THE UNIVERSITY OF CHICAGO

The Language of Bias: Examining the Influence of Political Bias and Electoral Context on Emotional, Moral, Intergroup Language and Sentiment in News Media

By

Yutong Jiang

June 2024

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts in Computational Social Science

Faculty Advisor: Joshua Conrad Jackson

Preceptor: Zhao Wang

Abstract

In an era of increasing political polarization, the language used by news media has become a crucial area of investigation. This study investigates the relationship between political bias and language usage in news articles, focusing on the use of emotional, moral, intergroup language, and sentiment. By analyzing a comprehensive dataset of 2,510,592 news articles from 234 publishers spanning the years 2010 to 2022, I seek to shed light on the linguistic strategies used by media outlets to influence public perception and opinion.

My approach combines a novel Transformer-Dictionary Hybrid method for identifying emotional, moral, and intergroup content with sentiment analysis using the VADER tool. The study addresses two key research questions: 1) How does the usage of emotional, moral, and intergroup language differ among news publishers with different political biases? 2) Do news outlets employ this language differently during election years compared to non-election years?

ANOVA and mixed linear model analyses reveal significant differences in the usage of emotional, moral, intergroup language, and sentiment across publishers with different political biases. Liberal publishers tend to use more emotional and moral language compared to left-leaning and neutral publishers. Intergroup language usage is significantly lower among neutral publishers compared to liberal ones. Sentiment analysis shows that conservative publishers express a more negative sentiment compared to liberal publishers.

A two-way ANOVA examining the effect of election years shows a significant interaction between language dimensions and election cycle. Mixed linear model analyses reveal a decrease in intergroup language usage and a slight decrease in positive sentiment during election years. The findings contribute to our understanding of how political bias and electoral context influence the linguistic choices in news media. The results suggest that partisan outlets strategically employ language to engage and potentially polarize their audience, while adapting to the heightened political climate of election years. These insights shed light on the complex interplay between language, media bias, and political polarization, underlining the importance of media literacy in navigating the contemporary media landscape.

Keywords: media bias, affective polarization, language usage, sentiment analysis, political communication, computational social science

1 Introduction

The increasing polarization of the American political landscape has drawn significant attention to the role of media in shaping public opinion and discourse. As partisanship becomes increasingly divided, the language employed by news outlets across the political spectrum has come under scrutiny. This study aims to investigate the relationship between political bias and language usage in news articles, focusing on the use of emotional, moral, and interpersonal language, as well as the overall sentiment expressed in the articles. By analyzing a comprehensive dataset of news articles from various publishers spanning the years 2010 to 2022, I seek to shed light on the linguistic strategies used by media outlets to influence public perception and opinion, and potentially the landscape of political polarization in the US.

The contemporary media environment has been characterized by the proliferation of partisan news outlets, which have been blamed for exacerbating affective polarization (Lelkes et al., 2017) As Iyengar and colleagues note, several features of the current landscape, such as sorted partisanship and the high-choice media environment, have contributed to the proclivity of partisans to divide the world into a liked in-group (one's own party) and a disliked out-group (the opposing party) (2019).

For example, partisan outlets often depict the opposing party in harsh terms and focus disproportionately on out-party scandals (Puglisi and Snyder, 2011). The lack of balanced content may inculcate hostility toward the out-group and persuade viewers to adopt extreme ideological positions (Iyengar et al., 2019). However, the exact mechanism and whether this relationship is causal are both unconfirmed (Iyengar et al., 2019). Despite content coverage choices, such as focus on out-party scandals, previous research has identified the usage of language, such as labeling and word choice, that journalists may use to bias news (Hamborg et al., 2019). Recent research has examined more aspects of language, such as the use of moral language and overall sentiment expressed in news media and their potential links to affective polarization. For example, Wang and Inbar (2021) found that moral language is used more frequently by media outlets when the opposing party is in power. This finding suggests that the political context may influence the linguistic strategies employed by media outlets to engage and persuade their audiences. Rozado et al. (2022) analyzed the sentiment trends in news articles from various outlets and found notable differences between left and right-leaning media. This suggests that the language used by media outlets across the political spectrum may contribute to the growing divide in public opinion. Furthermore, Ludwig et al. (2023) explored the connection between content algorithms, sentiment analysis, and affective polarization, highlighting the potential role of algorithmic amplification in exacerbating political divisions through the promotion of emotionally charged content.

To expand upon these previous analyses, I will examine language usage from three dimensions: emotional, moral, and interpersonal, as well as overall sentiment trends. This approach is informed by the work of Brady et al. (2023), who argue that during algorithm-mediated social learning, content algorithms on social media platforms exploit human social-learning biases and amplify prestigious, ingroup, moral, and emotional ('PRIME') information as a side-effect of goals to maximize engagement on the platforms. By applying this perspective to the analysis of news articles, I aim to uncover the potential role of emotional, moral, and interpersonal language in shaping public opinion and driving affective polarization, particularly in the context of news consumption.

Specifically, this study seeks to address the following research questions:

- RQ 1: How does the usage of emotional, moral, and interpersonal language differ among news publishers with different political biases?
- RQ 2: Do news outlets employ PRIME language differently during election years compared to non-election years?

2 Literature Review

2.1 Affective Polarization: Background and Origin

Affective polarization, the increasing emotional divide between individuals who identify with opposing political parties, has become a prominent feature of the American political landscape in recent years. This phenomenon is characterized by a strong preference for one's own party (the in-group) and a growing hostility towards the opposing party (the out-group) (Iyengar et al., 2019). The origins and causes of affective polarization are multifaceted, with several key factors contributing to its rise.

One of the primary drivers of affective polarization is the increasing prevalence of sorted partisanship. Over the past 50 years, the percentage of partisans who identify with the party that most closely reflects their ideology has steadily increased (Iyengar et al., 2019). This sorting process has made it easier for partisans to make generalized inferences about the opposing side, even if those inferences are inaccurate. As a result, partisans are more likely to view the opposing party as a homogeneous and threatening out-group, even though ideologies remain unchanged. This is considered as the social identity cause of rising affective polarization. Research also supports that apart from reinforcing social identities, there is also an increasing ideological divide (2019).

Another factor contributing to affective polarization is the high-choice media environment and the rise of partisan news outlets. These outlets are frequently blamed for the current polarized climate, as they tend to activate partisan identities and consequent feelings toward the political parties (Lelkes et al., 2017). Social identity theory suggests that individuals thrives to identify with exemplary in-group members. In the context of partisan media, this tendency can manifest in the form of news outlets portraying the opposing party in an extremely negative light. Moreover, these outlets tend to disproportionately focus on scandals, whether substantiated or not, involving members of the out-party (Puglisi and Snyder, 2011). Consequently, this type of biased coverage can foster and reinforce feelings of animosity and resentment towards the out-group, ultimately contributing to the exacerbation of affective polarization. As individuals are repeatedly exposed to such negative portrayals of the opposing party, they may internalize these sentiments, leading to a deepening emotional divide between adherents of different political ideologies.

