polarization and voter engagement in the United States.

REFERENCES

1. Bhatia, S., Lewandowsky, S., & Politzer, G. (2025). *Evidence for increasing intuitive reasoning in U.S. political language*. *Nature Human Behaviour*.
<https://www.nature.com/articles/s41562-025-02136-2>
2. Gidengil, E., & Stolle, D. (2022). Emotions, partisanship, and political behaviour. *Political Behavior*, 44, 1331–1354. <https://doi.org/10.1007/s11109-022-09784-4>
3. Gilder Lehrman Institute of American History. (n.d.). *Postwar politics and the Cold War*.

<https://www.gilderlehrman.org/history-resources/essays/postwar-politics-and-cold-war>

4. TrendSpider. (n.d.). *The post-war recession of 1948–1949*.
<https://trendspider.com/learning-center/the-post-war-recession-1948-1949/>
5. The National WWII Museum. (n.d.). *Episode 5: The strike wave*.
<https://www.nationalww2museum.org/war/podcasts/best-my-ability-podcast/season-2-archive/episode-5-strike-wave>
6. U.S. Department of State, Office of the Historian. (n.d.). *Introduction: 1945–1952*.
<https://history.state.gov/milestones/1945-1952/foreword>
7. Library of Congress. (n.d.). *Post-war United States (1945–1968): Overview*.
<https://www.loc.gov/classroom-materials/united-states-history-primary-source-timeline/post-war-united-states-1945-1968/overview/>
8. Peters, G., & Woolley, J. T. (n.d.). *Party platforms*. The American Presidency Project.
<https://www.presidency.ucsb.edu/documents/app-categories/elections-and-transitions/party-platforms>
9. Miller Center. (n.d.). *Harry S. Truman: Campaigns and elections*.
<https://millercenter.org/president/truman/campaigns-and-elections>
10. Peters, G., & Woolley, J. T. (n.d.). *Convention speeches*. The American Presidency Project.
<https://www.presidency.ucsb.edu/documents/app-categories/elections-and-transitions/convention-speeches>
11. The White House. (2000). *Helping to Build the New Economy*.
<https://clintonwhitehouse5.archives.gov/WH/new/html/20000225.html>
12. Borger, J. (2000, October 17). *Not voting for Gore*. The Guardian.
<https://www.theguardian.com/world/2000/oct/17/uselections2000.usa3>
13. OnTheIssues.org. (n.d.). *2000 Democratic National Platform: Environment*.
https://www.ontheissues.org/Archive/2000_DNC_Platform_Environment.htm
14. OnTheIssues.org. (n.d.). *2000 Republican National Platform: Tax reform*.
https://www.ontheissues.org/Archive/2000_RNC_Platform_Tax_Reform.htm
15. Tavits, M., & Potter, J. (2022). *The Nature and Impact of Emotional Content in Congressional Candidate Emails to Supporters*. Science Direct.
<https://www.sciencedirect.com/science/article/abs/pii/S0261379422000609>
16. Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22, 129–146.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC6690035>

17. Pew Research Center. (2023). *Americans' feelings about politics, polarization, and the tone of political discourse*.
<https://www.pewresearch.org/politics/2023/09/19/americans-feelings-about-politics-polarization-and-the-tone-of-political-discourse>
 18. Peters, G., & Woolley, J. T. (n.d.). *The post-convention bounce in voter preference*. The American Presidency Project.
<https://www.presidency.ucsb.edu/statistics/data/the-post-convention-bounce-voters-preference>
 19. Edwards-Lee, S. (2016, July 19). *Republican convention: Why some Republicans are not voting for Trump*. TIME Magazine.
<https://time.com/4433348/republican-convention-not-voting-trump/>
 20. Norman, J. (2016). *Trump, Clinton finish with historically poor images*. Gallup.
<https://news.gallup.com/poll/197231/trump-clinton-finish-historically-poor-images.asp>
- [X](#)

APPENDICES

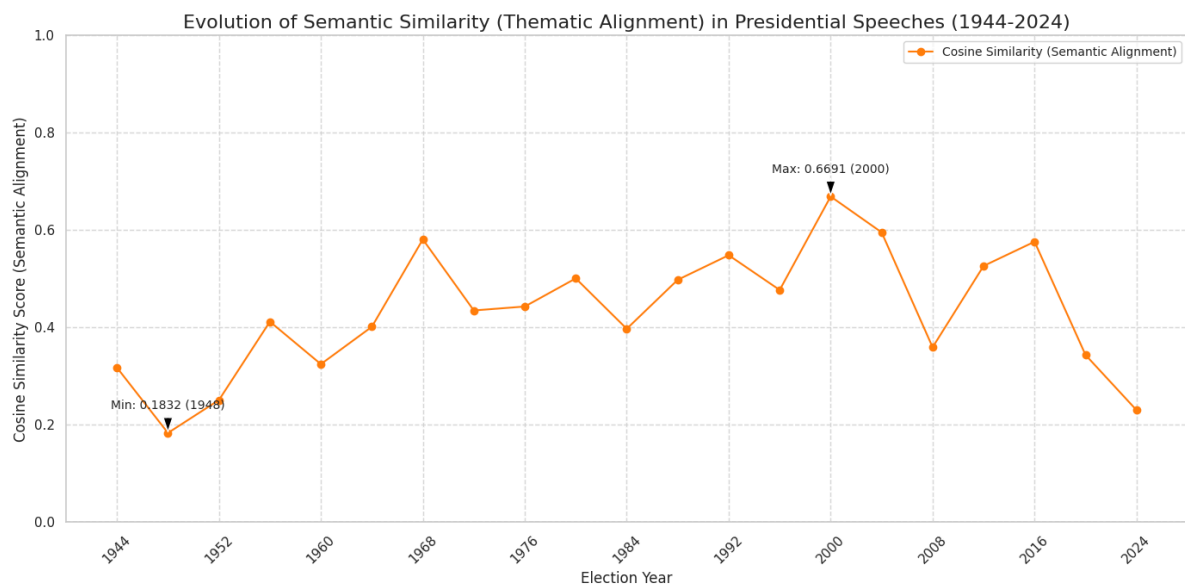


Figure 1 – Evolution of Semantic Similarity (Thematic Alignment) in Presidential Speeches (1944-2024)

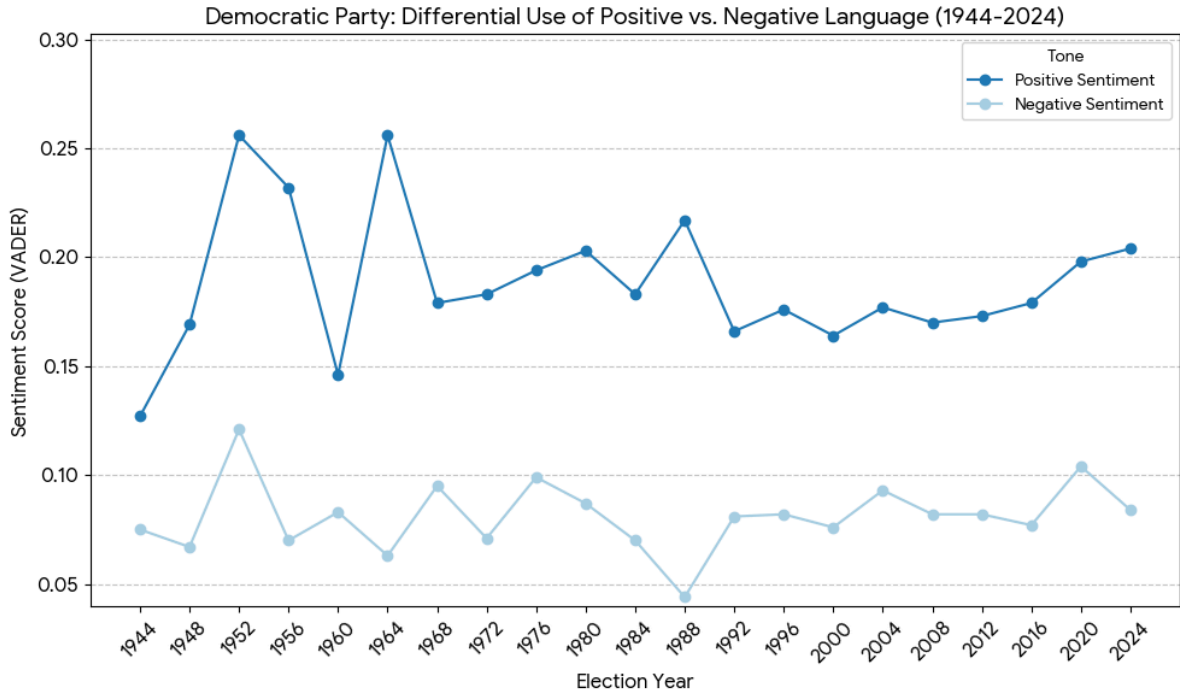


Figure 2.1 – Democratic Party: Differential Use of Positive vs. Negative Language (1944-2024)

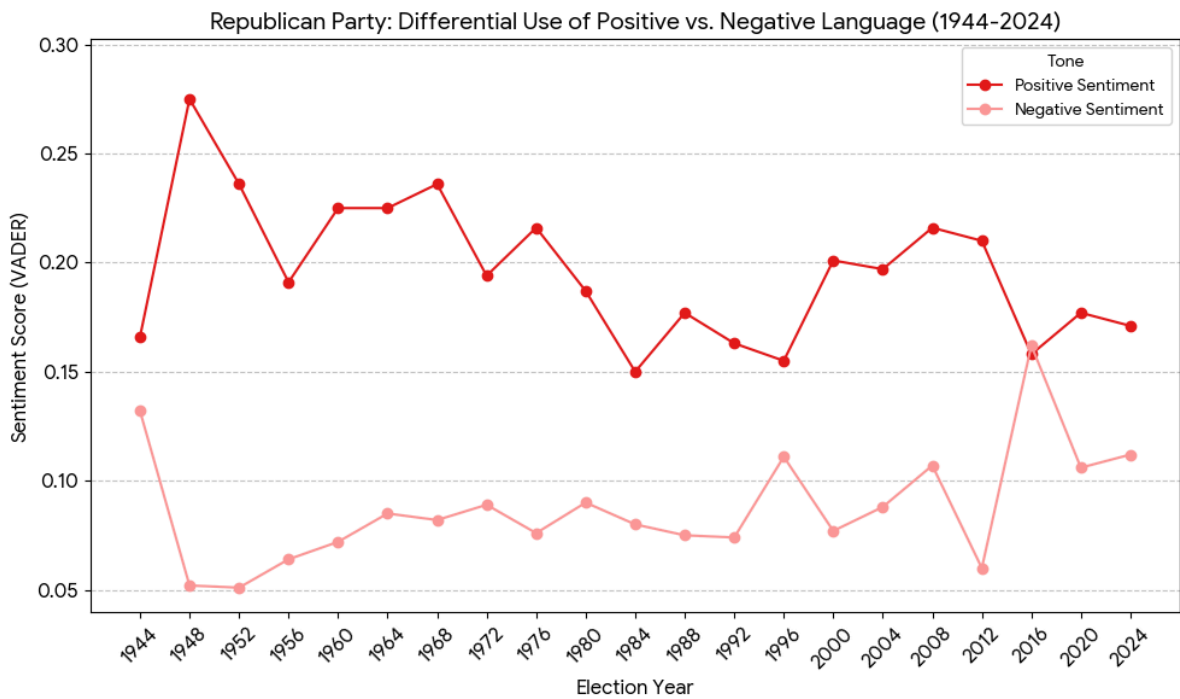


Figure 2.2 – Republican Party: Differential Use of Positive vs. Negative Language (1944-2024)

Key Evaluation Criteria

- Originality & Significance
 - Does the paper contribute new insights or perspectives to the field?
 - **Slightly**
- Clarity & Structure
 - Is the argument well-organized and easy to follow? Are ideas clearly presented?
 - **7/10**
- Use of Evidence & Research Methods
 - Are sources appropriately cited? Is their methodology sound and well-explained?
 - **5/10**
- Engagement with Literature
 - Does the paper demonstrate an understanding of relevant research in the field? Do they acknowledge known results and connect their findings well to them?
 - **3/10**
- Grammar & Language
 - Is the writing clear and professional? Minor grammatical and stylistic errors should be noted, but they should not be the main focus of the review.
 - **10/10**

Feedback

- Specific & Actionable Suggestions
- Encourage Revisions
 - What revisions are needed
 - What are they
 - Why are they needed
 - How might they be addressed

The methods employed in this paper belong to the field of Natural Language Processing for Computational Social Science. Overall, it is a good topic area with a nice dataset selection and methods.

The paper, while having computational elements, is written more like a history paper. A majority of the citations are historical with a bulk of the discussion section describing historical events. I would suggest moving the historical elements to the introduction as a means of providing motivation for this work. The rest should focus computationally with slight nods to historical elements where appropriate (paragraph max).

The introduction does not explicitly state the hypothesis. Therefore, it is difficult to grasp the novelty of this work in comparison to other NLP for CSS papers. What is the unique approach of this work? Have others analyzed this dataset before? What does this research find or investigate differently?

There is no “related works” or “background” section after the introduction. This **must** be included to demonstrate an understanding of other works in this field. Specifically computational social science papers, such as *Fightin’ Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict*. At least 10 citations in this section.

The methodology section is typically the second or third longest portion of the paper, with results and discussion coming ahead. Each experiment needs to be explicitly mentioned in the methodology section, with the choice of metrics and methods justified (citing previous work). What is the purpose of each experiment? How does it tie into your hypothesis? Can you provide a visual “methodology” that shows the process of speech □ text representation □ feature extraction (TF-IDF and Word embeddings) □ latent space representation □ Analysis Approaches (**Jaccard / Cosine Similarity / Etc.**). Reference other papers with “Visual Abstracts” or “Graphical Abstracts”. The Jaccard Index is mentioned and never used in the rest of the paper.

Within the discussion section, as a computational paper, this should include justifications and analysis of findings. Describing political characters and their style is interesting motivation. This can go in the introduction – providing a reason for why **computationally** we would expect to see differences statistically, etc. Discussion section needs to include analysis on the results from a computational side and how it represents the historical trends at that time, without going into incredible detail about the history. Discussion must include limitations of the work and how future researchers can expand on it.

Proposed method updates and experiments (keep what is in the paper – these are supplements).

- What are the most common phrases from electoral candidates over time per party? Include some visual representation (such as a word cloud) or a table.
- Intraparty embedding / vector similarity across time.
 - Compare the representation of the first speech to the representation of each speech after that. How does the cosine similarity change over time from the original? Is it in line with your historical analysis?
- Interparty embedding / vector similarity across time

- Compare the representation of the first speech for one party to the representation of each belonging to the other party.
- Provide tables for the intraparty and interparty experiments above, reporting the **Jaccard Index**, Cosine Similarity (TF-IDF), Cosine Similarity (word embedding)
- Visualize the TF-IDF vectors within the latent space and color them by year / party.
- Investigate the latent space representation with other (more contemporary) methods of feature extraction, such as word embeddings from large language models. What trends do you notice across time?

Trying to understand the semantic similarity graph is difficult because there is no clear experiment explained. The work is comparing the similarity between the candidates for each party? If so, explicitly list this in the methodology.

Also, your average reader is not going to understand the distinction between a cosine similarity of 0.6 to that of a 0.2. Can you provide a blurb from each speech (2 from 0.6 similarity score and 2 from 0.2 similarity score only) so the reader can “see” the difference?

Include graphics directly in the space mentioned and not in the appendix. Appendix should be only for overflow, supplemental items.