Besides negative portrayals, the lack of balanced content in partisan outlets may persuade viewers to adopt extreme ideological positions, which, in turn, can increase affective polarization. This suggests that exposure to partisan media may not only reinforce existing partisan attitudes but also drive individuals towards more extreme positions, further widening the gap between opposing partisans (Iyengar et al., 2019).

While the relationship between partisan media and affective polarization is well-established, the precise mechanisms through which this occurs remain unclear. It is argued that such a relationship is mitigated by individual's prior partisanship and polarizing attitudes. Those already highly polarized people will seek more partisan news (2019). This tendency can be linked to "filter bubble", given the rise of digital media, particularly algorithmically curated and individually customizable environments like social media platforms, news outlets, and search engines (Ludwig et al., 2023). The "filter bubble" hypothesis posits that personalized, algorithm-driven news recommendation systems (NRS) tend to favor news items that align with users' existing political attitudes, thus creating a homogeneous environment that increasingly drives polarization (2023).

Furthermore, the algorithmic curation of news content can create a feedback loop that intensifies the effects of partisan media on affective polarization. As users engage with content that confirms their existing beliefs and attitudes, algorithms may interpret this as a preference for similar content, leading to the delivery of even more ideologically aligned news and information. This process can result in the formation of echo chambers, where individuals are primarily exposed to ideas and arguments that support their existing views, while limiting their exposure to diverse perspectives (Iyengar et al., 2019). Consequently, this reinforcement of political attitudes can contribute to the exacerbation of affective polarization, as individuals become increasingly entrenched in their partisan identities and hostile towards the out-group.

2.2 The Interplay between Language, Algorithms, and Polarization

The combination of partisan media bias and algorithmic filter bubbles creates a potent environment for the growth of affective polarization. As individuals are exposed to a steady stream of negative portrayals of the opposing party, curated by algorithms designed to maximize engagement and reinforce existing beliefs, they may become increasingly emotionally invested in their partisan identities. This, in turn, can lead to a heightened sense of animosity and distrust towards the out-group, ultimately contributing to the widening emotional divide between political parties.

Content algorithms have been the subject of extensive research in recent years, focusing on various aspects of their design, implementation, and impact on user experiences and societal outcomes. Specifically, research has delved into how the linguistic features of a post or article influence its algorithmic amplification. Many research has shown that moralized and emotional information is highly likely to spread through online social network platforms (Brady et al., 2023). The use of specific language elements, such as emotional, moral, or inflammatory words and phrases, has been shown to be associated with increased engagement and dissemination by algorithms designed to maximize user attention and interaction (Brady et al., 2017). According to the study, the presence of moral-emotional language in social media communications significantly increases the diffusion of moral and political ideas within ideological group boundaries, a process the authors call ”moral contagion.” This finding suggests that the use of moralized and emotionally charged language can lead to the formation of echo chambers and the amplification of political polarization on social media platforms, as individuals are exposed to an increasingly homogeneous set of ideas that align with their preexisting beliefs (2017). More recently, Brady et al. (2020) explored the mechanism of increasing diffusion of emotional and moralized content on social media. They suggest that moral and emotional content captures people’s attention more than other content during political discourse on social media and that political leaders can leverage the attentional capture effects of moral and emotional language to increase the reach and impact of their messages.

Building upon these findings, Ludwig et al. (2023) investigate the link between content algorithms, sentiment analysis, and affective polarization. They suggest that sentiment analysis techniques can be used to identify and quantify the emotional tone of social media posts and news articles, providing insights into how the linguistic features of content influence its spread and impact on user attitudes. They find that more exposure to NRS enriched with negative sentiment lead to more affective polarization, whereas participants using an NRS incorporating balanced sentiment ideologically depolarized over time.

Besides moral, emotional language, and sentiment, a recent study has linked more language dimensions to algorithmic amplification because of their proven function on social learning bias (Brady et al., 2023). In their paper, the authors reviewed the concept of ”PRIME” information, which stands for ”Prestigious, Ingroup, Moral, and Emotional” content. The authors argue that during algorithm-mediated social learning, content algorithms on social media platforms exploit human social-learning biases and amplify PRIME information to maximize user engagement on the platforms. This process may bring side effects like promoting social misperceptions including conflicts and misinformation (2023).

Thus, it is plausible that the use of specific types of languages online can exacerbate polarization through algorithmic amplification and diffusion. The relationship between moralized and emotional language between algorithmic amplification and polarization is well-studied on social media platforms but has not been examined on news media outlets. The effect of other language dimensions, such as intergroup and prestige, on affective polarization is also a new research direction. One possible effect is that as algorithms prioritize content that generates high levels of engagement, news articles containing PRIME language may be more likely to be amplified and disseminated to ideologically aligned users. This increased exposure to PRIME content can reinforce partisan attitudes and contribute to the deepening of affective polarization.

2.3 Temporal Dynamic of Language Usage and Affective Polarization

Recent research has highlighted the potential influence of political context and electoral cycles on the use of moral language and sentiment in news media, as well as the dynamics of partisan attitudes. These findings provide a strong justification for comparing the use of PRIME (Prestigious, Ingroup, Moral, and Emotional) information by news outlets during election and non-election years and its potential impact on affective polarization.

Wang and Inbar (2021) demonstrate that moral language is used more frequently by media outlets when the opposing party is in power. This suggests that the strategic use of moral language by partisan media may be influenced by the broader political landscape, particularly the party in control of the government. In the context of election years, when the stakes are high and the competition for power is intense, it is plausible that news outlets may intensify their use of moral language to engage and mobilize their audiences, potentially contributing to heightened affective polarization.

The study by Michelitch and Utych (2018) provides additional evidence for the dynamic nature of partisan attitudes and the role of electoral cycles in shaping individuals’ attachment to political parties. The authors find that the predicted probability of feeling close to a political party increases by 6 percentage points from the midpoint of an electoral cycle to an election, an effect comparable to traditional determinants of partisanship. This finding suggests that partisan identities become more salient during election periods, which may amplify the impact of PRIME language used by news media on affective polarization.

The current study aims to address the gap in the literature by investigating the relationship between PRIME language usage in news media outlets and its potential impact on affective polarization through algorithmic amplification. By examining the presence of ingroup, moral, and emotional language in news articles across the political spectrum, this research seeks to shed light on how the linguistic features of partisan media content may contribute to the formation of echo chambers and the reinforcement of ideological divides. Furthermore, by comparing the use of PRIME language during election and non-election years, the study aims to provide insights into how the political context may influence the strategic use of language by media outlets to engage and mobilize their audiences.

3 Data and Methods

3.1 Data

This study utilizes a comprehensive dataset of news articles obtained from the News on the Web (NOW) corpus (https://www.english-corpora.org/now/). After removing non-US news, the data contains 11,688,074 articles. I removed articles from publishers that published less than 2000 articles over the entire dataset. This significantly reduces the number of publishers, thus making manual labeling of publisher bias easier. The resulting dataset consists of 9,658,888 articles published between January 2010 and July 2022, sourced from 649 unique publishers.

To analyze the relationship between language usage and political bias, I augmented the dataset with publisher metadata on partisan slant using ratings from AllSides (www.allsides.com). The use of AllSides ratings for determining media bias has been justified in previous studies, such as Huszár et al. (2022) and Rozado et al. (2022). I manually labeled the publishers and found 234 publishers with available bias ratings on AllSides. To focus on general news content, I excluded publishers that solely concentrate on entertainment, science, technology, and sports. After filtering out articles without bias ratings and deduplicating articles by full content, I have 2,510,592 articles for analysis.

Table 1 presents a sample of the dataset used in this study, showcasing the structure and variables included for each article.

The first column, "ID" shows the unique identifier assigned to the article. The second column, "Title" displays the title of the article. The "Text" column contains the full text of the article, which is the main source of data for analyzing language usage and sentiment. The "Standardized Publisher" column indicates the standardized name of the publisher or news outlet that published the article. The "Rating_Num" column represents the partisan bias rating of the publisher on a scale from 1 to 5, with 1 being the most liberal and 5 being the most conservative. The example article has a rating of 3, suggesting a centrist bias.

The "Year" and "Month" columns provide information about the publication date of the article.

ID	Title	Text	Standardized Publisher	Rating_Num	Year	Month
1334669	Article Title	Article Full Text	thenextweb	3	2010	8

Table 1: Example Article Data

The dataset consist of 234 publishers with partisan slant (bias rating) ranging from 1 (most liberal) to 5 (most conservative). A detailed distribution of number of articles for each bias rating is shown in 2. Additionally, this table 3 shows the number of publishers

Bias Rating	**Article Count**
1.0	190,669
2.0	1,006,330
3.0	1,078,157
4.0	140,803
5.0	94,633

Table 2: Article count by publisher rating number

for each partisan slant.

Rating Number	**Number of Publishers**	**Percentage (%)**
1.0	20	8.55
2.0	66	28.21
3.0	130	55.56
4.0	12	5.13
5.0	6	2.56

Table 3: Number and Percentage of Publishers by Rating Number.

Out of the total, publishers with a moderate bias (rating 3.0) constitute the majority, accounting for 55.56% of the dataset. This substantial proportion underscores the centrality of perspectives within the analyzed corpus. Liberal publishers (rating 1.0) and slightly liberal publishers (rating 2.0) together make up 36.75%, indicating a significant representation of left-leaning viewpoints. In contrast, conservative (rating 4.0) and highly conservative (rating 5.0) publishers are markedly less represented, comprising only 7.69% of the dataset. This uneven distribution of articles could skew analyses towards the language usage represented by the centrist and liberal publishers, potentially under-representing those from the conservative publishers.

Here, I show the top 3 publishers (ranked by total number of articles published) for each bias rating as an example of the type of news outlet included in this dataset. 3

Rating Number	Publisher	Article Count
1	Huffington Post	32,847
	Newyork Daily News	25,974
	The Boston Globe	20,275
2	Yahoo News	198,747
	Business Insider	89,659
	The Washington Post	60,511
3	Forbes	76,418
	Reuters	57,274
	Marketwatch	51,828
4	Newyork Post	66,015
	Washington Times	22,442
	Boston Herald	12,489
5	Fox News	57,022
	Washington Examiner	24,310
	Dailycaller	6,824

Table 4: Top three publishers by article count for each rating number.

3.2 Measures

3.2.1 PRIME Labeling

To label the presence of PRIME (Prestigious, Ingroup, Moral, and Emotional) content in the news articles, I used a novel Transformer-Dictionary Hybrid Approach developed by the Political NLP lab at University of Chicago. This approach combines the strengths of pre-trained language models and domain-specific dictionaries to accurately identify words associated with each PRIME dimension.

The first step in this approach was to develop an initial vocabulary of 71,879 words. The lab, including I, sourced these words from two main domains: political subreddits and our news dataset. By including words from both sources, we ensured that our vocabulary was comprehensive and representative of the language used in political discourse and news media.

It is important to note that because of the lack of standards for determining what kind of words are prestigious, I will only analyze the remaining three dimensions, intergroup, moral, and emotional. This is because prestige often comes from a specific person or organization's status. Thus, it is difficult to curate a list of prestige names and give the transformer an accurate definition as a training label.

Next, we fine-tuned pre-trained BERT models to detect PRIME content words. We utilized two specific models: Roberta-base and Distilbert-base-uncased. These models have been widely used in natural language processing tasks and have demonstrated strong performance in text classification.

To fine-tune the models, we developed training sets for each PRIME dimension. The training sets consisted of 206 moral words, 215 emotional words, and 206 intergroup words. These carefully curated sets of words served as the ground truth for training the models to recognize words associated with each dimension. The moral words are created using GPT4 and manual labeling, referencing foundations from the Moral Foundations Dictionary, developed by Frimer (2019). The intergroup words are also curated by GPT4 and manually selected. The emotional words are sourced from dictionary created by Shaver et al. (1987). We also used a dictionary API (https://dictionaryapi.dev/) to tag each word with dictionary definitions to improve model performance.

The model will generate probabilities for all non-stopwords in the vocabulary of 71,879 words. By applying the fine-tuned models to each word, I obtained probabilities indicating the likelihood of a word belonging to each PRIME dimension. These probabilities served as the basis for quantifying the presence of PRIME content in the news articles. Then, I added a sigmoid layer to enable multiple label output that classify a word to be emotional, intergroup and/or moral. For example, the output for word 'hate' could be: 1, 0, 1 (emotional, intergroup, moral), indicating 'hate' is a emotional and moral word.

We evaluated the performance of the fine-tuned models using a train-test split approach. The models achieved impressive results, with an F1 score of 0.83 for morality, 0.83 for emotion, and 0.82 for intergroup. These high F1 scores indicate that the models were able to accurately identify words belonging to each PRIME dimension.