Overall

- Accept as is (very unlikely)
- Accept with minor revisions (acceptance conditional on satisfactory minor revisions)
- Accept with major revisions (acceptance conditional on satisfactory major revisions)
- Revise and resubmit (major revisions needed, acceptance not guaranteed)
- Reject (with constructive feedback provided)

Great work on this paper, which uses the semantic similarity and sentiment analysis of presidential nomination speeches to identify ideological differences between political parties in the United States over time.

This paper is relevant and timely, and significant in the fields of political science and linguistics. This paper fulfills the requirements of originality and significance, as well as clarity and structure. In terms of use of evidence and research methods, the Abstract identifies a combination of methods, describing analysis of the historical and political contexts. However, it seems the former takes up the majority of the analysis in the paper, with 2 brief sections on historical and political analysis.

For example, in the Results section, you describe 3 different epochs across time. Perhaps you could return to those epochs in the Discussion section to provide more historical and political context.

Also, rather than just discussing the lowest and highest cosine similarity scores, provide an analysis of each of the epochs, or each of the presidencies from 1948 until the present. Or, a decade-by-decade analysis of the last 50 years of presidencies might be a more reasonable workload. You might decide instead of a broad analysis of speeches beginning in 1948, you might choose a shorter time frame to engage in deeper analysis.

Also, some thoughts to consider: for each speech, where did they take place: in red states, blue states, or swing states? Does the location of the speech have an impact on the language used? The locations for each speech are listed in the database that you used, alongside the dates, which span from May through September.

- Are there differences in language based on the month? For example, more positive patriotic speeches in July around Independence Day?

Another thought to consider: is there more to “feelings” than just positive vs negative?

The intro does a good job using examples of less reliance on words like “data” and more reliance on “feel.” But is it possible to feel something beyond pos/neg? Are there neutral feelings, like the French saying “comme ci comme ca”? Both the Intro and the Conclusion sections need to be expanded.

In the intro, write a bit about the anatomy of a speech, and how any speech (not just political ones) will follow certain formats and emotional cues. Perhaps consider a different frame for your analysis, such as comparing nomination speeches to inauguration speeches, or winning candidates only, etc.

Decision: This paper is ready to be accepted for publication, with major revisions.

Semantic Polarization in U.S. Politics: A Computational Analysis of Republican and Democratic Speeches

[Author names redacted]

[Author schools redacted]

ABSTRACT

This research investigates the evolution of U.S. political rhetoric over the past eighty years, focusing on how thematic priorities and emotional framing have influenced partisan discourse. The study applies computational methods, including semantic similarity and sentiment analysis, to presidential nomination speeches from 1944 to 2024, with the aim of understanding patterns of ideological alignment, divergence, and rhetorical strategy. Beyond identifying these patterns in party messaging, the research examines how differences in thematic emphasis and emotional tone shape voter perception, engagement, and political polarization. By combining quantitative linguistic analysis with historical and political context, the study provides insight into the mechanisms through which political language both mirrors and reinforces broader social and ideological divisions in the United States.

INTRODUCTION

Political discourse in the United States has undergone significant transformation over the past century, affecting not only policy content but also rhetorical style, emotional tone, and approaches to persuasion. Recent empirical research indicates that political language is shifting away from evidence-based reasoning toward more emotionally charged, subjective rhetoric, with important implications for democratic governance.

An essential part of understanding political rhetoric is recognizing that all speeches — whether political, academic, or ceremonial—follow a structured format designed to achieve specific communicative goals. Across contexts, public speaking pedagogy emphasises that a speech typically consists of three core components: an introduction that captures attention and establishes purpose, a body where the substantive arguments and evidence are presented in logical progression, and a conclusion that reinforces the key message and leaves a lasting effect on the audience. This structural organisation supports clarity,

coherence, and persuasion, and is central to effective communication in diverse domains (Language Through Literature, 2024).

Beyond structural elements, speeches also deploy rhetorical appeals that blend credibility, logic, and emotion to engage audiences. Classical rhetorical theory, rooted in Aristotle, highlights *ethos* (the speaker's credibility), *logos* (logical argumentation), and *pathos* (emotional resonance) as foundational to persuasive discourse; these appeals continue to inform modern speech analysis and illustrate how speakers adapt style and content to influence interpretation and judgment (Robinson & Jerskey, 2021). This framework not only applies to individual speeches but also enables meaningful comparison across genres of political oratory—for instance, contrasting nomination acceptance speeches with inauguration addresses reveals differences in purpose, audience engagement, and emotional framing, with acceptance speeches often designed to mobilise partisan support and inaugurals shaped to articulate vision and national unity (Koller & Semino, 2019).

This shift in “feeling” is not just a move toward extreme positivity or negativity. As the French expression “comme ci, comme ça” suggests, most of human sentiment stays in the middle ground – a state of calculative neutrality. Computationally, this neutrality acts as a proxy for policy-heavy discourse. In the 1940s and 50s, a high ratio of neutral scores in VADER reflected a focus on administrative governance that surpassed partisan emotion.

A study published in *Nature Human Behaviour* (2025) analyzed over 8 million congressional speeches from 1879 to 2022 (Bhatia, Lewandowsky, & Politzer, 2025). The results show a decline since the 1970s in the use of fact-based language, such as “data,” “analysis,” “evidence,” and “investigation,” alongside a rise in opinion-based expressions, including “point of view,” “believe,” “common sense,” and “feel.” The authors argue that this move away from evidence-based rhetoric correlates with increasing political polarization, reduced legislative productivity, and growing socio-economic inequalities.

This trend has coincided with broader phenomena, including increasing political polarization, declining legislative cooperation, and rising socio-economic inequality. Empirical studies suggest that these patterns are linked to changes in political language, which has become more emotionally charged and less grounded in shared facts or evidence. Research on partisan speech over the past 130 years indicates that the phrasing and rhetorical markers associated with each major party have grown increasingly distinct and divergent. This growing divergence reflects not only differences in policy but also strategic choices in how parties communicate, reinforcing partisan identities and shaping public perception.

Furthermore, these shifts raise critical questions about the health of democratic discourse. When political communication moves away from shared reference points and parties adopt increasingly divergent rhetorical registers, opportunities for consensus or mutual understanding may diminish. Scholars have described this process as Affective Polarization, in which partisans not only disagree on issues but also increasingly distrust and dislike members of the opposing party (Gidengil & Stolle, 2022). Such polarization is part of a broader transformation in U.S. political communication, with implications for governance, civic engagement, and social cohesion.

This research tests the hypothesis that U.S presidential rhetorics have undergone significant semantic drift, where the lexical overlap between the party's current platform and historical baseline minimizes over time, while opinionated deviation increases. By applying a NLP pipeline to a corpus spanning 80 years, this study investigates whether these drifts are outcomes of historical epochs or variable to granular metadata such as the location of the convention or temporal timing.

Unlike previous works that focus on keyword frequency or specific policies, this research adopts a longitudinal vector-space analysis. The unique contribution of this research lies in the comparative quantification methodology. Unlike previous NLP studies that measure similarity in a vacuum, this paper introduced a synchronous pairwise analysis.

The analysis reveals that partisan divergence is not just cyclical but represents a statistically significant departure from the 'floor' of shared lexicon, an observation that contradicts more conservative models of rhetorical consistency.

RELATED WORKS

The intersection of computational linguistics and political science has produced a robust body of literature focused on identifying and quantifying ideological shifts. This study situates itself with three primary pillars: historical language analysis, semantic similarity modelling, and sentiment analysis.

Recent scholarship has increasingly utilized large-scale datasets to track the evolution of political discourse over centuries (Aroyehun et al., 2025) (Ash, Morelli, & Vannoni, 2024) demonstrate a shift in evidence-based reasoning to intuition-based decision making in over 8 million congressional speeches, a trend they associate with rising polarization. Likewise, researchers have utilized (Card et al., 2022) 140 years of U.S political addresses to show immigration framing has become more positive. This study builds on these frameworks by

focusing on the critical environment of the presidential nomination speeches from 1944 to 2024.

Furthermore, the quantification of partisan differences relies on high-dimensional text-as-data methods. Researchers have developed (Gentzkow, Shapiro, & Taddy, 2019) metrics for partisan classification, highlighting that political speeches can be classified by party with a high level of accuracy due to lexical divergence. Adding on, this is further expanded when a 2020 review of computational text analysis gives the theoretical justifications for applying TF-IDF, word embeddings to social sciences (Gentzkow, Kelly, & Taddy, 2019). Moreover, the application of word embeddings and Cosine Similarity to quantify ideological placement has been validated (Rheault & Cochrane, 2020) and this study uses Jaccard score and TF-IDF Cosine Similarity to measure the degree of “thematic alignment” versus “ideological siloing” between candidates.

Beyond thematic content, the emotional valence of political speech is a primary driver of polarization. A 2019 review synthesizes research on affective polarization, where the partisans exhibit emotional dislike beyond just policy disagreements (Iyengar et al., 2019) and has provided the theoretical grounding for the findings on emotional framing in nomination speeches, emphasizing how rhetorical strategies fuel polarization in U.S politics over time. The use of VADER (Valence Aware Dictionary and Sentiment Reasoner) has become a standard of measuring emotional intensity in politically charged texts (Hutto & Gilbert, 2014). Recent uses of VADER to State of the Union addresses (Gustafsson, 2025) and candidate communications (Tavits & Potter, 2022) suggest that emotional tones may lead to opinionated engagement.

METHODOLOGY

The research presented in this paper employs a quantitative approach to analyze political polarization through the lens of language. The methodology is designed to collect, clean, and analyze a large corpus of political discourse using advanced natural language processing techniques. This section details the steps taken to construct the analytical framework.

Data Collection

The corpus for this study was sourced from the American Presidency Project (APP) at the University of California, Santa Barbara. This archive provides a comprehensive collection of presidential public documents, including speeches, letters, and conference transcripts. The dataset selected for this research consisted of convention speeches, specifically those accepting presidential nominations, from 1944 to 2024.

Once the raw data was collected, it underwent a preprocessing pipeline to prepare the texts for linguistic analysis. The speeches were stored as .docx files and thoroughly cleaned to ensure accurate results. The first step removed capitalization and punctuation. All characters were then converted to lowercase to ensure consistency within the document. Next, stop word filtering removed common words with minimal semantic value, leaving only terms with true meaning. The final step was lemmatization, which converted words to their base form (e.g., “running,” “ran,” and “runs” all become “run”). This process grouped semantically equivalent terms, which was essential for accurate embeddings and similarity analysis.

Algorithmic Formalism Explained

1. Jaccard Similarity

Jaccard Similarity (J) measures the overlap between the unique vocabulary of two datasets, in this case, the Democratic and Republican speeches. This metric quantifies the ratio of shared unique words relative to all unique words in both speeches. Let V_D represent the cleaned and lemmatized words in the Democrat dataset, and V_R the same for the Republican dataset.

The Jaccard Similarity is defined as the size of the intersection divided by the size of the union.

$$J(V_D, V_R) = \frac{|V_D \cap V_R|}{|V_D \cup V_R|}$$

The score ranges from 0, indicating very high vocabulary divergence, to 1.0, indicating nearly identical vocabulary with minimal divergence.

2. Cosine Similarity on TF-IDF Vectors

Semantic similarity is assessed using Cosine Similarity applied to document vectors weighted by Term Frequency-Inverse Document Frequency (TF-IDF). This method emphasizes words that are unique to one document while down-weighting common words shared between both speeches.

2a. TF-IDF Weight Calculation

The weight $W_{t,d}$ of any pre-processed term t in a speech d is the product of its Term Frequency (TF) and Inverse Document Frequency (IDF):

$$W_{t,d} = \text{TF}(t, d) \times \text{IDF}(t)$$

TF is the raw count of the term in the document, with higher values indicating centrality to that speech.

IDF measures the term's rarity across the two-speech corpus ($N = 2$) using a smoothed formula:

$$\text{IDF}(t) = 1 + \ln\left(\frac{N}{\text{DF}(t)}\right)$$

This is the formula when $\text{DF}(t)$ is the number of speeches (one or two) containing the term t . Terms that appear in only one of the two speeches receive the highest weight, amplifying their impact on semantic comparison.

2b. Cosine Similarity Formula

After vectorization, Cosine Similarity measures the cosine of the angle between the two speech vectors. The result is calculated as:

$$\text{Sim}_c(V_D, V_R) = \frac{V_D \cdot V_R}{\|V_D\| \cdot \|V_R\|}$$

The resulting score ranges from 0 to 1.0, where 1.0 indicates highly similar speech content and word usage.

3. VADER Sentiment Analysis

VADER (Valence Aware Dictionary and Sentiment Reasoner) is a rule-based sentiment analysis tool originally designed for social media text, though effective for speeches due to its handling of intensifiers. VADER provides four scores for the text as a whole:

- Positive (pos): proportion of text with a positive tone.
- Neutral (neu): proportion of text with a neutral tone.
- Negative (neg): proportion of the text with a negative tone.
- Compound (compound): a single, non-proportional score summarizing overall sentiment strength and direction, ranging from -1 (extremely negative) to $+1$ (extremely positive).

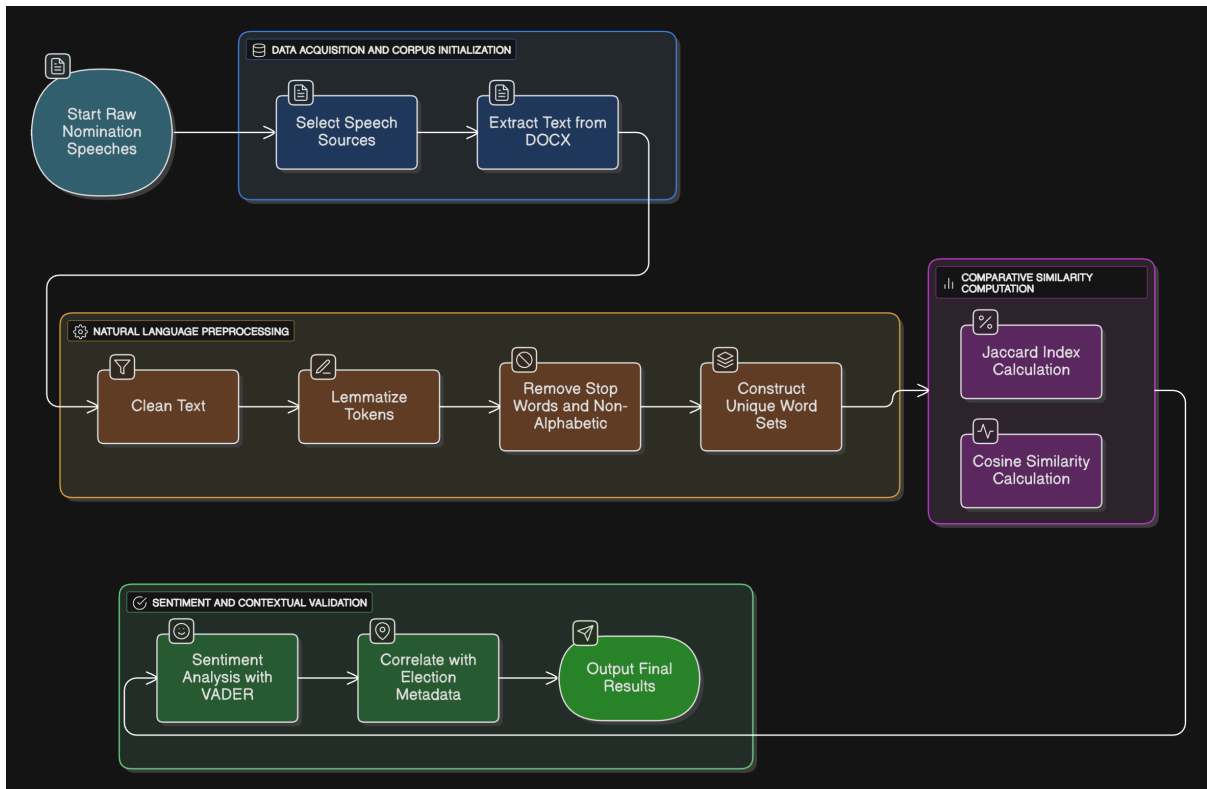


Figure 1 – Methodology Flowchart

The use of Term Frequency-Inverse Document Frequency (TF-IDF) weighting is key to isolating distinctive partisan “signals” from the generic linguistic “noise” present in presidential speeches. Politically charged speeches are inherently dense with boilerplate language – high-frequency terms such as “America”, “people”, and “future” that appear across every speech regardless of the party. A simple word frequency count would result in synthetically high similarity scores, masking the true ideological differences across different candidates. TF-IDF addresses this by mathematically penalizing terms that occur frequently across the speech, using Inverse Document Frequency, while amplifying the terms that are central to the candidate’s platform, using Term Frequency, and this results in the Cosine Similarity to reflect a meaningful thematic alignment and allows the analysis to identify nuanced shifts in how two candidates frame a similar set of issues in the same election cycle.

While the TF-IDF and Cosine Similarity measure the direction of themes, the Jaccard score is used as the second metric to quantify the raw lexical overlap. In the context of political oratories, Jaccard score serves as a proxy for a shared set of vocabulary. By calculating the ratio of the intersection of unique words to their union, it can be analyzed how two opposing candidates in the same election cycle utilize a similar set of lexicons to describe the bigger

picture. A high Jaccard score suggests that the candidates are debating on shared issues using a similar set of terms; likewise, a lower score highlights semantic decoupling.

In this study, the key analytical unit is the Election pair cycle. For every four-year interval, the Democrat's speech is directly compared to the Republican's speech. This approach isolates the specific rhetorical climate of each election cycle, which ensures that the scores are not skewed by the evolution of language through the decades. By comparing candidates from each election cycle, this study ensures that the two candidates face the same historical events, this reveals a precise measurement of "inter-party" divergence.

RESULTS

This section of the paper presents the empirical findings derived from the analysis of the Democratic and Republican speeches. The core analysis focuses on two comparisons: semantic similarity and emotional valence. The semantic similarity is calculated using the Cosine Similarity of the speeches' aggregated word vectors. The emotional metric is quantified using VADER polarity scores, which capture the emotional valence of each speech. It is important to note that the VADER analysis is conducted on the raw text, whereas the semantic similarity analysis uses lemmatized text.

Semantic Similarity

The data from figure 2 reveals three epochs, each marked by distinct mathematical trends: an initial period of volatility, followed by a trajectory toward convergence, and concluding with steep divergence. The overall mean between 1944 and 2024 is 0.4319, with a notably high standard deviation of 0.1295, indicating significant political shifts over time.

Epoch 1: Early Volatility (1944 – ≈ 1960)

The absolute minimum occurs in 1948 (Truman v. Dewey), with a score of 0.1832—the lowest point on the graph. This reflects maximum thematic divergence, meaning the angle between the two parties' topic vectors was widest, indicating the most dissimilar issue sets in the entire eighty-year span. After 1948, the similarity rose sharply (e.g., 0.4119 in 1956) before falling again, reflecting a highly fluctuating environment as the political center of gravity continued to shift.

Epoch 2: Trend Towards Convergence (≈ 1964 – 1996)

This period spans over three decades and shows a move toward shared thematic ground, despite major events such as the Vietnam War and the end of the Cold War. The scores oscillate (e.g., 0.4014 in 1964, peaking briefly at 0.5369 in 1972, then falling in the 1980s), but the underlying trend is upward. Mathematical estimation returns a slope of 0.00236. This

positive slope confirms a steady increase in semantic similarity, culminating in a peak score of 0.4770, showing that the two parties most closely addressed similar sets of issues during this era.

Epoch 3: Modern Divergence (2000 – 2024)

After 2000, the trend reversed sharply. Using the same slope formula, the average slope is -0.0157 , indicating divergence at a rate more than six times greater than the previous period's convergence. The score reached 0.2301, representing a substantial decline in thematic agreement in contrast with the 2000 peak.

Comparative Sentiment: Differential Use of Positive vs. Negative Language

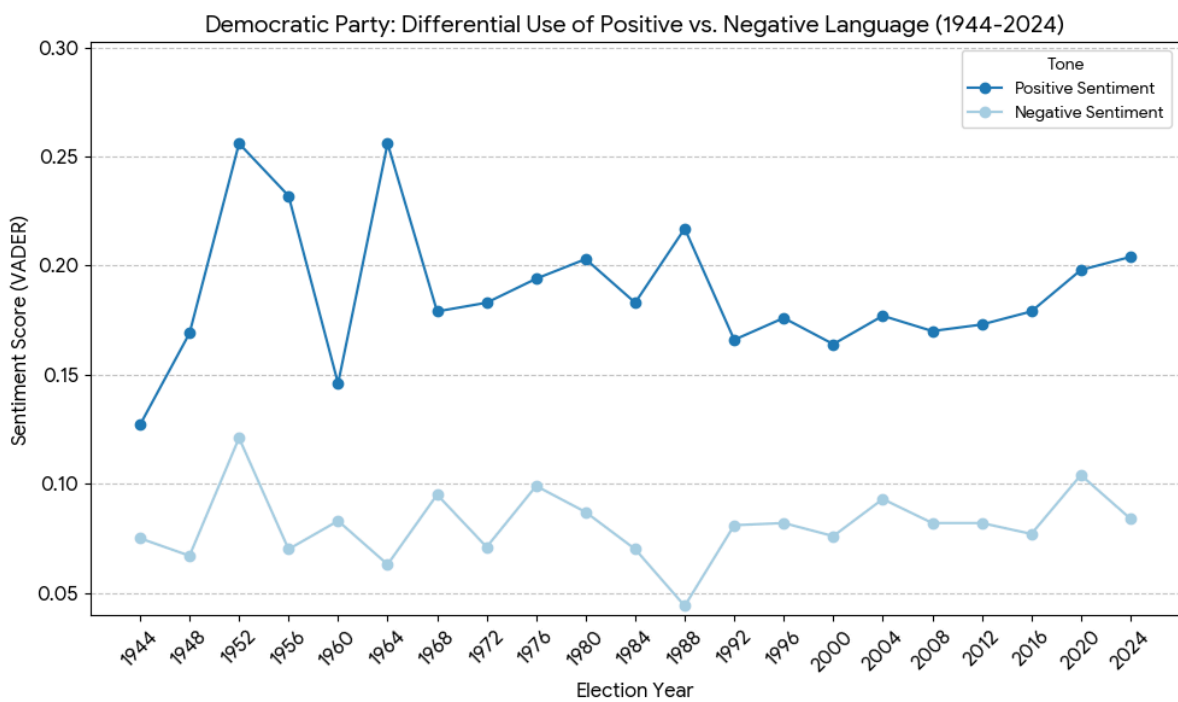


Figure 3.1 – Democratic Party: Differential Use of Positive vs. Negative Language (1944-2024)

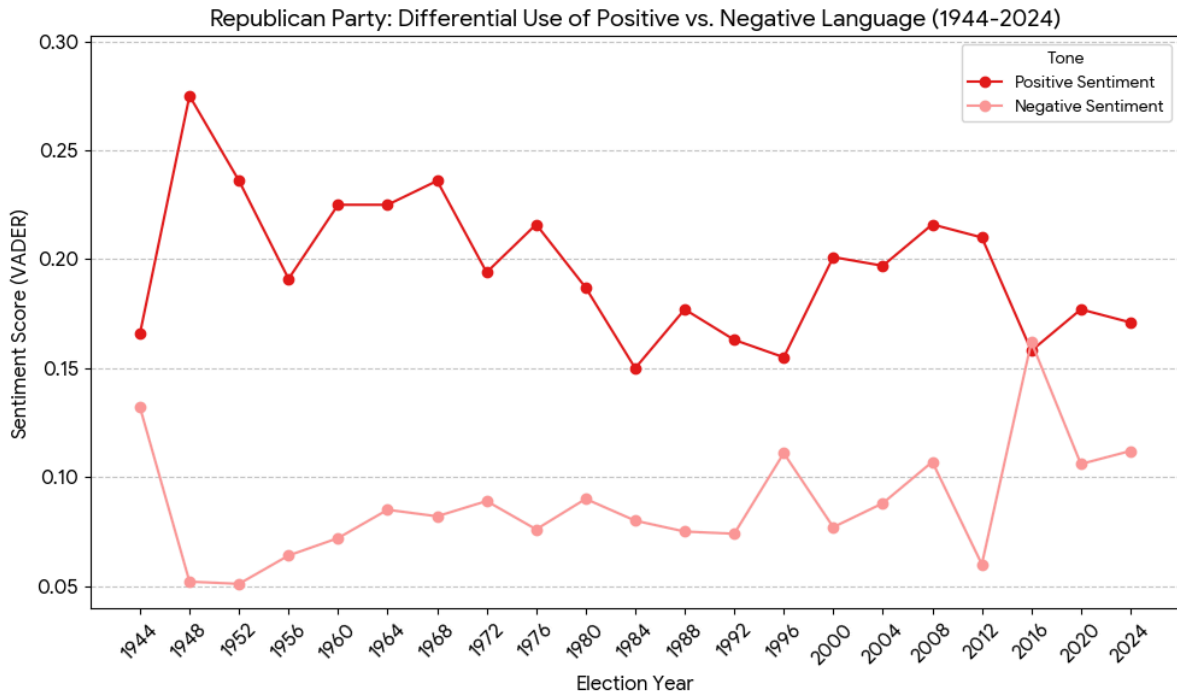


Figure 3.2 – Republican Party: Differential Use of Positive vs. Negative Language (1944-2024)

The time-series analysis of party-specific emotional trends, quantified using VADER sentiment scores (Figures 3.1 and 3.2), establishes the parameters for differential language use by locating the centroid of emotional scores and assessing temporal volatility.

The emotional rhetoric of the Democratic and Republican Parties, as illustrated in Figure 3.1 and Figure 3.2, shows a highly stable positive differential. The key observation is the structural maintenance of this positive differential across the eighty-year period for all but one speech. The Positive Sentiment line consistently maintains a visible, non-zero distance from the Negative Sentiment line. This universal positivity of the Framing Score indicates a bounded emotional space: the emotional centroid remains anchored in the net-positive quadrant. The absence of cross-overs shows that party communication adheres to a minimum positivity threshold that prevents sustained shifts into negative valence. Although both Positive and Negative Sentiment lines trend upward—indicating growing emotional variance—the integrity of the positive differential is preserved.

However, the 2016 election cycle marks an exception to this rhetorical constraint. This year shows an ephemeral valence inversion that reduces the stability of the positive differential. The raw VADER scores reveal an atypical rise in Negative Sentiment to 0.162, coinciding with a contraction of Positive Sentiment to 0.158. The brief moment where the Negative score exceeds the Positive score produces a nominally negative Framing Score.

Furthermore, the Compound VADER score for this year falls to -0.997 , reflecting extreme negative sentiment and demonstrating the party's minimum positive boundary.

Lexical Overlap Measurement: Jaccard Similarity

The analysis of presidential nomination speeches shows a significant fluctuation in raw lexical overlap. By calculating the intersection of unique words, this metric reveals to which extent candidates work within a shared linguistic framework versus isolated one-sided lexicons.

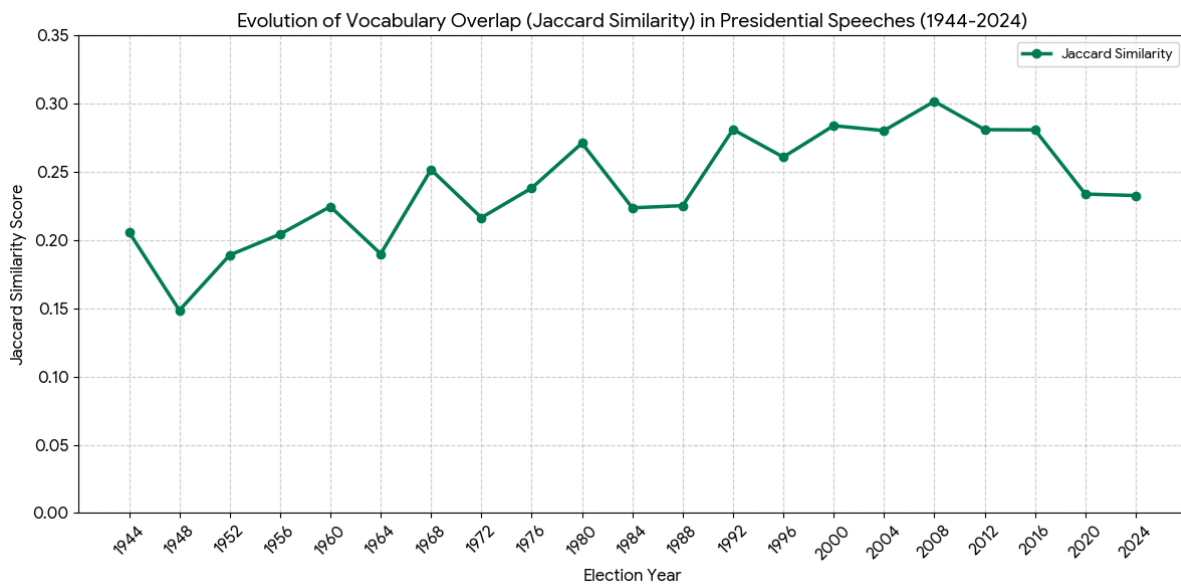


Figure 4 – Evolution of Vocabulary Overlap (Jaccard Similarity) in Presidential Speeches (1944-2024)

The 1948 Lexical Floor: The analysis reveals 1948 (Truman vs. Dewey) as the point of maximum lexical divergence across the 80-year corpus. During this cycle, the Jaccard Similarity reached a historic floor of 0.1485. This reveals that the two candidates had a collective unique vocabulary of less than 15%. The rhetoric isolation of this score is highlighted by Truman’s aggressive focus on “Eight-Nothin’ Congress” and “labor rights”, while Dewey’s strategy focused on consolidating blue states (at the time) in the Northeast through optimising framing of “unity” and “national progress”.

The 2008 Convergence Apex: A distinct reversal in the long-time trend occurred during the 2008 election cycle, where the Jaccard Similarity reached its peak of 0.3016. This convergence was motivated by a high degree of lexical overlap regarding the Great Recession and foreign policy. Particularly, McCain and Obama campaigned extensively in swing states (Florida, Ohio and Pennsylvania) using a more shared vocabulary of “economic crisis” and “reform” to appeal to the median voter, which resulted in the highest recorded lexical commonality in the study.

The data also suggests that the timing of the convention and the location of the speech also acted as variables in terms of the rhetoric distance. Speeches that were delivered later in the election cycle, usually in August or September, tend to show a minimal decrease in the commonality of vocabulary compared to the early-summer addresses. As the election date approaches, the language shifts from a general platform-building to more aggressive contrast-marking. Historically, when conventions took place in critical swing states, there was a visible “thematic pull” that stabilized Jaccard scores as both parties attempted to convey the regional economic concerns. However, the post-2008 data suggests that this effect is gradually fading; candidates maintain their partisan lexicon regardless of the location of the convention, prioritizing national audiences over regional interests.

The fluctuations in Jaccard scores have been persistent, which confirms that political polarization is fundamentally linguistic. Even during years of high positive sentiment – such as 1956 – the shared unique vocabulary has remained low, at 0.2043, suggesting the vocabulary of priority has stayed divided.