To quantify the presence of PRIME, specifically intergroup, moral, and emotional, content in each article, I use the predicted categories from the labeled vocabulary list to count the occurrences of each PRIME dimension in the article. The counts are normalized into percentage scores by dividing them by the total number of words in the article.

Our Transformer-Dictionary Hybrid Approach offers several advantages. By leveraging pre-trained language models, we benefit from their ability to capture complex linguistic patterns and semantic relationships. The fine-tuning process allows us to adapt these models to the specific task of detecting PRIME content words. Additionally, the use of domain-specific dictionaries ensures that our approach is grounded in the relevant vocabulary and language used in political discourse and news media.

3.2.2 Sentiment Analysis

For sentiment analysis, I employed the VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis tool, which is a lexicon and rule-based sentiment analysis library specifically attuned to sentiments expressed in social media-style text (Hutto and Gilbert, 2014). VADER provides a compound score that represents the overall sentiment of a given text, ranging from -1 (most negative) to +1 (most positive). The compound score is calculated by summing the valence scores of each word in the lexicon, adjusted according to the rules, and then normalized to be between -1 and +1. To obtain the sentiment score for each article, I applied the SentimentIntensityAnalyzer from the VADER library to the full text of the article. The polarity_scores method was used to compute the sentiment scores, and we specifically extracted the compound score. This compound score serves as a standardized measure of the overall sentiment expressed in each article, allowing for comparisons across different publishers and time periods.

3.3 Analysis

3.3.1 Descriptive Analysis

I first present some descriptive analysis of the measures. This includes PRIME scores and sentiment scores for all 2,510,592 articles across 12 years (2010-2022). Here is a plot visualizing distribution of number of articles published each year 1. The plot shows a substantial increase in the number of articles starting in 2019, with a peak in 2021. The years preceding 2019 demonstrate relatively fewer publications, suggesting a potential expansion in media output or data availability in the later years.

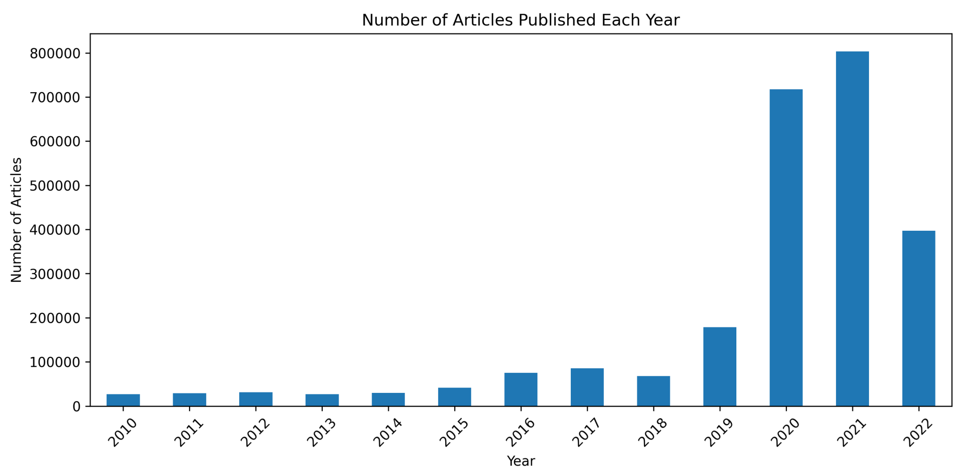

Figure 1: Number of Articles Published Each Year

The table 5 presents some basic statistics of the measures, with emotion (M = 0.023, SD = 0.014), intergroup (M = 0.111, SD = 0.046), and moral (M = 0.018, SD = 0.013) dimensions. The low scores on these dimensions are due to the fact that many words in an article do not belong to any of the three dimensions and, thus are labeled as other. The large percentage of other words makes scores of these dimensions small. The sentiment scores, (M = 0.334, SD = 0.848), show that the average sentiment in all articles is slightly positive. The standard deviation indicates a broader variability in the dataset's sentiment orientation, ranging from strongly negative to strongly positive expressions.

Statistic	Emotion	Intergroup	Moral	Sentiment
Mean	0.026	0.126	0.021	0.334
Standard Deviation	0.013	0.037	0.014	0.848
Minimum	0.000	0.000	0.000	-1.000
Maximum	0.207	0.588	0.220	1.000

Table 5: Descriptive statistics of Measures

To visualize the general trend in PRIME and sentiment, I aggregated data by yearly level and computed yearly average and confidence intervals for all measures. The trend is visualized in this figure 2.

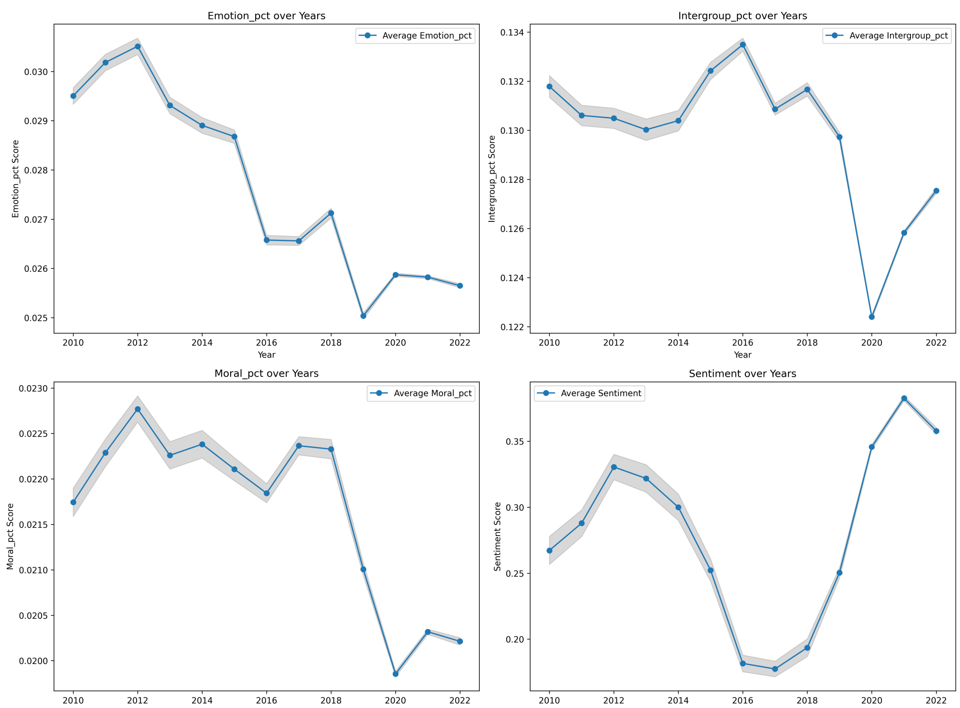

Figure 2: PRIME and Sentiment Trend in News from 234* Popular News Outlets

The figure depicts the trend of different dimensions—Emotion, Intergroup, Moral, and Sentiment—over a span from 2010 to 2022. Over all, PRIME language usage and sentiment score have decreased since 2012. This suggests news media are using less of these languages over the years. All graphs exhibit undulating patterns, suggesting that these contents in news media have fluctuated over the years. Peaks in these graphs may correspond to periods of heightened emotional intensity or moral engagement in the public discourse, potentially aligning with election cycles, major events, or crises that provoke public and media attention.