DISCUSSION

This section of the paper interprets the empirical results of the semantic similarity scores and the emotional valence scores of the Democratic and Republican convention addresses. First, epoch wise explanations of Cosine Similarity scores are illustrated by outlining the political climate of each period, giving brief context on both party platforms, and assessing how closely the nominees’ convention speeches align with their respective platforms, thereby explaining their points of convergence or divergence. Second, the results of the Jaccard Similarity are qualitatively validated by interpreting the observed patterns of lexical similarity. Third, the reasons, implications, and limits behind the overall net positive tone of all speeches and the distinct net negative tone of Trump’s convention address are deciphered using Figures 3.1 and 3.2.

Semantic Similarity

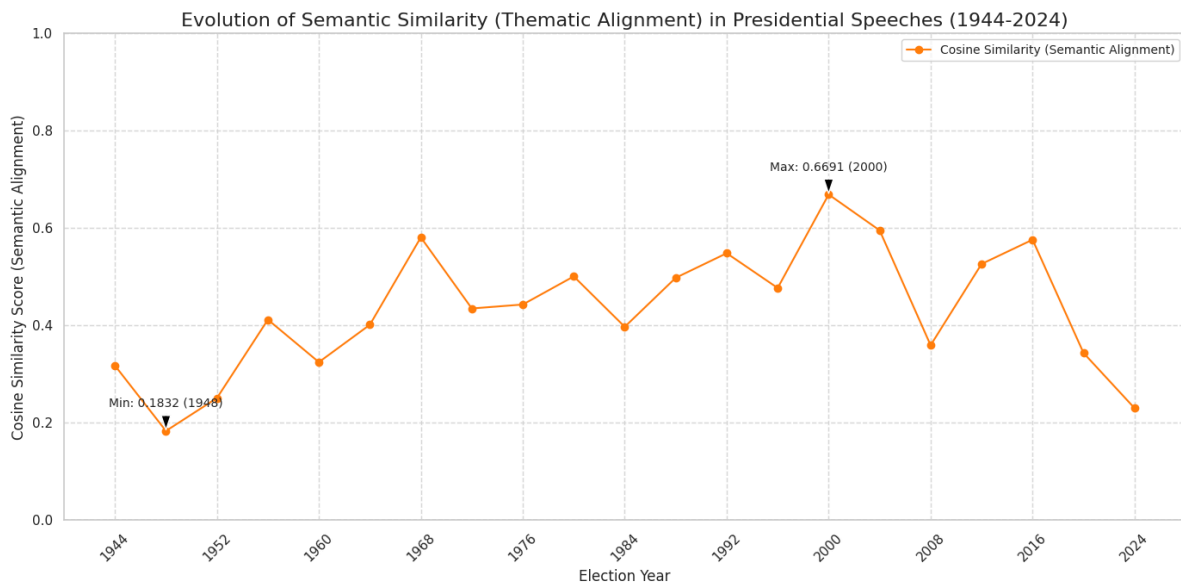


Figure 2 – Evolution of Semantic Similarity (Thematic Alignment) in Presidential Speeches (1944-2024)

Epoch 1: Early Volatility (1944 – ≈ 1960):

Immediately after World War II, the United States was tasked with transitioning from a wartime to a peacetime economy. This rapid shift—from military production to civilian industry, and from full employment to managing inflation and labor unrest—produced an unstable domestic environment in which political rhetoric lacked consistent framing. At the same time, the emergence of early Cold War tensions with the Soviet Union introduced new foreign-policy priorities that were not uniformly emphasized across party lines. These economic uncertainties, coupled with uneven postwar prosperity and emerging social change, contributed to a fragmented political landscape in which shared thematic ground between parties was limited.

Democratic nomination speeches during this period placed strong emphasis on domestic critique and reform. Candidates often directly confronted Republican platforms and advocated expansive federal intervention to address issues such as unemployment, housing shortages, consumer protection, and economic instability. The global context was often framed in explicitly moral terms, linking domestic justice to international peace and economic cooperation. These speeches were typically dense with policy detail, outlining specific proposals across a wide range of areas including employment, wages, agriculture, education, civil rights, environmental conservation, and federal aid programs. (Peters & Woolley, n.d.)

Republican nomination speeches, by contrast, relied more heavily on broad themes of leadership, national security, and future optimism. Rather than enumerating extensive

domestic policy programs, Republican candidates emphasized abstract ideals such as freedom, responsibility, peace, and national strength. Their rhetoric focused on America's role in the world and its long-term destiny, often prioritizing general principles over concrete legislative agendas. Their platforms in this period similarly articulated commitments to defense, economic growth, conservation, and civil rights in broad, value-oriented language. (The American Presidency Project, n.d.)

These stylistic and thematic differences help explain the pronounced divergence observed in this epoch. Democrats' reliance on detailed reform language introduced a wide array of specific policy themes that were largely absent from Republican speeches, while Republicans' preference for generalized rhetoric centered on unity and strength reduced thematic overlap between the parties. Across individual election cycles, the range and emphasis of topics also shifted. In 1948, Democratic rhetoric foregrounded civil rights, labor issues, and domestic economic concerns, while Republicans reflected contrasting framing of issues. By 1952, Eisenhower's nomination speech focused heavily on leadership and peace, diverging from policy-heavy Democratic discourse. In 1960, John F. Kennedy's acceptance speech blended a forward-looking vision with global change, once again producing thematic vectors distinct from those of Republican counterparts. (The American Presidency Project, n.d.)

Epoch 2: Trend Towards Convergence (≈ 1964 – 1996)

Across the three decades, several core issues consistently appeared in the rhetoric of both major parties, creating a degree of issue-level overlap despite opposing ideological framing. The Vietnam War, in particular, became a central political concern due to its profound domestic impact, elevating foreign policy to a dominant position in national discourse. Civil rights and social equality also emerged as defining issues, especially following the Civil Rights Act of 1964, which accelerated political realignment and forced parties to articulate contrasting positions on race, equality, and voting rights. Alongside these, broader economic and cultural concerns—including domestic welfare and social programs, national security and Cold War strategy, and questions of government reform and public trust—remained persistent elements of political debate. (Peters & Woolley, n.d.)

Despite addressing key areas of discourse, the parties adopted divergent rhetorical and policy stances. On the Vietnam War, Democratic platforms increasingly called for de-escalation, peace negotiations, and, in some cases, the end of conscription, with 1968 highlighting significant internal divisions within Democratic rhetoric itself. Republican speeches, by contrast, emphasized law and order, military strength, and a firm approach to

foreign commitments, particularly under the leadership of Nixon and Reagan. In response to civil rights agitations, Democrats progressively integrated themes of social justice, anti-discrimination, and expanded political participation into their platforms. Republicans, meanwhile, shifted toward rhetoric centered on states' rights, individual liberty, and market freedom, particularly as conservative leadership consolidated and the Southern Strategy took shape. Across economic and cultural issues more broadly, Democrats framed their arguments around economic security and social protection, whereas Republicans emphasized market opportunity and personal responsibility. (Peters & Woolley, n.d.) (The American Presidency Project, n.d.)

Overall, this epoch exhibits greater thematic similarity than the previous epoch, as both parties shared the similarity of addressing specific national issues even while offering contrasting policies. This shared engagement helps explain the oscillatory pattern observed during these decades: although the specific issues evolved significantly over time, each election cycle featured both parties actively engaging with the same central topics in public discourse.

Epoch 3: Modern Divergence (2000 – 2024)

Epoch 3 reflects differences in both core national concerns and in the way parties framed their agendas. Together, these shifts contributed to reduced thematic overlap across 2 decades, despite starting at the point of maximum thematic convergence in 80 years.

Differences across key issues emerged most sharply in the areas of healthcare and cultural rights. Democratic platforms repeatedly placed healthcare access, public health, and social safety nets at the center of their agendas, emphasizing protections for the Affordable Care Act and addressing post-pandemic health challenges in 2020. Republicans, by contrast, moved away from comprehensive healthcare policy articulation, especially in 2020, where they did not release a full new platform, opting instead for a brief resolution endorsing the incumbent administration's priorities. This structural absence of detailed healthcare policy language reflects a broader shift away from issue-specific framing. Similarly, Democrats expanded the language of social rights over time, particularly with respect to LGBTQ+ equality and civil liberties from 2012 onward, embedding these issues as core components of their platforms. Republican rhetoric increasingly concentrated on cultural and identity-based issues such as border security, immigration enforcement, and law-and-order themes, which became central to party identity in the 2010s and 2020s, including an emphasis on strict immigration control and election integrity. (Peters & Woolley, n.d.)

In other domains, both parties addressed central issues with diametrically opposing views. On economic matters, Republican platforms from Bush through Trump emphasized tax cuts, deregulation, and free-market principles, with later rhetoric linking tariffs and immigration policy to economic nationalism under the “America First” framework. These platforms consistently prioritized lower taxes and a reduced role for government. Democratic platforms, from Gore through Biden and Harris, focused on economic fairness and middle-class security, advocating investment in infrastructure, healthcare, and economic relief, particularly after the 2008 financial crisis and the COVID-19 pandemic. Democratic economic language explicitly connected fiscal policy to equity and social support. In foreign policy, Democrats expressed commitments to international cooperation and alliance-building, including support for global institutions and re-engagement with agreements such as the Paris Climate Accord. Republicans, particularly under Trump, emphasized national sovereignty, border security, and defense priorities framed through an America First lens rather than multilateral engagement. (Peters & Woolley, n.d.) (The American Presidency Project, n.d.)

These stylistic differences were reinforced by contrasting communication strategies. Democratic platforms consistently engaged with a broader and more detailed set of policy issues, while Republican platforms became increasingly compact and selectively focused. This narrowing of issue articulation reduced thematic overlap between the parties, contributing to greater separation in topic space and reinforcing the pattern of modern rhetorical divergence.

Quantitative Analysis of Longitudinal Semantic and Lexical Drift

Era	Key Thematic Phrases (Democrats)	Key Thematic Phrases (Republicans)	Analysis
Post-War (1944-1956)	<i>Common man, labor, social security, prosperity</i>	<i>Spiritual unity, free men, faith, liberty</i>	Low Jaccard (0.14-0.18): High ideological divergence.
Cold War (1960-1988)	<i>New frontier, great society, justice, world</i>	<i>Peace through strength, tax cuts, freedom</i>	Moderate Jaccard: National security acts as a unifying anchor.

Modern (1992-2008)	<i>Working families, healthcare, middle-class, jobs</i>	<i>Small business, opportunity, mortgage, security</i>	High Jaccard (0.30): High lexical overlap due to shared economic crises.
Current (2012-2024)	<i>Democracy, rights, future, diversity</i>	<i>Make America Great, borders, strength, fight</i>	Declining Jaccard: Re-emergence of partisan "siloed" language.

Table 1 - Thematic Frequency Table (1944–2024)

The frequency analysis reveals a significant shift in how both parties construct their speeches. While the lemma “America” stayed the most frequent term across the corpus at 1.34%, its usage density varied depending on the political climate and degree of partisan polarization.

During the mid-20th century, the analysis shows a high degree of thematic overlap, characterized by Cosine Similarity and Jaccard scores. From a computational perspective, this shows that the priority vectors of both the parties were closely aligned. This thematic overlap was likely influenced by existential threats, such as World War II and the onsets of the Cold War, and this resulted in a unified national drive which led to candidates using common vocabulary to describe immediate issues.

Starting in the late 1960s, the results show the beginnings of semantic decoupling. While the Cosine Similarity remained fluctuating, the Jaccard score slowly began minimizing. The divergence highlights a shift in strategy, while still conveying similar vague themes, began utilizing specific, party-based lexicons. This period reflects the transition from a broadcast model of political communications to a more fragmented approach.

In the modern era, both metrics reached lows, highlighting that the candidates were moving to a state of orthogonality. The analysis reveals that the parties are no longer debating on a shared set of issues; rather, they are operating in their own distinct latent spaces. This trend is evident as it stays constant across multiple election cycles, regardless of the individual candidates’ campaigning styles, indicating a gradual shift towards permanent rhetorical fragmentation.

Comparative Sentiment: Differential Use of Positive vs. Negative Language

An analysis of presidential nomination acceptance speeches over the past eighty years shows that nearly all candidates delivered addresses with a net positive emotional tone, with one exception: Donald Trump's 2016 Republican National Convention speech. Four main reasons explain why every other speech conveyed net positive sentiment.

First, nomination acceptance speeches are moments of celebration, gratitude, and party unity. Candidates traditionally begin with expressions of thanks—using phrases such as “I accept your nomination” or “I would like to thank”—which contain lexicon items tagged as positive in VADER. Second, these speeches are designed to be forward-looking. Candidates outline promises, opportunities, and hope, as the primary goal is to inspire and mobilize voters. Third, nominees address the entire nation instead of narrow voter blocs. They emphasize shared values, the common good, and the American Dream, themes that naturally produce positive sentiment in lexicon-based methods. Finally, sustained use of inclusive pronouns such as “we,” “our,” and “us,” alongside words like “proud,” “honoured,” and “grateful,” helps maintain an overall positive emotional profile. (Peters & Woolley, n.d.)

Positive emotion in acceptance speeches has significant public effects. Inspiring language strengthens enthusiasm among existing supporters and offers undecided voters reasons to align with the candidate or party. Such speeches can increase turnout, volunteering, donations, political discussion, and overall engagement. They also help set the agenda for the election by communicating a candidate's priorities, values, and policy direction, shaping how voters interpret the campaign. When used consistently, positive emotional messaging can strengthen party reputation and build trust, ultimately supporting electoral success. (Tavits & Potter, 2022)

However, this strategy has limitations. Positive rhetoric can appear unrealistic or disconnected when the country is facing a crisis, war, or economic instability. In such times, voters may perceive overly optimistic messaging as incompetent or out of touch. Furthermore, positive emotion does not guarantee political advantage. Research shows that negative emotions spread faster and evoke stronger reactions, which is why fear-based appeals dominate political advertising and social media. Given the current climate of intense polarization and public frustration in the United States, a purely positive speech may not effectively mobilize voters who feel exhausted, angry, and disappointed with the government. (Iyengar, Lelkes, Levendusky, Malhotra, & Westwood, 2019) (Pew Research Center, 2023)

By contrast, Donald Trump's 2016 acceptance address showed a clear net negative sentiment, the only instance of its kind in the history of presidential nomination speeches. The dominant tone of his speech was rooted in fear. He framed the United States as being in

a state of crisis, repeatedly referencing terrorism, violence, and attacks on police. His use of intense descriptors—such as “brutally murdered,” “barbarians,” “savage killers,” “horrible,” “unfair,” “disaster,” “ruins,” and “chaos”—created vivid, violent imagery that contributed to the overall negative emotional score. More importantly, Trump devoted substantial time to detailing the failures of existing policies and explicitly blamed Democratic leaders. He directly criticized Hillary Clinton, Bill Clinton, and Barack Obama for rising crime, terrorism, immigration problems, violence involving undocumented immigrants, corruption, economic stagnation, unemployment, failing infrastructure, weakened military strength, national-security threats, poor trade deals, foreign-policy shortcomings, and political elitism. Although presidents often point to systemic problems as a way of introducing their own plans, Trump’s emphasis on failure was unusually detailed and accusatory. (Peters & Woolley, n.d.)

The public response to the speech was mixed in both the short and long term. Immediately after the conventions, Trump experienced a larger polling “bump” than Hillary Clinton—his support rose by 3 percent, compared to her 2 percent increase. However, because Clinton was already ahead before the conventions, this translated into only a 1-point net gain for Trump, leaving her still in the lead. (Peters & Woolley, n.d.) A Gallup poll found that after the conventions, voters were 15 percent less likely to support Trump and 4 percent more likely to support Clinton. It was the first time since Gallup began asking the question in 1984 that a candidate left a convention with a net negative effect. (Edwards-Lee, 2016)

Despite these shifts, both candidates ended the 2016 campaign with historically low favorability ratings. The year produced two of the least liked nominees in modern election history, suggesting that convention speeches did little to reshape voters’ deeper opinions. Scholars note that conventions can influence short-term enthusiasm, media narratives, and undecided voters, but long-term turnout and final vote decisions depend on broader factors such as economic conditions, campaign activity, and later events. In 2016, the small convention effects did not meaningfully predict the election outcome. Experts therefore view the speech’s direct influence on turnout as limited and overshadowed by subsequent developments. (Norman, 2016)

Qualitative Validation: Interpretability Blurbs of Lexical Similarity

To illustrate the 0.3016 Jaccard score identified in the 2008 election cycle, the following excerpts from Barack Obama and John McCain show a significant convergence of lexical priorities. Despite their ideological differences, both candidates have utilized a nearly identical “lexicon of crisis” to show the immediate economic anxieties of the American crowd.

Snippet A: Barack Obama (Democrat, 2008)

"We meet at one of those defining moments – a moment when our nation is at war, our economy is in turmoil... Tonight, more Americans are out of work and more are working harder for less. More of you have lost your homes and even more are watching your home values plummet. More of you have cars you can't afford to drive, credit card bills you can't afford to pay".

Snippet B: John McCain (Republican, 2008)

"These are tough times for many of you. You're worried about keeping your job or finding a new one, and you're struggling to put food on the table and stay in your home... We're not going to allow that while you struggle to buy groceries, fill your gas tank, and make your mortgage payment"

The following table shows the high-value lemmas that appeared with high-frequencies in both the texts, contributing to the study-wide similarity score:

Keyword	Obama	McCain
Mortgage	"...job that pays the mortgage"	"...make your mortgage payment"
Home	"...lost your homes"	"...stay in your home"
Work/Job	"...out of work... 23 million new jobs"	"...keeping your job... create new jobs"
Economy	"...economy is in turmoil"	"...rescue our economy"

Table 2 – High-value lemmas from the 2008 election cycle

The record low similarity in 1948, of 0.1485, is the result of two candidates addressing the nation from two completely different perspectives: Dewey from a highly philosophical perspective while Truman having a policy-heavy stance.

Snippet C: Thomas Dewey (Republican, 1948)

"Our people... have a yearning to move to higher ground, to find a common purpose in the finer things which unite us. The unity we seek is more than material. It is more than a matter of things and measures. It is most of all spiritual... We have not yet found the spiritual means to put together the world's broken pieces."

Snippet D: Harry Truman (Democrat, 1948)

"The Republican Party... favors the privileged few and not the common everyday man... They proved it by the things they did to the people. Prices have gone all the way off the chart... I am calling this Congress back into session to pass laws to halt rising prices, to meet the housing crisis... and an increase in the minimum wage."

While the 2008 high-similarity apex was driven by a common vocabulary of crisis, the 1948 floor is defined by thematic isolation. The candidates did not agree on what the nation's problems were; therefore, this caused them to have a very small intersection of shared word sets, resulting in the low Jaccard score.

LIMITATIONS AND METHODOLOGICAL CONSTRAINTS

While this study uses a rigorous framework for measuring polarization, several computational limitations must be acknowledged. Firstly, incorporating TF-IDF and Jaccard similarity relies on a “bag-of-words” model, which completely ignores syntax and word order, and most importantly: context. For example, the phrase “I support healthcare reform” and “I oppose healthcare reform” would yield a high degree of lexical similarity while representing completely opposing ideological stances. Second, while VADER Sentiment Analysis is a standard for identifying texts with explicit tones, it will struggle to identify subtle hints of political sarcasm or “dog-whistle” rhetoric, which need a more in-depth understanding than simply rule-based models. Finally, such speeches represent a very specific, formalized genre of political oratory and does not capture the full extent to which polarization occurs which are found in more informal settings, such as social media.

RESEARCH SCOPE

Future studies should transition from static vector representations (TF-IDF) to context-based embeddings using Transformer-based architectures such as BERT (Bidirectional Encoder Representations from Transformers). Unlike the current bag-of-words approach, using the BERT architecture will capture the polysemy of the text, allowing researchers to identify how the meaning of the word changes based on the text around it. This would allow researchers to distinguish between shared vocabulary in opposition compared to the shared vocabulary in agreement. This would provide a metric that would capture the intent of the speaker rather than just the lexical frequency.

Furthermore, to identify the migration of political issues, researchers can deploy pipelines that utilize Dynamic Topic Modeling (DTM). While this study showed that parties are diverging, DTM would allow the tracking of specific topic trajectories over a period. This would reveal which topics are of higher interest to one party, such as the Republican “capture” of Border Security, and how the semantic weight would vary across the different election cycles.

Finally, while this paper used VADER for sentiment analysis, researchers in the future could incorporate Affective Computing to measure rhetorical aggression. Utilizing Emotion AI, researchers could classify the texts into more specific categories such as fear, anger or

hope; this would allow researchers to determine if the decline in similarity is due to emotional framing instead of the policy substance.

CONCLUSION

This analysis of U.S. presidential election speeches from 1944 to 2024 identifies clear patterns in both thematic alignment and emotional framing across time. The semantic similarity results indicate that political rhetoric has moved through distinct phases: an early period of volatility (1944–1960) marked by sharp fluctuations in issue emphasis, a prolonged phase of convergence (≈1964–1996) during which parties addressed substantially overlapping national concerns, and a modern period of divergence (1996–2024) characterized by rapidly increasing thematic distance. The 1948 election reflects an extreme point of divergence shaped by post-war political, economic, and social uncertainty, whereas the 2000 election illustrates high semantic alignment despite policy differences, demonstrating how shared national priorities can persist during times of economic stability and globalization.

The comparative sentiment analysis shows that presidential nomination speeches have historically maintained a net-positive emotional tone aimed at reinforcing party unity and encouraging public mobilization. Deviations from this pattern, such as Donald Trump's 2016 address, reveal how highly negative, fear-based rhetoric can shift short-term public perceptions without necessarily determining final electoral outcomes. Together, these trends show how political communication balances substantive issue framing with the strategic use of emotion.

Overall, the findings highlight the evolving nature of U.S. political discourse. While parties may converge around core national issues during certain historical moments, the methods through which they frame and address these concerns—both semantically and emotionally—can differ substantially. Understanding these dynamics offers insight into historical electoral outcomes and broader patterns of political polarization and voter engagement in the United States.

REFERENCES

1. Bhatia, S., Lewandowsky, S., & Politzer, G. (2025). *Evidence for increasing intuitive reasoning in U.S. political language*. *Nature Human Behaviour*.
<https://www.nature.com/articles/s41562-025-02136-2>
2. Gidengil, E., & Stolle, D. (2022). Emotions, partisanship, and political behaviour. *Political Behavior*, 44, 1331–1354. <https://doi.org/10.1007/s11109-022-09784-4>

3. Gilder Lehrman Institute of American History. (n.d.). *Postwar politics and the Cold War*.
<https://www.gilderlehrman.org/history-resources/essays/postwar-politics-and-cold-war>
4. TrendSpider. (n.d.). *The post-war recession of 1948–1949*.
<https://trendspider.com/learning-center/the-post-war-recession-1948-1949/>
5. The National WWII Museum. (n.d.). *Episode 5: The strike wave*.
<https://www.nationalww2museum.org/war/podcasts/best-my-ability-podcast/season-2-archive/episode-5-strike-wave>
6. U.S. Department of State, Office of the Historian. (n.d.). *Introduction: 1945–1952*.
<https://history.state.gov/milestones/1945-1952/foreword>
7. Library of Congress. (n.d.). *Post-war United States (1945–1968): Overview*.
<https://www.loc.gov/classroom-materials/united-states-history-primary-source-timeline/post-war-united-states-1945-1968/overview/>
8. Peters, G., & Woolley, J. T. (n.d.). *Party platforms*. The American Presidency Project.
<https://www.presidency.ucsb.edu/documents/app-categories/elections-and-transitions/party-platforms>
9. Miller Center. (n.d.). *Harry S. Truman: Campaigns and elections*.
<https://millercenter.org/president/truman/campaigns-and-elections>
10. Peters, G., & Woolley, J. T. (n.d.). *Convention speeches*. The American Presidency Project.
<https://www.presidency.ucsb.edu/documents/app-categories/elections-and-transitions/convention-speeches>
11. The White House. (2000). *Helping to Build the New Economy*.
<https://clintonwhitehouse5.archives.gov/WH/new/html/20000225.html>
12. Borger, J. (2000, October 17). *Not voting for Gore*. The Guardian.
<https://www.theguardian.com/world/2000/oct/17/uselections2000.usa3>
13. OnTheIssues.org. (n.d.). *2000 Democratic National Platform: Environment*.
https://www.ontheissues.org/Archive/2000_DNC_Platform_Environment.htm

14. OnTheIssues.org. (n.d.). *2000 Republican National Platform: Tax reform*.
https://www.ontheissues.org/Archive/2000_RNC_Platform_Tax_Reform.htm
15. Tavits, M., & Potter, J. (2022). *The Nature and Impact of Emotional Content in Congressional Candidate Emails to Supporters*. Science Direct.
<https://www.sciencedirect.com/science/article/abs/pii/S0261379422000609>
16. Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22, 129–146.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC6690035>
17. Pew Research Center. (2023). *Americans' feelings about politics, polarization, and the tone of political discourse*.
<https://www.pewresearch.org/politics/2023/09/19/americans-feelings-about-politics-polarization-and-the-tone-of-political-discourse>
18. Peters, G., & Woolley, J. T. (n.d.). *The post-convention bounce in voter preference*. The American Presidency Project.
<https://www.presidency.ucsb.edu/statistics/data/the-post-convention-bounce-voters-preference>
19. Edwards-Lee, S. (2016, July 19). *Republican convention: Why some Republicans are not voting for Trump*. TIME Magazine.
<https://time.com/4433348/republican-convention-not-voting-trump/>
20. Norman, J. (2016). *Trump, Clinton finish with historically poor images*. Gallup.
<https://news.gallup.com/poll/197231/trump-clinton-finish-historically-poor-images.aspx>
21. Aroyehun, S. T., Simchon, A., Carrella, F., Lasser, J., Lewandowsky, S., & Garcia, D. (2025). *Computational analysis of US congressional speeches reveals a shift from evidence to intuition*. *Nature Human Behaviour*.
<https://www.nature.com/articles/s41562-025-02136-2>
22. Card, D., Boydston, A. E., Gross, J. H., Resnik, P., & Smith, N. A. (2022). *Computational analysis of 140 years of US political speeches reveals more positive but increasingly polarized framing of immigration*. *Proceedings of the National Academy of Sciences*. <https://www.pnas.org/doi/full/10.1073/pnas.2120510119>

23. Gentzkow, M., Shapiro, J. M., & Taddy, M. (2019). *Measuring group differences in high-dimensional data: Applications to congressional speech and text classification*. *American Economic Review*.
<https://www.aeaweb.org/articles?id=10.1257/aer.20180412>
24. Ash, E., Morelli, M., & Vannoni, M. (2024). *Computational analysis of US Congressional speeches reveals a shift from evidence-based to intuition-based reasoning*. <https://www.nature.com/articles/s41562-025-02136-2>
25. Gentzkow, M., Kelly, B., & Taddy, M. (2019). *Text as data*. *Journal of Economic Literature*. <https://www.aeaweb.org/articles?id=10.1257/jel.20181020>
26. Rheault, L., & Cochrane, C. (2020). *Word embeddings for the analysis of ideological placement in political texts*. *Political Analysis*.
<https://www.cambridge.org/core/journals/political-analysis/article/abs/word-embeddings-for-the-analysis-of-ideological-placement-in-parliamentary-corpora/017F0CEA9B3DB6E1B94AC36A509A8A7B>
27. Hutto, C. J., & Gilbert, E. (2014). *VADER: A parsimonious rule-based model for sentiment analysis of social media text*. *Proceedings of the International AAAI Conference on Web and Social Media*.
<https://ojs.aaai.org/index.php/ICWSM/article/view/14550>
28. Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). *The origins and consequences of affective polarization in the United States*. *Annual Review of Political Science*.
<https://www.annualreviews.org/content/journals/10.1146/annurev-polisci-051117-073034>
29. Tavits, M., & Potter, J. (2022). *The nature and impact of emotional content in congressional candidate emails to supporters*. *ScienceDirect*.
<https://www.sciencedirect.com/>
30. Gustafsson, L. (2025). *NLP analysis of sentiment and rhetorical tone trends in U.S. presidential State of the Union addresses*. DiVA Portal (Master's thesis, Uppsala University).
<https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1980354&dswid=-618>
31. The American Presidency Project. (n.d.). *Convention Speeches*. University of California, Santa Barbara.

<https://www.presidency.ucsb.edu/documents/app-categories/elections-and-transitions/convention-speeches>

32. Language Through Literature. (2024). *Organizing rhetoric: Structuring speeches*. Gender & Language Studies.
<https://gender.study/language-through-literature/organizing-rhetoric-structuring-speeches/>
33. Robinson, M. B., & Jerskey, M. (2021). *Writing, Speaking, and Activism*. OpenStax.
<https://openstax.org/books/writing-guide/pages/19-1-writing-speaking-and-activism>
34. Koller, V., & Semino, E. (2019). *Organizing rhetoric as a political resource: A multimodal analysis of political speeches* (White Rose ePrints). University of Sheffield & University of Lancaster.
https://eprints.whiterose.ac.uk/id/eprint/113234/1/FinalVersion_PDF.pdf

Comment 1: “The introduction does not explicitly state the hypothesis. Therefore, it is difficult to grasp the novelty of this work in comparison to other NLP for CSS papers. What is the unique approach of this work? Have others analyzed this dataset before? What does this research find or investigate differently?”

My Response: Thank you for raising this concern. 3 paragraphs have been added to the introduction section to clarify the hypotheses, its difference from previous works, and its investigative objectives.

“This research tests the hypothesis that U.S presidential rhetorics have undergone significant semantic drift, where the lexical overlap between the party’s current platform and historical baseline minimizes over time, while opinionated deviation increases. By applying a NLP pipeline to a corpus spanning 80 years, this study investigates whether these drifts are outcomes of historical epochs or variable to granular metadata such as the location of the convention or temporal timing.

Unlike previous works that focus on keyword frequency or specific policies, this research adopts a longitudinal vector-space analysis. The unique contribution of this research lies in the comparative quantification methodology. Unlike previous NLP studies that measure similarity in a vacuum, this paper introduced a synchronous pairwise analysis.

The analysis reveals that partisan divergence is not just cyclical but represents a statistically significant departure from the ‘floor’ of shared lexicon, an observation that contradicts more conservative models of rhetorical consistency.”

Comment 2: There is no “related works” or “background” section after the introduction. This must be included to demonstrate an understanding of other works in this field. Specifically computational social science papers, such as Fightin’ Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict. At least 10 citations in this section.

My Response: A 'Related Works' section has been added, with the following citations:

“The intersection of computational linguistics and political science has produced a robust body of literature focused on identifying and quantifying ideological shifts. This study situates itself with three primary pillars: historical language analysis, semantic similarity modelling, and sentiment analysis.

Recent scholarship has increasingly utilized large-scale datasets to track the evolution of political discourse over centuries (Aroyehun et al., 2025) (Ash, Morelli, & Vannoni, 2024) demonstrate a shift in evidence-based reasoning to intuition-based decision making in over 8 million congressional speeches, a trend they associate with rising polarization. Likewise, researchers have utilized (Card et al., 2022) 140 years of U.S political addresses to show immigration framing has become more positive. This study builds on these frameworks by focusing on the critical environment of the presidential nomination speeches from 1944 to 2024.