On the other hand, the sentiment scores have been decreasing after 2012 but started to increase after 2018. This suggests that news articles' sentiment was becoming negative until 2018, in which sentiment started to become more positive.

3.3.2 Analyzing Partisan Bias

To address the research question "How does the usage of emotional, moral, and interpersonal language differ among news publishers with different political biases?", I focused on examining the differences in language usage among news outlets with varying partisan slants.

I began by applying an ANOVA framework to assess differences across political biases. In this approach, PRIME language dimensions—emotion, intergroup, moral, and sentiment—were treated as dependent variables, while political bias served as the independent variable. This allowed me to test the null hypothesis that there were no differences in PRIME language usage between publishers of varying political orientations. The results of this analysis are described in 4.

To further enhance the analysis and account for the individual variability of each outlet, I employed a linear mixed-effects model (LMM). The LMM incorporates both fixed effects, which represent the factors of interest (in this case, political bias), and random effects, which account for the variability introduced by individual publishers. By including publisher-specific random intercepts, the LMM acknowledges the hierarchical structure of the data, where multiple observations from the same publisher are nested within the broader categories of political bias.

3.3.3 Analyzing Election Year Effect

The second research question (RQ2) explores the influence of election cycles on the use of PRIME (emotion, intergroup, moral) language and sentiment expression in news media content. The focus is on distinguishing language patterns during election years from those in non-election years to determine if the political climate surrounding elections has a discernible effect on the linguistic strategies employed by news publishers.

To facilitate the analysis, years were classified as 'election' or 'non-election' years. This classification was based on the established national election schedule, with federal election years in the United States occurring biennially. Given the time range of this dataset, the three election years would be 2012, 2016, and 2020.

To determine the impact of election years on PRIME language usage, I employed a two-way Analysis of Variance (ANOVA), with PRIME language dimensions as within-subjects factors and election year classification as a between-subjects factor. The model also accounted for the interaction between PRIME language types and the election year variable, allowing for a nuanced interpretation of how language usage may vary specifically during election years.

To account for the potential influence of individual publishers and the general passage of time on the relationship between election years and language usage, I also conducted a linear mixed-effects model analysis. This approach allows for a more nuanced examination of the election year effect while controlling for the variability introduced by publisher-specific characteristics and temporal trends.

By incorporating publisher-specific random effects, the LMM accounts for the hierarchical structure of the data, where articles are nested within publishers. This approach recognizes that language usage patterns may vary systematically between publishers due to factors such as editorial policies, target audience, or ideological leanings. The inclusion of random slopes for the year variable allows for the possibility that the temporal trends in language usage may differ across publishers, capturing potential heterogeneity in how language evolves over time within each publisher.

The results from both analyses will contribute to a more nuanced interpretation of how the political climate surrounding elections shapes the linguistic strategies employed by news publishers across different language dimensions.

4 Results

4.1 Partisan Bias and PRIME

4.1.1 ANOVA

The results of the ANOVA are shown here 6. For all four measure, the F-statistic are all extremely high, with significant p-value below 0.01 threshold. This result suggests that the group means for each of the dimensions are significantly different. Thus, the way emotion, intergroup, moral, and sentiment are presented in news articles varies significantly depending on the political bias of the publisher. This indicates that the political slant of a news outlet may influence how these psychological and linguistic dimensions are represented in news content.

Here, I visualize group differences with boxplot 3

Each box plot corresponds to a specific bias rating (1.0, 2.0, 3.0, 4.0, and 5.0) and displays the median, interquartile range, and potential outliers for the respective dimension.

The figure suggest a subtle yet consistent decrease in emotional language as we move

Dimension	F-statistic	p-value
Emotion	9450.97	< 0.001
Intergroup	6844.70	< 0.001
Moral	6420.95	< 0.001
Sentiment	5461.18	< 0.001

Table 6: ANOVA results for emotion, intergroup, moral, and sentiment dimensions.

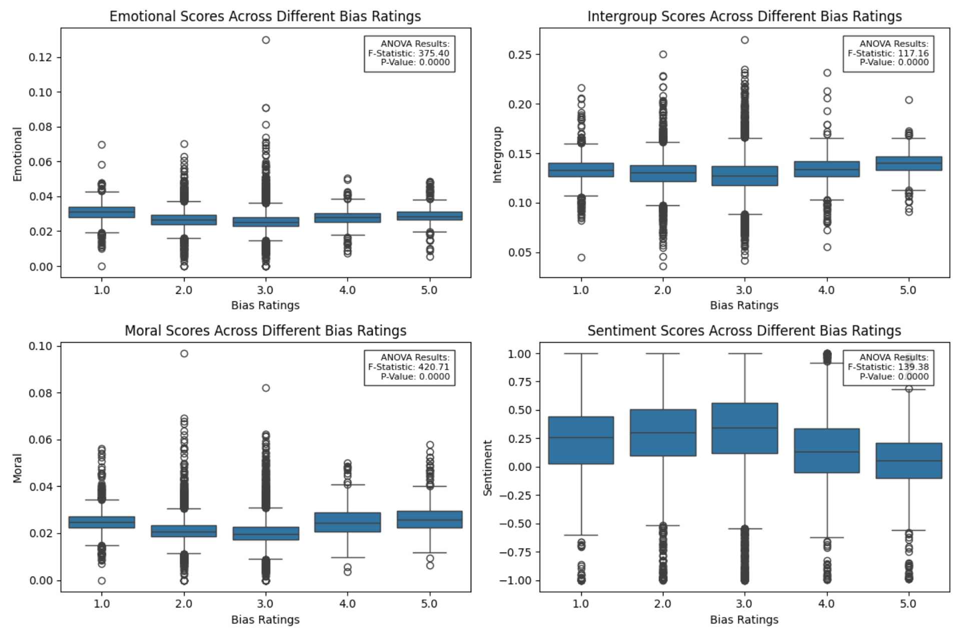

Figure 3: Distribution of emotional, intergroup, moral, and sentiment scores across different bias ratings.

from more partisan (both liberal and conservative) towards centrist outlets. This might indicate that more extreme political biases are associated with heightened emotional expression, potentially as a strategy to engage readers more strongly or to elicit specific emotional responses that align with their ideological stance.