Furthermore, the quantification of partisan differences relies on high-dimensional text-as-data methods. Researchers have developed (Gentzkow, Shapiro, & Taddy, 2019) metrics for partisan classification, highlighting that political speeches can be classified by party with a high level of accuracy due to lexical divergence. Adding on, this is further expanded when a 2020 review of computational text analysis gives the theoretical justifications for applying TF-IDF, word embeddings to social sciences (Gentzkow, Kelly, & Taddy, 2019). Moreover, the application of word embeddings and Cosine Similarity to quantify ideological placement has been validated (Rheault & Cochrane, 2020) and this study uses Jaccard score and TF-IDF Cosine Similarity to measure the degree of “thematic alignment” versus “ideological siloing” between candidates.

Beyond thematic content, the emotional valence of political speech is a primary driver of polarization. A 2019 review synthesizes research on affective polarization, where the partisans exhibit emotional dislike beyond just policy disagreements (Iyengar et al., 2019) and has provided the theoretical grounding for the findings on emotional framing in nomination speeches, emphasizing how rhetorical strategies fuel polarization in U.S politics over time. The use of VADER (Valence Aware Dictionary and Sentiment Reasoner) has become a standard of measuring emotional intensity in politically charged texts (Hutto & Gilbert, 2014). Recent uses of VADER to State of the Union addresses (Gustafsson, 2025) and candidate communications (Tavits & Potter, 2022) suggest that emotional tones may lead to opinionated engagement.”

Comment 3: Each experiment needs to be explicitly mentioned in the methodology section, with the choice of metrics and methods justified (citing previous work). What is the purpose of each experiment? How does it tie into your hypothesis?

My Response: 2 paragraphs explicitly mentioning the purpose of each experiment have been added towards the end of the ‘methodology’ section:

“The use of Term Frequency-Inverse Document Frequency (TF-IDF) weighting is key to isolating distinctive partisan “signals” from the generic linguistic “noise” present in presidential speeches. Politically charged speeches are inherently dense with boilerplate language – high-frequency terms such as “America”, “people”, and “future” that appear across every speech regardless of the party. A simple word frequency count would result in synthetically high similarity scores, masking the true ideological differences across different candidates. TF-IDF addresses this by mathematically penalizing terms that occur frequently across the speech, using Inverse Document Frequency, while amplifying the terms that are central to the candidate’s platform, using Term Frequency, and this results in the Cosine Similarity to reflect a meaningful thematic alignment and allows the analysis to identify nuanced shifts in how two candidates frame a similar set of issues in the same election cycle.

While the TF-IDF and Cosine Similarity measure the direction of themes, the Jaccard score is used as the second metric to quantify the raw lexical overlap. In the context of political oratories, Jaccard score serves as a proxy for a shared set of vocabulary. By calculating the ratio of the intersection of unique words to their union, it can be analyzed how two opposing candidates in the same election cycle utilize a similar set of lexicons to describe the bigger picture. A high

Jaccard score suggests that the candidates are debating on shared issues using a similar set of terms; likewise, a lower score highlights semantic decoupling.”

Additionally, **‘Trying to understand the semantic similarity graph is difficult because there is no clear experiment explained. The work is comparing the similarity between the candidates for each party? If so, explicitly list this in the methodology’**, this comment has also been addressed in the added paragraphs.

Comment 4: Can you provide a visual “methodology” that shows the process of speech text representation feature extraction (TF-IDF and Word embeddings) latent space representation Analysis Approaches (Jaccard / Cosine Similarity / Etc.). Reference other papers with “Visual Abstracts” or “Graphical Abstracts”.

My Response: ‘Figure 1 — Methodology Flowchart’ has been added to visualize the full methodological process of the chosen methods.

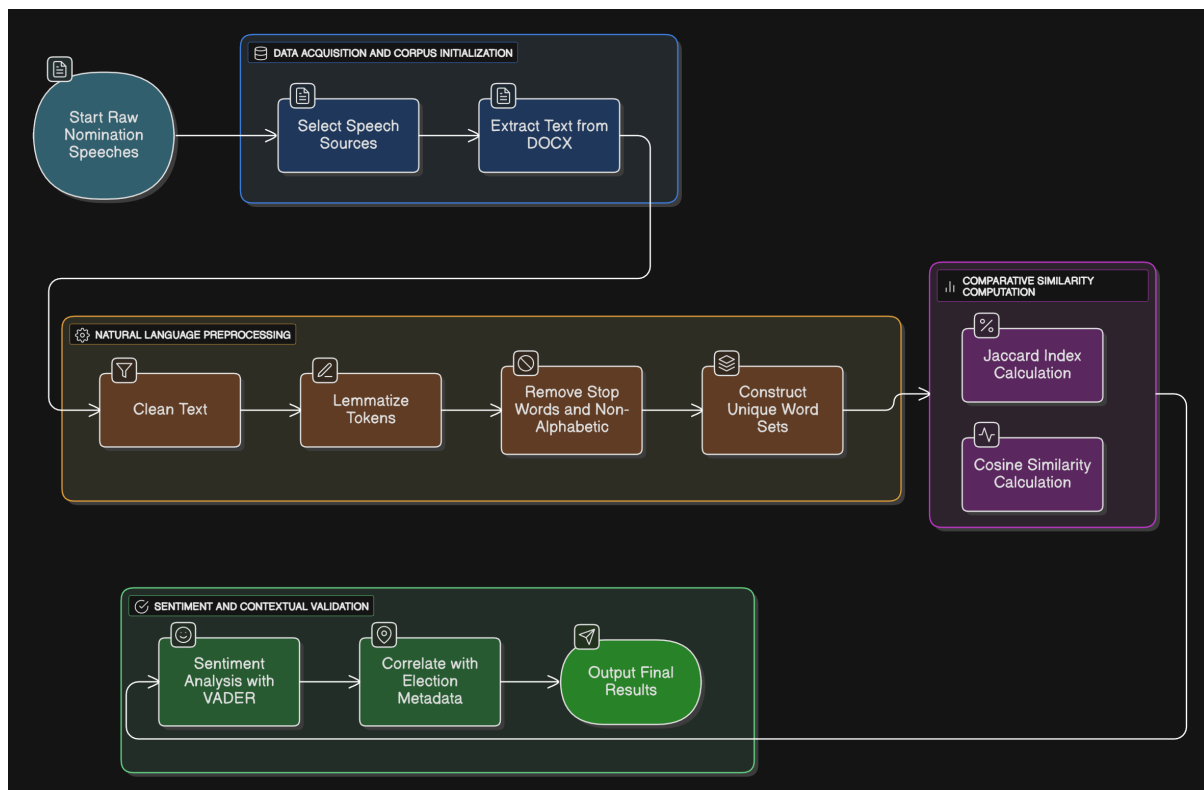


Figure 1 – Methodology Flowchart

Comment 5: Within the discussion section, as a computational paper, this should include justifications and analysis of findings. The discussion section needs to include analysis on the

results from a computational side and how it represents the historical trends at that time, without going into incredible detail about the history.

My Response: The discussion section has been amended accordingly. Previous explanations of the highest and lowest point on the Cosine Similarity Graph have been removed, and epoch wise explanation of the graph has been added. Although there still exist slight historical references to contextualize the nature of the 3 epochs, the historical detail of each analysis has been limited to a large extent. The description of the epochs have been revised to come from a computational perspective, instead of a historical one. The aim of the explanation lies in the use of specific vocabulary and issues referred/not referred to in both parties' nomination speeches as reasons for justifying the variation in cosine similarity scores, across the 80 year span. The revised section is as follows:

“Epoch 1: Early Volatility (1944 – ≈ 1960):

Immediately after World War II, the United States was tasked with transitioning from a wartime to a peacetime economy. This rapid shift—from military production to civilian industry, and from full employment to managing inflation and labor unrest—produced an unstable domestic environment in which political rhetoric lacked consistent framing. At the same time, the emergence of early Cold War tensions with the Soviet Union introduced new foreign-policy priorities that were not uniformly emphasized across party lines. These economic uncertainties, coupled with uneven postwar prosperity and emerging social change, contributed to a fragmented political landscape in which shared thematic ground between parties was limited.

Democratic nomination speeches during this period placed strong emphasis on domestic critique and reform. Candidates often directly confronted Republican platforms and advocated expansive federal intervention to address issues such as unemployment, housing shortages, consumer protection, and economic instability. The global context was often framed in explicitly moral terms, linking domestic justice to international peace and economic cooperation. These speeches were typically dense with policy detail, outlining specific proposals across a wide range of areas including employment, wages, agriculture, education, civil rights, environmental conservation, and federal aid programs. (Peters & Woolley, n.d.)

Republican nomination speeches, by contrast, relied more heavily on broad themes of leadership, national security, and future optimism. Rather than enumerating extensive domestic policy programs, Republican candidates emphasized abstract ideals such as freedom, responsibility, peace, and national strength. Their rhetoric focused on America's role in the world and its long-term destiny, often prioritizing general principles over concrete legislative agendas. Their platforms in this period similarly articulated commitments to defense, economic growth, conservation, and civil rights in broad, value-oriented language. (The American Presidency Project, n.d.)

These stylistic and thematic differences help explain the pronounced divergence observed in this epoch. Democrats' reliance on detailed reform language introduced a wide array of specific policy themes that were largely absent from Republican speeches, while Republicans'

preference for generalized rhetoric centered on unity and strength reduced thematic overlap between the parties. Across individual election cycles, the range and emphasis of topics also shifted. In 1948, Democratic rhetoric foregrounded civil rights, labor issues, and domestic economic concerns, while Republicans reflected contrasting framing of issues. By 1952, Eisenhower's nomination speech focused heavily on leadership and peace, diverging from policy-heavy Democratic discourse. In 1960, John F. Kennedy's acceptance speech blended a forward-looking vision with global change, once again producing thematic vectors distinct from those of Republican counterparts. (The American Presidency Project, n.d.)

Epoch 2: Trend Towards Convergence (≈ 1964 – 1996)

Across the three decades, several core issues consistently appeared in the rhetoric of both major parties, creating a degree of issue-level overlap despite opposing ideological framing. The Vietnam War, in particular, became a central political concern due to its profound domestic impact, elevating foreign policy to a dominant position in national discourse. Civil rights and social equality also emerged as defining issues, especially following the Civil Rights Act of 1964, which accelerated political realignment and forced parties to articulate contrasting positions on race, equality, and voting rights. Alongside these, broader economic and cultural concerns—including domestic welfare and social programs, national security and Cold War strategy, and questions of government reform and public trust—remained persistent elements of political debate. (Peters & Woolley, n.d.)

Despite addressing key areas of discourse, the parties adopted divergent rhetorical and policy stances. On the Vietnam War, Democratic platforms increasingly called for de-escalation, peace negotiations, and, in some cases, the end of conscription, with 1968 highlighting significant internal divisions within Democratic rhetoric itself. Republican speeches, by contrast, emphasized law and order, military strength, and a firm approach to foreign commitments, particularly under the leadership of Nixon and Reagan. In response to civil rights agitations, Democrats progressively integrated themes of social justice, anti-discrimination, and expanded political participation into their platforms. Republicans, meanwhile, shifted toward rhetoric centered on states' rights, individual liberty, and market freedom, particularly as conservative leadership consolidated and the Southern Strategy took shape. Across economic and cultural issues more broadly, Democrats framed their arguments around economic security and social protection, whereas Republicans emphasized market opportunity and personal responsibility. (Peters & Woolley, n.d.) (The American Presidency Project, n.d.)

Overall, this epoch exhibits greater thematic similarity than the previous epoch, as both parties shared the similarity of addressing specific national issues even while offering contrasting policies. This shared engagement helps explain the oscillatory pattern observed during these decades: although the specific issues evolved significantly over time, each election cycle featured both parties actively engaging with the same central topics in public discourse.

Epoch 3: Modern Divergence (2000 – 2024)

Epoch 3 reflects differences in both core national concerns and in the way parties framed their agendas. Together, these shifts contributed to reduced thematic overlap across 2 decades, despite starting at the point of maximum thematic convergence in 80 years.

Differences across key issues emerged most sharply in the areas of healthcare and cultural rights. Democratic platforms repeatedly placed healthcare access, public health, and social safety nets at the center of their agendas, emphasizing protections for the Affordable Care Act and addressing post-pandemic health challenges in 2020. Republicans, by contrast, moved away from comprehensive healthcare policy articulation, especially in 2020, where they did not release a full new platform, opting instead for a brief resolution endorsing the incumbent administration's priorities. This structural absence of detailed healthcare policy language reflects a broader shift away from issue-specific framing. Similarly, Democrats expanded the language of social rights over time, particularly with respect to LGBTQ+ equality and civil liberties from 2012 onward, embedding these issues as core components of their platforms. Republican rhetoric increasingly concentrated on cultural and identity-based issues such as border security, immigration enforcement, and law-and-order themes, which became central to party identity in the 2010s and 2020s, including an emphasis on strict immigration control and election integrity. (Peters & Woolley, n.d.)

In other domains, both parties addressed central issues with diametrically opposing views. On economic matters, Republican platforms from Bush through Trump emphasized tax cuts, deregulation, and free-market principles, with later rhetoric linking tariffs and immigration policy to economic nationalism under the "America First" framework. These platforms consistently prioritized lower taxes and a reduced role for government. Democratic platforms, from Gore through Biden and Harris, focused on economic fairness and middle-class security, advocating investment in infrastructure, healthcare, and economic relief, particularly after the 2008 financial crisis and the COVID-19 pandemic. Democratic economic language explicitly connected fiscal policy to equity and social support. In foreign policy, Democrats expressed commitments to international cooperation and alliance-building, including support for global institutions and re-engagement with agreements such as the Paris Climate Accord. Republicans, particularly under Trump, emphasized national sovereignty, border security, and defense priorities framed through an America First lens rather than multilateral engagement. (Peters & Woolley, n.d.) (The American Presidency Project, n.d.)

These stylistic differences were reinforced by contrasting communication strategies. Democratic platforms consistently engaged with a broader and more detailed set of policy issues, while Republican platforms became increasingly compact and selectively focused. This narrowing of issue articulation reduced thematic overlap between the parties, contributing to greater separation in topic space and reinforcing the pattern of modern rhetorical divergence."

Comment 6: Discussion must include limitations of the work and how future researchers can expand on it.

My Response: The 2 sections have been added accordingly:

“LIMITATIONS AND METHODOLOGICAL CONSTRAINTS

While this study uses a rigorous framework for measuring polarization, several computational limitations must be acknowledged. Firstly, incorporating TF-IDF and Jaccard similarity relies on a “bag-of-words” model, which completely ignores syntax and word order, and most importantly: context. For example, the phrase “I support healthcare reform” and “I oppose healthcare reform” would yield a high degree of lexical similarity while representing completely opposing ideological stances. Second, while VADER Sentiment Analysis is a standard for identifying texts with explicit tones, it will struggle to identify subtle hints of political sarcasm or “dog-whistle” rhetoric, which need a more in-depth understanding than simply rule-based models. Finally, such speeches represent a very specific, formalized genre of political oratory and does not capture the full extent to which polarization occurs which are found in more informal settings, such as social media.

RESEARCH SCOPE

Future studies should transition from static vector representations (TF-IDF) to context-based embeddings using Transformer-based architectures such as BERT (Bidirectional Encoder Representations from Transformers). Unlike the current bag-of-words approach, using the BERT architecture will capture the polysemy of the text, allowing researchers to identify how the meaning of the word changes based on the text around it. This would allow researchers to distinguish between shared vocabulary in opposition compared to the shared vocabulary in agreement. This would provide a metric that would capture the intent of the speaker rather than just the lexical frequency.