A similar decreasing trend is observed in the moral scores from more partisan to centrist outlets, implying that publishers with strong political biases may more frequently invoke moral language.

The decrease in intergroup scores from more partisan to centrist publishers may suggest that publishers at the ends of the political spectrum focus more on language that identifies or discusses group dynamics, possibly highlighting in-group and out-group distinctions to a greater extent than centrist outlets.

Interestingly, sentiment scores tend to increase moving towards the center. This could mean that centrist outlets maintain a more positive or neutral tone in their reporting, avoiding the more polarized and potentially negative sentiment often found in partisan content. This might appeal to a broader audience seeking a balanced perspective, or it could indicate a strategic editorial choice to project impartiality and objectivity.

These analyses suggest that political bias influences the usage of emotional, moral, and interpersonal language as well as the sentiment of news content.

4.1.2 Mixed-effect for Emotion

The analysis of emotional language across different partisan biases was performed using a mixed linear model, whose results are summarized here 7. This model evaluated the effect of the partisan bias rating of news publishers on the percentage of emotional language in their content, controlling for inter-publisher variability.

Table 7: Mixed Linear Model Regression Results for Emotional Language

Parameter	**Coefficient**	**Std.Err.**	**z**	**P$>|z|$**	**[0.025**	**0.975]**
Intercept	0.030	0.001	48.790	0.000	0.029	0.031
C(Rating_Num)[T.2.0]	-0.004	0.001	-5.553	0.000	-0.005	-0.003
C(Rating_Num)[T.3.0]	-0.005	0.001	-7.362	0.000	-0.006	-0.004
C(Rating_Num)[T.4.0]	-0.003	0.001	-2.761	0.006	-0.005	-0.001
C(Rating_Num)[T.5.0]	-0.001	0.001	-0.497	0.619	-0.003	0.002
Group Var	0.000	0.000				

The model results indicate that there are significant differences in emotional language usage across partisan bias categories. Compared to the most liberal publishers (rating 1.0), publishers with a left-leaning bias (rating 2.0) and neutral publishers (rating 3.0) have significantly lower percentages of emotional language in their content, as indicated by the negative coefficients and significant p-values (p <0.001). Publishers with a right-leaning bias (rating 4.0) also have a significantly lower percentage of emotional language compared to the most liberal publishers, albeit with a smaller magnitude (p = 0.006).

Interestingly, the difference in emotional language usage between the most liberal publishers (rating 1.0) and the most conservative publishers (rating 5.0) is not statistically significant (p = 0.619). This suggests that while there is a general trend of decreasing emotional language usage as we move from left to right on the political spectrum, the most conservative publishers may employ emotional language to a similar extent as the most liberal publishers.

4.1.3 Mixed-effect for Intergroup

The analysis of intergroup language across different partisan biases was also performed using a mixed linear model. The results are summarized in Table 8.

Table 8: Mixed Linear Model Regression Results for Intergroup Language

Parameter	**Coefficient**	**Std.Err.**	**z**	**P>** $\vert z \vert$	**[0.025**	**0.975]**
Intercept	0.134	0.002	55.046	0.000	0.129	0.139
C(Rating_Num)[T.2.0]	-0.005	0.003	-1.772	0.076	-0.010	0.001
C(Rating_Num)[T.3.0]	-0.007	0.003	-2.520	0.012	-0.012	-0.001
C(Rating_Num)[T.4.0]	-0.006	0.004	-1.387	0.165	-0.013	0.002
C(Rating_Num)[T.5.0]	0.005	0.005	0.987	0.324	-0.005	0.015
Group Var	0.000	0.000				

The intercept coefficient (0.134) represents the estimated percentage of intergroup language for the reference category, which is the liberal publishers (rating 1.0).

Compared to the liberal publishers, left-leaning publishers (rating 2.0) have a 0.005 lower percentage of intergroup language, although this difference is not statistically significant (p = 0.076). Neutral publishers (rating 3.0) have a 0.007 lower percentage of intergroup language compared to the most liberal publishers, which is statistically significant (p = 0.012).

The difference in intergroup language usage between the liberal publishers and right-leaning publishers (rating 4.0) is not statistically significant (p = 0.165). Similarly, the difference between the liberal publishers and the conservative publishers (rating 5.0) is not statistically significant (p = 0.324).

These results suggest that while there are some differences in intergroup language usage across partisan bias categories, the overall pattern is less clear compared to the findings for emotional language. The most notable difference is between the most liberal publishers and neutral publishers, with neutral publishers using significantly less intergroup language.

4.1.4 Mixed-effect for Moral

The analysis results of moral language are summarized in Table 9.

The intercept coefficient estimated 0.024 percentage of moral language usage in the liberal publishers (rating 1.0). The coefficients for each rating category represent the difference in moral language usage compared to the reference category.

Compared to the liberal publishers, left-leaning publishers (rating 2.0) have a 0.003 lower percentage of moral language, which is statistically significant (p = 0.001). Similarly, centrist publishers (rating 3.0) have a 0.004 lower percentage of moral language compared to the liberal publishers, which is also statistically significant (p <0.001).

Table 9: Mixed Linear Model Regression Results for Moral Language

Parameter	**Coefficient**	**Std.Err.**	**z**	**P$>|z|$**	**[0.025**	**0.975]**
Intercept	0.024	0.001	33.583	0.000	0.022	0.025
C(Rating_Num)[T.2.0]	-0.003	0.001	-3.372	0.001	-0.004	-0.001
C(Rating_Num)[T.3.0]	-0.004	0.001	-4.739	0.000	-0.005	-0.002
C(Rating_Num)[T.4.0]	0.000	0.001	0.178	0.859	-0.002	0.002
C(Rating_Num)[T.5.0]	0.002	0.001	1.572	0.116	-0.001	0.005
Group Var	0.000	0.000				

The difference in moral language usage between the liberal publishers and right-leaning publishers (rating 4.0) is not statistically significant (p = 0.859). Likewise, the difference between the most liberal publishers and the most conservative publishers (rating 5.0) is not statistically significant (p = 0.116).

These results suggest that there are significant differences in moral language usage between the liberal publishers and both left-leaning and neutral publishers. The liberal publishers tend to use a higher percentage of moral language compared to these two categories. However, the differences between the most liberal publishers and right-leaning or the right publishers are not statistically significant.

4.1.5 Mixed-effect for Sentiment

The mixed linear model 10 assessing the relationship between partisan bias and sentiment in news content reveals some interesting patterns. The intercept coefficient (0.265) represents the estimated sentiment score for the publishers with a rating of 1.0, indicating a slightly positive sentiment on average.

Compared to the liberal publishers, the differences in sentiment scores for left-leaning publishers (rating 2.0), neutral publishers (rating 3.0), and right-leaning publishers (rating 4.0) are not statistically significant (p = 0.584, p = 0.160, and p = 0.270, respectively). This suggests that the sentiment expressed in news content does not differ significantly between the most liberal publishers and these other categories.

However, there is a significant difference in sentiment between the most liberal publishers and the conservative publishers (rating 5.0). The coefficient for the conservative publishers is -0.179 (p = 0.048), indicating that their news content tends to express a more negative sentiment compared to the liberal publishers.

The random effects part of the model, represented by the ”Group Var” term (0.038), suggests that there is considerable variability in sentiment scores attributed to differences between individual publishers within each bias rating category.

While there are no significant differences in sentiment between the liberal publishers and left-leaning, neutral, or right-leaning publishers, the conservative publishers tend to express a more negative sentiment in their news content. This suggests that the sentiment expressed in news articles may be more polarized at the extremes of the political spectrum.

Table 10: Mixed Linear Model Regression Results for Sentiment

Parameter	**Coefficient**	**Std. Err.**	**z**	**P> $|z|$**	**[0.025**	**0.975]**
Intercept	0.265	0.043	6.092	0.000	0.179	0.350
C(Rating_Num)[T.2.0]	0.027	0.050	0.547	0.584	-0.070	0.124
C(Rating_Num)[T.3.0]	0.066	0.047	1.406	0.160	-0.026	0.157
C(Rating_Num)[T.4.0]	-0.078	0.071	-1.102	0.270	-0.217	0.061
C(Rating_Num)[T.5.0]	-0.179	0.090	-1.976	0.048	-0.356	-0.001
Group Var	0.038	0.004				

When viewed together, these results underscore the profound impact of political bias on media content across multiple language dimensions. The ANOVA and mixed linear model analyses collectively suggest that political bias is a significant determinant in the representation of emotional, moral, and intergroup narratives, as well as sentiment within news articles.

The findings from the mixed linear models provide a more nuanced understanding of the relationship between partisan bias and language usage while accounting for publisher variability. The results indicate that liberal publishers tend to employ more emotional and moral language compared to left-leaning and neutral publishers. However, the usage of emotional and moral language does not differ significantly between liberal publishers and those on the right side of the political spectrum. For intergroup language, the most notable difference is between liberal publishers and neutral publishers, with neutral publishers using significantly less intergroup language.

In terms of sentiment, the mixed linear model reveals that conservative publishers tend to express a more negative sentiment in their news content compared to liberal publishers. This suggests that the sentiment expressed in news articles may be more polarized at the extremes of the political spectrum.

These findings support the hypothesis that partisan bias influences the use of PRIME language dimensions and sentiment in news content.

Building upon these findings, the next subsection delves into the relationship between PRIME language usage and the electoral context. By examining how the usage of emotional, moral, and intergroup language, as well as sentiment, varies between election and non-election years, I aim to uncover the potential influence of the political climate on the linguistic strategies employed by news publishers. This analysis will provide further insights into the dynamic nature of news content and its potential role in shaping public discourse during periods of heightened political activity.

4.2 PRIME and Election

4.2.1 ANOVA

The analysis was structured around a two-way Analysis of Variance (ANOVA) model, which allows for the assessment of the main effects of both PRIME language type (emotional, intergroup, moral), sentiment, and the temporal context (election vs. non-election years), as well as their interaction effect.

Table 11: ANOVA Results for the Impact of PRIME Type and Election Year on Language Usage

Source	Sum of Squares	df	F-value	P-value
C(PRIME_Type)	1.614×10^5	3	298458.23	< 0.001
C(Election_Year)	13.993	1	77.61	1.26×10^{-18}
C(PRIME_Type):C(Election_Year)	11.927	3	22.05	2.85×10^{-14}
Residual	1.811×10^6	10042352	NA	NA

The ANOVA results presented in Table 11 reveal significant effects of both PRIME Type and Election Year on language usage, as well as their interaction.

The significant F-value (77.61) for Election Year highlights a clear effect of the electoral cycle on language usage. The very low p-value (around zero) indicates that this effect is statistically significant, suggesting heightened or altered use of language during election years, which may relate to increased political activity and public engagement.

The interaction term has an F-value of 22.05, with a p-value of 2.85×10^{-14}, indicating that the influence of PRIME Type on language usage varies in election versus non-election years. This might suggest that certain language types are particularly emphasized in election years, possibly due to their greater persuasive or emotional impact during these politically charged periods.

4.2.2 Mixed-effect for Emotion

To further investigate the relationship between emotional language usage and election years, a mixed linear model was employed. The model output is presented in Table 12.

The mixed linear model results indicate a significant effect of Election Year on emotional language usage. The p-value (0.000) suggest that there is a statistically significant difference in the percentage of emotional language used during election years compared to non-election years. However, the coefficient of 0 suggests a statistically significant but practically negligible difference in emotion percentage between election and non-election years.

Interestingly, the Year variable, which represents the overall temporal trend, does not

Table 12: Mixed Linear Model Regression Results for Emotional Language Usage

Parameter	**Coefficient**	**Std.Err.**	**z**	**P>** $\lvert z \rvert$	**[0.025**	**0.975]**
Intercept	0.235	0.031	7.598	0.000	0.174	0.296
C(Election_Year)[T.True]	0.000	0.000	16.530	0.000	0.000	0.000
Year	-0.000	0.008	-0.034	0.973	-0.016	0.016
Group Var	0.114	0.433				
Group x Year Cov	-0.000	0.048				
Year Var	0.050					

have a significant effect on emotional language usage (p = 0.973). This suggests that while there may be differences in emotional language usage between election and non-election years, there is no consistent year-over-year trend in the data.

The random effects in the model account for variability in emotional language usage attributed to differences between individual publishers (Group Var) and potential interactions between publishers and the temporal trend (Group x Year Cov).

These findings support the hypothesis that the electoral context influences the use of emotional language in news content. However, the small magnitude of the effect suggests that while statistically significant, the practical difference in emotional language usage between election and non-election years may be limited.

4.2.3 Mixed-effect for Intergroup

To analyze the relationship between intergroup language usage and election years, a mixed linear model was employed. The model output is presented in Table 13.