Furthermore, to identify the migration of political issues, researchers can deploy pipelines that utilize Dynamic Topic Modeling (DTM). While this study showed that parties are diverging, DTM would allow the tracking of specific topic trajectories over a period. This would reveal which topics are of higher interest to one party, such as the Republican “capture” of Border Security, and how the semantic weight would vary across the different election cycles.

Finally, while this paper used VADER for sentiment analysis, researchers in the future could incorporate Affective Computing to measure rhetorical aggression. Utilizing Emotion AI, researchers could classify the texts into more specific categories such as fear, anger or hope; this would allow researchers to determine if the decline in similarity is due to emotional framing instead of the policy substance.”

Comment 7: What are the most common phrases from electoral candidates over time per party? Include some visual representation (such as a word cloud) or a table.

Comment 8: Interparty embedding / vector similarity across time. Compare the representation of the first speech for one party to the representation of each belonging to the other party.

My Response: ‘Table 1 - Thematic Frequency Table (1944–2024)’ showcasing the most common phrases from electoral candidates over time has been added to the discussion section, along with its explanation and inter-party comparison, which is as follows:

Era	Key Thematic Phrases (Democrats)	Key Thematic Phrases (Republicans)	Analysis
Post-War (1944-1956)	<i>Common man, labor, social security, prosperity</i>	<i>Spiritual unity, free men, faith, liberty</i>	Low Jaccard (0.14-0.18): High ideological divergence.
Cold War (1960-1988)	<i>New frontier, great society, justice, world</i>	<i>Peace through strength, tax cuts, freedom</i>	Moderate Jaccard: National security acts as a unifying anchor.
Modern (1992-2008)	<i>Working families, healthcare, middle-class, jobs</i>	<i>Small business, opportunity, mortgage, security</i>	High Jaccard (0.30): High lexical overlap due to shared economic crises.
Current (2012-2024)	<i>Democracy, rights, future, diversity</i>	<i>Make America Great, borders, strength, fight</i>	Declining Jaccard: Re-emergence of partisan "siloes" language.

“The frequency analysis reveals a significant shift in how both parties construct their speeches. While the lemma “America” stayed the most frequent term across the corpus at 1.34%, its usage density varied depending on the political climate and degree of partisan polarization.

During the mid-20th century, the analysis shows a high degree of thematic overlap, characterized by Cosine Similarity and Jaccard scores. From a computational perspective, this shows that the priority vectors of both the parties were closely aligned. This thematic overlap was likely influenced by existential threats, such as World War II and the onsets of the Cold War, and this resulted in a unified national drive which led to candidates using common vocabulary to describe immediate issues.

Starting in the late 1960s, the results show the beginnings of semantic decoupling. While the Cosine Similarity remained fluctuating, the Jaccard score slowly began minimizing. The divergence highlights a shift in strategy, while still conveying similar vague themes, began utilizing specific, party-based lexicons. This period reflects the transition from a broadcast model of political communications to a more fragmented approach.

In the modern era, both metrics reached lows, highlighting that the candidates were moving to a state of orthogonality. The analysis reveals that the parties are no longer debating on a shared set of issues; rather, they are operating in their own distinct latent spaces. This trend is evident as it stays constant across multiple election cycles, regardless of the individual candidates' campaigning styles, indicating a gradual shift towards permanent rhetorical fragmentation."

Comment 9: Also, your average reader is not going to understand the distinction between a cosine similarity of 0.6 to that of a 0.2. Can you provide a blurb from each speech (2 from 0.6 similarity score and 2 from 0.2 similarity score only) so the reader can "see" the difference?

My Response: A sub-section 'Qualitative Validation: Interpretability Blurbs of Lexical Similarity' has been added under the discussion section, providing the required snippets:

"To illustrate the 0.3016 Jaccard score identified in the 2008 election cycle, the following excerpts from Barack Obama and John McCain show a significant convergence of lexical priorities. Despite their ideological differences, both candidates have utilized a nearly identical "lexicon of crisis" to show the immediate economic anxieties of the American crowd.

Snippet A: Barack Obama (Democrat, 2008)

"We meet at one of those defining moments – a moment when our nation is at war, our economy is in turmoil... Tonight, more Americans are out of work and more are working harder for less. More of you have lost your homes and even more are watching your home values plummet. More of you have cars you can't afford to drive, credit card bills you can't afford to pay".

Snippet B: John McCain (Republican, 2008)

"These are tough times for many of you. You're worried about keeping your job or finding a new one, and you're struggling to put food on the table and stay in your home... We're not going to allow that while you struggle to buy groceries, fill your gas tank, and make your mortgage payment"

The following table shows the high-value lemmas that appeared with high-frequencies in both the texts, contributing to the study-wide similarity score:

<i>Keyword</i>	<i>Obama</i>	<i>McCain</i>
<i>Mortgage</i>	<i>"...job that pays the mortgage"</i>	<i>"...make your mortgage payment"</i>
<i>Home</i>	<i>"...lost your homes"</i>	<i>"...stay in your home"</i>
<i>Work/Job</i>	<i>"...out of work... 23 million new jobs"</i>	<i>"...keeping your job... create new jobs"</i>
<i>Economy</i>	<i>"...economy is in turmoil"</i>	<i>"...rescue our economy"</i>

Table 2 – High-value lemmas from the 2008 election cycle

The record low similarity in 1948, of 0.1485, is the result of two candidates addressing the nation from two completely different perspectives: Dewey from a highly philosophical perspective while Truman having a policy-heavy stance.

Snippet C: Thomas Dewey (Republican, 1948)

"Our people... have a yearning to move to higher ground, to find a common purpose in the finer things which unite us. The unity we seek is more than material. It is more than a matter of things and measures. It is most of all spiritual... We have not yet found the spiritual means to put together the world's broken pieces."

Snippet D: Harry Truman (Democrat, 1948)

"The Republican Party... favors the privileged few and not the common everyday man... They proved it by the things they did to the people. Prices have gone all the way off the chart... I am calling this Congress back into session to pass laws to halt rising prices, to meet the housing crisis... and an increase in the minimum wage."

While the 2008 high-similarity apex was driven by a common vocabulary of crisis, the 1948 floor is defined by thematic isolation. The candidates did not agree on what the nation's problems were; therefore, this caused them to have a very small intersection of shared word sets, resulting in the low Jaccard score."

Comment 10: Include graphics directly in the space mentioned and not in the appendix. Appendix should be only for overflow, supplemental items.

My Response: This has been done accordingly. The 'Appendix' section has been removed and all graphs have been moved to their respective sections.

Comment 11: The Jaccard Index is mentioned and never used in the rest of the paper.

My Response: The sub-section 'Lexical Overlap Measurement: Jaccard Similarity' has been added to the 'Results' section to show the use of Jaccard Similarity, which is as follows:

"Lexical Overlap Measurement: Jaccard Similarity

The analysis of presidential nomination speeches shows a significant fluctuation in raw lexical overlap. By calculating the intersection of unique words, this metric reveals to which extent candidates work within a shared linguistic framework versus isolated one-sided lexicons.

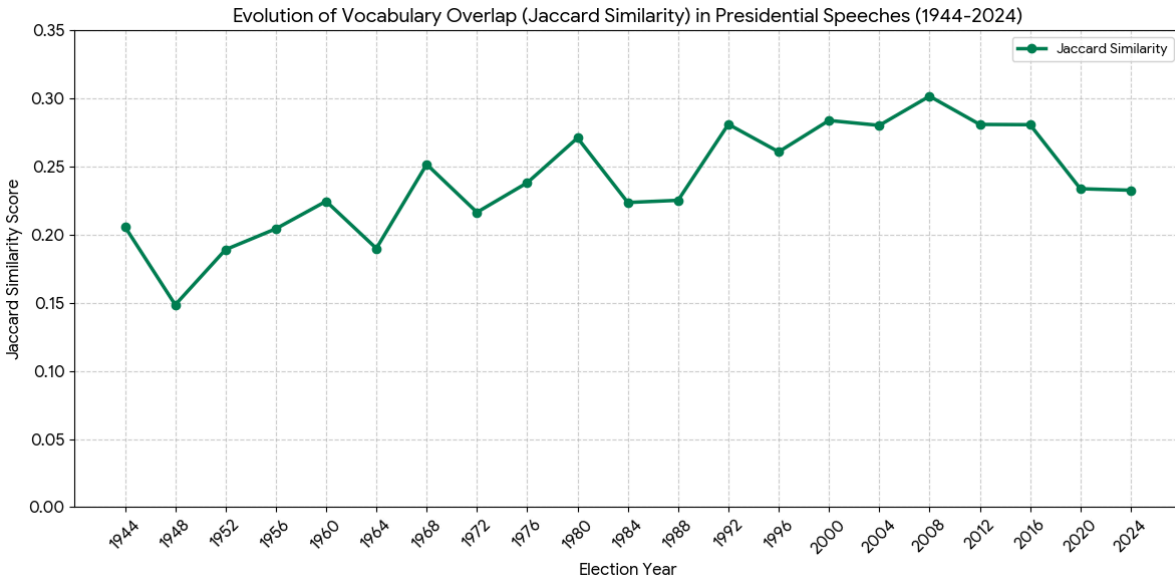


Figure 4 – Evolution of Vocabulary Overlap (Jaccard Similarity) in Presidential Speeches (1944-2024)

The 1948 Lexical Floor: The analysis reveals 1948 (Truman vs. Dewey) as the point of maximum lexical divergence across the 80-year corpus. During this cycle, the Jaccard Similarity reached a historic floor of 0.1485. This reveals that the two candidates had a collective unique vocabulary of less than 15%. The rhetoric isolation of this score is highlighted by Truman’s aggressive focus on “Eight-Nothin’ Congress” and “labor rights”, while Dewey’s strategy focused on consolidating blue states (at the time) in the Northeast through optimising framing of “unity” and “national progress”.

The 2008 Convergence Apex: A distinct reversal in the long-time trend occurred during the 2008 election cycle, where the Jaccard Similarity reached its peak of 0.3016. This convergence was motivated by a high degree of lexical overlap regarding the Great Recession and foreign policy. Particularly, McCain and Obama campaigned extensively in swing states (Florida, Ohio and Pennsylvania) using a more shared vocabulary of “economic crisis” and “reform” to appeal to the median voter, which resulted in the highest recorded lexical commonality in the study.

The data also suggests that the timing of the convention and the location of the speech also acted as variables in terms of the rhetoric distance. Speeches that were delivered later in the election cycle, usually in August or September, tend to show a minimal decrease in the commonality of vocabulary compared to the early-summer addresses. As the election date approaches, the language shifts from a general platform-building to more aggressive contrast-marking. Historically, when conventions took place in critical swing states, there was a visible “thematic pull” that stabilized Jaccard scores as both parties attempted to convey the regional economic concerns. However, the post-2008 data suggests that this effect is gradually fading; candidates maintain their partisan lexicon regardless of the location of the convention, prioritizing national audiences over regional interests.

The fluctuations in Jaccard scores have been persistent, which confirms that political polarization is fundamentally linguistic. Even during years of high positive sentiment – such as

1956 – the shared unique vocabulary has remained low, at 0.2043, suggesting the vocabulary of priority has stayed divided.”

Comment 1: For example, in the Results section, you describe 3 different epochs across time. Perhaps you could return to those epochs in the Discussion section to provide more historical and political context.

Also, rather than just discussing the lowest and highest cosine similarity scores, provide an analysis of each of the epochs, or each of the presidencies from 1948 until the present. Or, a decade-by-decade analysis of the last 50 years of presidencies might be a more reasonable workload. You might decide instead of a broad analysis of speeches beginning in 1948, you might choose a shorter time frame to engage in deeper analysis.

My Response: The discussion section has been amended accordingly. Previous explanations of the highest and lowest point on the Cosine Similarity Graph have been removed, and epoch wise explanation of the graph has been added. Although there still exist slight historical references to contextualize the nature of the 3 epochs, the historical detail of each analysis has been limited to a large extent. The description of the epochs have been revised to come from a computational perspective, instead of a historical one. The aim of the explanation lies in the use of specific vocabulary and issues referred/not referred to in both parties' nomination speeches as reasons for justifying the variation in cosine similarity scores, across the 80 year span. The revised section is as follows:

“Epoch 1: Early Volatility (1944 – ≈ 1960):

Immediately after World War II, the United States was tasked with transitioning from a wartime to a peacetime economy. This rapid shift—from military production to civilian industry, and from full employment to managing inflation and labor unrest—produced an unstable domestic environment in which political rhetoric lacked consistent framing. At the same time, the emergence of early Cold War tensions with the Soviet Union introduced new foreign-policy priorities that were not uniformly emphasized across party lines. These economic uncertainties, coupled with uneven postwar prosperity and emerging social change, contributed to a fragmented political landscape in which shared thematic ground between parties was limited.

Democratic nomination speeches during this period placed strong emphasis on domestic critique and reform. Candidates often directly confronted Republican platforms and advocated expansive federal intervention to address issues such as unemployment, housing shortages, consumer protection, and economic instability. The global context was often framed in explicitly moral terms, linking domestic justice to international peace and economic cooperation. These speeches were typically dense with policy detail, outlining specific proposals across a wide range of areas including employment, wages, agriculture, education, civil rights, environmental conservation, and federal aid programs. (Peters & Woolley, n.d.)

Republican nomination speeches, by contrast, relied more heavily on broad themes of leadership, national security, and future optimism. Rather than enumerating extensive domestic policy programs, Republican candidates emphasized abstract ideals such as freedom, responsibility, peace, and national strength. Their rhetoric focused on America's role in the world and its long-term destiny, often prioritizing general principles over concrete legislative agendas.

Their platforms in this period similarly articulated commitments to defense, economic growth, conservation, and civil rights in broad, value-oriented language. (The American Presidency Project, n.d.)

These stylistic and thematic differences help explain the pronounced divergence observed in this epoch. Democrats' reliance on detailed reform language introduced a wide array of specific policy themes that were largely absent from Republican speeches, while Republicans' preference for generalized rhetoric centered on unity and strength reduced thematic overlap between the parties. Across individual election cycles, the range and emphasis of topics also shifted. In 1948, Democratic rhetoric foregrounded civil rights, labor issues, and domestic economic concerns, while Republicans reflected contrasting framing of issues. By 1952, Eisenhower's nomination speech focused heavily on leadership and peace, diverging from policy-heavy Democratic discourse. In 1960, John F. Kennedy's acceptance speech blended a forward-looking vision with global change, once again producing thematic vectors distinct from those of Republican counterparts. (The American Presidency Project, n.d.)

Epoch 2: Trend Towards Convergence (≈ 1964 – 1996)

Across the three decades, several core issues consistently appeared in the rhetoric of both major parties, creating a degree of issue-level overlap despite opposing ideological framing. The Vietnam War, in particular, became a central political concern due to its profound domestic impact, elevating foreign policy to a dominant position in national discourse. Civil rights and social equality also emerged as defining issues, especially following the Civil Rights Act of 1964, which accelerated political realignment and forced parties to articulate contrasting positions on race, equality, and voting rights. Alongside these, broader economic and cultural concerns—including domestic welfare and social programs, national security and Cold War strategy, and questions of government reform and public trust—remained persistent elements of political debate. (Peters & Woolley, n.d.)

Despite addressing key areas of discourse, the parties adopted divergent rhetorical and policy stances. On the Vietnam War, Democratic platforms increasingly called for de-escalation, peace negotiations, and, in some cases, the end of conscription, with 1968 highlighting significant internal divisions within Democratic rhetoric itself. Republican speeches, by contrast, emphasized law and order, military strength, and a firm approach to foreign commitments, particularly under the leadership of Nixon and Reagan. In response to civil rights agitations, Democrats progressively integrated themes of social justice, anti-discrimination, and expanded political participation into their platforms. Republicans, meanwhile, shifted toward rhetoric centered on states' rights, individual liberty, and market freedom, particularly as conservative leadership consolidated and the Southern Strategy took shape. Across economic and cultural issues more broadly, Democrats framed their arguments around economic security and social protection, whereas Republicans emphasized market opportunity and personal responsibility. (Peters & Woolley, n.d.) (The American Presidency Project, n.d.)

Overall, this epoch exhibits greater thematic similarity than the previous epoch, as both parties shared the similarity of addressing specific national issues even while offering contrasting

policies. This shared engagement helps explain the oscillatory pattern observed during these decades: although the specific issues evolved significantly over time, each election cycle featured both parties actively engaging with the same central topics in public discourse.

Epoch 3: Modern Divergence (2000 – 2024)

Epoch 3 reflects differences in both core national concerns and in the way parties framed their agendas. Together, these shifts contributed to reduced thematic overlap across 2 decades, despite starting at the point of maximum thematic convergence in 80 years.

Differences across key issues emerged most sharply in the areas of healthcare and cultural rights. Democratic platforms repeatedly placed healthcare access, public health, and social safety nets at the center of their agendas, emphasizing protections for the Affordable Care Act and addressing post-pandemic health challenges in 2020. Republicans, by contrast, moved away from comprehensive healthcare policy articulation, especially in 2020, where they did not release a full new platform, opting instead for a brief resolution endorsing the incumbent administration's priorities. This structural absence of detailed healthcare policy language reflects a broader shift away from issue-specific framing. Similarly, Democrats expanded the language of social rights over time, particularly with respect to LGBTQ+ equality and civil liberties from 2012 onward, embedding these issues as core components of their platforms. Republican rhetoric increasingly concentrated on cultural and identity-based issues such as border security, immigration enforcement, and law-and-order themes, which became central to party identity in the 2010s and 2020s, including an emphasis on strict immigration control and election integrity. (Peters & Woolley, n.d.)

In other domains, both parties addressed central issues with diametrically opposing views. On economic matters, Republican platforms from Bush through Trump emphasized tax cuts, deregulation, and free-market principles, with later rhetoric linking tariffs and immigration policy to economic nationalism under the "America First" framework. These platforms consistently prioritized lower taxes and a reduced role for government. Democratic platforms, from Gore through Biden and Harris, focused on economic fairness and middle-class security, advocating investment in infrastructure, healthcare, and economic relief, particularly after the 2008 financial crisis and the COVID-19 pandemic. Democratic economic language explicitly connected fiscal policy to equity and social support. In foreign policy, Democrats expressed commitments to international cooperation and alliance-building, including support for global institutions and re-engagement with agreements such as the Paris Climate Accord. Republicans, particularly under Trump, emphasized national sovereignty, border security, and defense priorities framed through an America First lens rather than multilateral engagement. (Peters & Woolley, n.d.) (The American Presidency Project, n.d.)

These stylistic differences were reinforced by contrasting communication strategies. Democratic platforms consistently engaged with a broader and more detailed set of policy issues, while Republican platforms became increasingly compact and selectively focused. This narrowing of issue articulation reduced thematic overlap between the parties, contributing to greater separation in topic space and reinforcing the pattern of modern rhetorical divergence."

Comment 2: Also, some thoughts to consider: for each speech, where did they take place: in red states, blue states, or swing states? Does the location of the speech have an impact on the language used? The locations for each speech are listed in the database that you used, alongside the dates, which span from May through September. Are there differences in language based on the month? For example, more positive patriotic speeches in July around Independence Day? Another thought to consider: is there more to “feelings” than just positive vs negative?

My Response: A sub-section ‘Lexical Overlap Measurement: Jaccard Similarity’ has been added to the ‘Results’ section describing the results of Jaccard Similarity in detail. These results have been described with respect to blue/red/swing states, location, time period, date/month, and significance of surrounding events like election day/independence day. Additionally, the impact of “feelings” have been described beyond positive/negative sentiment, and on the basis of issues/vocabulary mentioned in every speech, which increased voter engagement towards their associated party. The section is as follows:

“Lexical Overlap Measurement: Jaccard Similarity

The analysis of presidential nomination speeches shows a significant fluctuation in raw lexical overlap. By calculating the intersection of unique words, this metric reveals to which extent candidates work within a shared linguistic framework versus isolated one-sided lexicons.

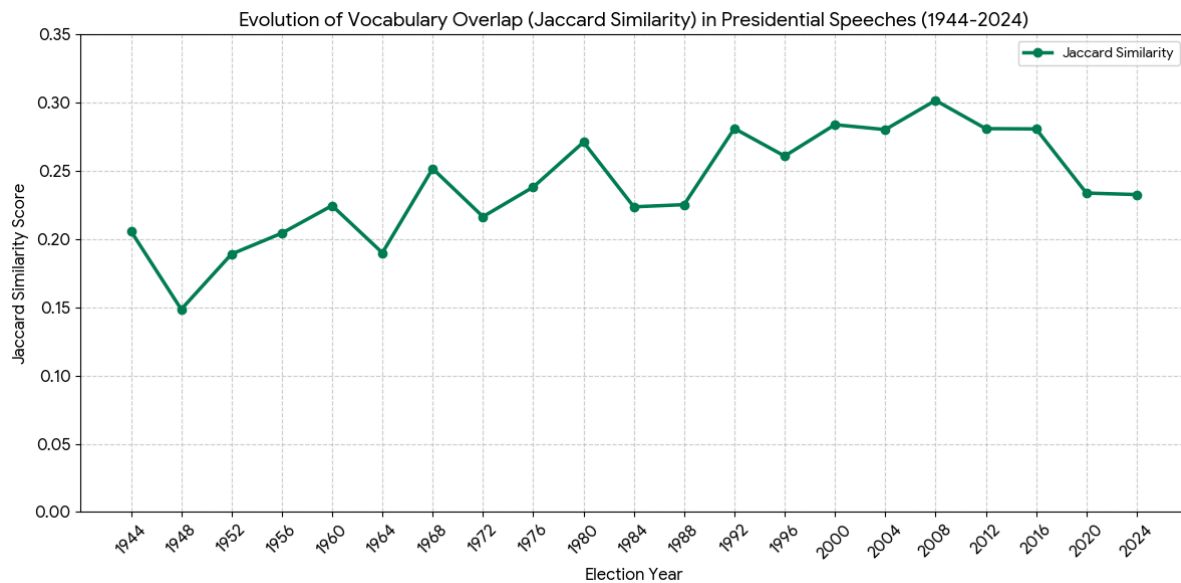


Figure 4 – Evolution of Vocabulary Overlap (Jaccard Similarity) in Presidential Speeches (1944-2024)

The 1948 Lexical Floor: The analysis reveals 1948 (Truman vs. Dewey) as the point of maximum lexical divergence across the 80-year corpus. During this cycle, the Jaccard Similarity reached a historic floor of 0.1485. This reveals that the two candidates had a collective unique vocabulary of less than 15%. The rhetoric isolation of this score is highlighted by Truman’s aggressive focus on “Eight-Nothin’ Congress” and “labor rights”, while Dewey’s strategy focused

on consolidating blue states (at the time) in the Northeast through optimising framing of “unity” and “national progress”.

The 2008 Convergence Apex: *A distinct reversal in the long-time trend occurred during the 2008 election cycle, where the Jaccard Similarity reached its peak of 0.3016. This convergence was motivated by a high degree of lexical overlap regarding the Great Recession and foreign policy. Particularly, McCain and Obama campaigned extensively in swing states (Florida, Ohio and Pennsylvania) using a more shared vocabulary of “economic crisis” and “reform” to appeal to the median voter, which resulted in the highest recorded lexical commonality in the study.*

The data also suggests that the timing of the convention and the location of the speech also acted as variables in terms of the rhetoric distance. Speeches that were delivered later in the election cycle, usually in August or September, tend to show a minimal decrease in the commonality of vocabulary compared to the early-summer addresses. As the election date approaches, the language shifts from a general platform-building to more aggressive contrast-marking. Historically, when conventions took place in critical swing states, there was a visible “thematic pull” that stabilized Jaccard scores as both parties attempted to convey the regional economic concerns. However, the post-2008 data suggests that this effect is gradually fading; candidates maintain their partisan lexicon regardless of the location of the convention, prioritizing national audiences over regional interests.

The fluctuations in Jaccard scores have been persistent, which confirms that political polarization is fundamentally linguistic. Even during years of high positive sentiment – such as 1956 – the shared unique vocabulary has remained low, at 0.2043, suggesting the vocabulary of priority has stayed divided.”

Comment 3: The intro does a good job using examples of less reliance on words like “data” and more reliance on “feel.” But is it possible to feel something beyond pos/neg? Are there neutral feelings, like the French saying “comme ci comme ça”?”

My Response: A paragraph has been added in the introduction section incorporating the suggested change:

“This shift in “feeling” is not just a move toward extreme positivity or negativity. As the French expression “comme ci, comme ça” suggests, most of human sentiment stays in the middle ground – a state of calculative neutrality. Computationally, this neutrality acts as a proxy for policy-heavy discourse. In the 1940s and 50s, a high ratio of neutral scores in VADER reflected a focus on administrative governance that surpassed partisan emotion.”

Comment 4: In the intro, write a bit about the anatomy of a speech, and how any speech (not just political ones) will follow certain formats and emotional cues. Perhaps consider different frame for your analysis, such as comparing nomination speeches to inauguration speeches, or winning candidates only, etc.

My Response: 2 paragraphs explaining the structural and rhetorical anatomy of a speech have been added to the ‘introduction’ section. Additionally, its application has been compared

between convention speeches and inauguration speeches to briefly justify how the same structural and rhetorical frameworks fulfill specific purposes of each type of speech in politics. The added paragraphs are as follows:

“An essential part of understanding political rhetoric is recognising that all speeches — whether political, academic, or ceremonial—follow a structured format designed to achieve specific communicative goals. Across contexts, public speaking pedagogy emphasises that a speech typically consists of three core components: an introduction that captures attention and establishes purpose, a body where the substantive arguments and evidence are presented in logical progression, and a conclusion that reinforces the key message and leaves a lasting effect on the audience. This structural organisation supports clarity, coherence, and persuasion, and is central to effective communication in diverse domains (Language Through Literature, 2024).

Beyond structural elements, speeches also deploy rhetorical appeals that blend credibility, logic, and emotion to engage audiences. Classical rhetorical theory, rooted in Aristotle, highlights ethos (the speaker’s credibility), logos (logical argumentation), and pathos (emotional resonance) as foundational to persuasive discourse; these appeals continue to inform modern speech analysis and illustrate how speakers adapt style and content to influence interpretation and judgment (Robinson & Jerskey, 2021). This framework not only applies to individual speeches but also enables meaningful comparison across genres of political oratory—for instance, contrasting nomination acceptance speeches with inauguration addresses reveals differences in purpose, audience engagement, and emotional framing, with acceptance speeches often designed to mobilise partisan support and inaugurals shaped to articulate vision and national unity (Koller & Semino, 2019).”

Decision: Accept with major revisions.

Review: This manuscript investigates longitudinal semantic divergence in U.S. presidential nomination speeches (1944-2024) using TF-IDF cosine similarity, Jaccard overlap, and VADER sentiment analysis. The topic is very timely, and the longitudinal framing is an intriguing and useful approach. Excellent job also responding to the two peer reviewers.

However, I would want to see certain major revisions before the manuscript can be published. The core issue is not the ideas but the precision of execution, interpretation of results, and how claims were phrased. I do not think you need to redo the current analysis, but some things should be reframed, and certain details and tests should ideally be added.

Major concerns:

1. Overstatement of novelty and claims.

The manuscript repeatedly claims methodological novelty (e.g., “the unique contribution of this research lies in the comparative quantification methodology”), but after a quick search, pairwise comparison of political speeches using TF-IDF and cosine similarity is known in computational political research. Some examples:

<https://web.stanford.edu/~gentzkow/research/text-as-data.pdf>,

https://ideas.repec.org/a/cup/polals/v28y2020i1p112-133_6.html). There is nothing wrong with what you did, and the *application* is novel, but the methodology is not.

I would reframe the contribution as a focused application over a specific corpus (i.e., nomination speeches over a specific time period) with an emphasis on pairing your analysis with historical and political context. Avoid implying methodological novelty unless a clear algorithmic extension is introduced.

2. More depth of statistical analysis.

The paper paints a nice qualitative narrative, and it reports means, slopes across historical epochs, and claims of statistically significant divergence. However, the paper presents no regression models, standard errors, confidence intervals, or hypothesis tests. I would want more robust tests instead of, say, merely observing that the divergence occurred “six times greater” in one period vs. another.

Some ideas:

- Linear regressions for each epoch, reported with slope coefficients, standard errors, p-values, and model fit statistics.
- Permutation tests or other robustness checks to demonstrate that observed trends differ from what would arise under random variation.

Without formal statistics, the results remain descriptive, and it is difficult to substantiate claims of statistical significance, even if the underlying methods are of note!

3. Corpus description and preprocessing analysis.

I may have missed this, but it remains unclear how many speeches were included per election cycle, whether third-party candidates were excluded, how TF-IDF's inverse document frequency component was computed (per election cycle or across the full corpus), whether document vectors were normalized, and how differences in speech length were handled. From my understanding, these can affect similarity scores and should be explicitly stated to ensure reproducibility.

The authors should provide a subsection (perhaps in an appendix?) describing the corpus size, token counts, preprocessing libraries and parameters, stopword handling, vectorization technique, and any smoothing applied to IDF. In any sort of computational science, computational details like these are important.

Relatedly, I have requested a code and data availability section.

4. Be wary of your interpretation, particularly of causality and correlation between certain factors.

I would argue that statements like “political polarization is fundamentally linguistic” or that parties are “operating in their own distinct latent spaces” are not supported by bag-of-words similarity measures alone. TF-IDF cosine similarity captures lexical proximity, but it is a bit of a stretch to extend this to assumptions on ideological structure, voter cognition, or causal political dynamics. I may even go as far as to say that different lexical structures may sometimes encode similar thoughts and feelings, just filtered through a different sociocultural context.

The manuscript should adopt a more careful, measured, and less certain language. You should more carefully distinguish lexical divergence from ideological polarization and electoral behavior. As of now, some of the conclusions risk appear rhetorically inflated.

Additionally, claims regarding voter engagement or turnout effects associated with positive or negative rhetoric are speculative and not empirically tested. You should heavily temper or remove these.

Minor concerns:

1. A compound VADER score of -0.997 for the 2016 Republican nomination speech is extremely close to the theoretical lower bound of -1, and is a bit suspicious. Was sentiment computed at the document level or averaged across sentences? Was preprocessing applied before sentiment scoring? The authors should perhaps show a full distribution of sentiment scores across years and remove or temper claims about behavioral consequences unless additional data are introduced.

2. References include several non-scholarly or questionable sources, such as financial education websites and pedagogical web materials. All citations should be carefully verified, and non-peer-reviewed sources removed or replaced.
3. After addressing the major revisions, the manuscript would still benefit from structural tightening. The introduction contains an extended discussion of Aristotelian rhetoric and speech structure that does not add too much to the computational analysis. Some lengthy historical narratives could also be condensed or partially moved to an appendix. Overall, greater concision would improve focus and readability.

Overall, the manuscript has a clear study design and motivation, complementary and modern similarity metrics, thoughtful and careful historical framing, and a transparent limitations section. However, I still think significant improvements could be made to the statistical rigor, methodological transparency, interpretation, and source quality. In its current form, it is promising and of good quality, but not yet publication-ready.