Table 13: Mixed Linear Model Regression Results for Intergroup Language Usage

Parameter	**Coefficient**	**Std.Err.**	**z**	**P>** $\lvert z \rvert$	**[0.025**	**0.975]**
Intercept	1.091	0.417	2.615	0.009	0.273	1.909
C(Election_Year)[T.True]	-0.003	0.000	-59.875	0.000	-0.003	-0.003
Year	0.057					
Group Var	38.172	76.367				
Group x Year Cov	0.103					
Year Var	2.206					

The mixed linear model results indicate a significant effect of Election Year on intergroup language usage. The coefficient for Election Year (-0.003) and its associated p-value (0.000) suggest that there is a statistically significant difference in the percentage of intergroup language used during election years compared to non-election years. The negative coefficient indicates that intergroup language usage tends to be lower during election years.

The random effects in the model account for variability in intergroup language usage attributed to differences between individual publishers (Group Var) and potential interactions between publishers and the temporal trend (Group x Year Cov). The large variance estimates for these random effects suggest considerable variability in intergroup language usage across publishers and over time.

4.2.4 Mixed-effect for Moral

The mixed linear model regression results for moral language usage are presented in Table 14. This model evaluated the effect of election years and the passage of time (Year) on the percentage of moral language used in news content, while controlling for variations across different publishers.

Table 14: Mixed Linear Model Regression Results for Moral Language

Parameter	**Coefficient**	**Std.Err.**	**z**	**P$>|z|$**	**[0.025**	**0.975]**
Intercept	0.555	0.087	6.372	<0.001	0.384	0.726
Election Year (True)	-0.000	0.000	-20.614	<0.001	-0.000	-0.000
Year	0.003	–	–	–	–	–

Note: Group Variance = 1.481, Group x Year Covariance = -0.001, Year Variance = 0.752. Model converged: Yes.

The coefficient for Election Year indicate a significant effect of Election Year on moral language usage. Despite the statistical significance, the practical impact of this change is very minimal, as the coefficient is close to zero.

Group Variance (1.481) suggests significant variability in the baseline percentages of moral language usage among different publishers. Group x Year Covariance (-0.001) and Year Variance (0.752) indicate that there is variation in how moral language percentage changes over the years within publishers.

4.2.5 Mixed-effect for Sentiment

The model results for sentiment language usage across election cycles are summarized here 15

Table 15: Mixed Linear Model Regression Results for Sentiment

Parameter	**Coefficient**	**Std.Err.**	**z**	**P**$>\|z\|$	**[0.025**	**0.975]**
Intercept	-21.337	7.562	-2.822	0.005	-36.158	-6.516
Election Year (True)	-0.011	0.001	-8.529	< 0.001	-0.013	-0.008
Year	0.005					

Note: Group Variance = 12041.344, Group x Year Covariance = -8.513, Year Variance = 135.956. Model converged: Yes.

The coefficient for Election Year is -0.011, suggesting a slight but statistically significant decrease in sentiment during election years compared to non-election years. This implies that election years generally see a decrement in positive sentiment, possibly due to increased political tension and conflict featured in media content.

The Year coefficient, represented as 0.005, indicates a slight upward trend in sentiment scores over the years, although specific statistical details (p-value, standard error) are not provided in the summary. This suggests a gradual improvement in sentiment over time, which could reflect changes in media coverage or societal shifts in attitudes.

The model also incorporates random effects to account for variability across different publishers and over time within these groups:

- **Group Variance (12041.344):** This large variance suggests significant differences in baseline sentiment scores between publishers, indicating that each publisher has a distinct sentiment trend that can vary widely from others.
- **Group x Year Covariance (-8.513):** The negative covariance value implies that the direction of change in sentiment over the years may differ among publishers. This indicates that some publishers may show increasing sentiment over time, while others could show decreasing trends.

In conclusion, the results of this study provide compelling evidence for the influence of both political bias and electoral cycles on the use of emotional, moral, intergroup language, and sentiment in news media content. The ANOVA and mixed linear model analyses reveal significant differences in language usage across the political spectrum and between election and non-election years. This relationship is especially clear for usage of intergroup language, in which is reduced in election years, and more negative sentiment in election years.

Moreover, the variability in language usage patterns across individual publishers and over time highlights the complex interplay between media outlets, political contexts, and linguistic choices. The random effects in the mixed linear models indicate that publishers exhibit distinct baseline levels and temporal trends in their use of PRIME language and sentiment.

5 Discussion

The findings of this study provide valuable insights into the complex relationship between partisan bias, electoral cycles, and the use of emotional, moral, intergroup language, and sentiment in news media content. The results underscore the significant influence of political bias on the linguistic choices made by news publishers and highlight the dynamic nature of these choices in response to the political climate during election years.

The ANOVA and mixed linear model analyses reveal that the usage of emotional, moral, and intergroup language varies significantly across the political spectrum. Liberal publishers tend to employ more emotional and moral language compared to left-leaning and neutral publishers, suggesting that they may rely on these linguistic strategies to engage their audience and reinforce partisan narratives. Also, the lack of significant differences between liberal and conservative publishers in emotional and moral language usage indicates that these strategies may be employed by both end of the spectrum to reinforce partisanship.

Interestingly, the analysis of intergroup language reveals that neutral publishers use significantly less intergroup language compared to liberal publishers. This finding suggests that centrist outlets may be more cautious in their use of language that highlights group distinctions, possibly in an effort to maintain a more balanced and inclusive perspective. The decreased usage of intergroup language during election years, as revealed by the mixed linear model, further supports the notion that publishers may strategically adapt their language to navigate the heightened political sensitivity of election periods.

The sentiment analysis provides additional nuance to the understanding of partisan bias in news content. The analysis found that conservative publishers tend to express a more negative sentiment compared to other publishers. This difference in sentiment may contribute to the polarization of public opinion, as exposure to consistently negative content can reinforce partisan attitudes and increase hostility towards the out-group (Lelkes et al., 2017).

The examination of language usage in the context of electoral cycles reveals the dynamic nature of media strategies. The ANOVA results indicate a significant interaction between language dimensions and election years, suggesting that publishers may emphasize certain types of language more heavily during these politically charged periods. The mixed linear model for intergroup language provides further evidence of this strategic adaptation, showing a significant decrease in intergroup language usage during election years. This shift may reflect an attempt by publishers to avoid exacerbating group divisions and maintain a more inclusive narrative during a time of heightened political sensitivity.

However, the mixed linear model for sentiment reveals a slight decrease in positive sentiment during election years, which may be attributed to the increased coverage of political conflict and negative campaigning that often characterizes election periods. This finding highlights the complex interplay between the political climate and media content, as publishers navigate the competing demands of engaging audiences, maintaining journalistic integrity, and responding to the heightened intensity of election cycles.

These findings on election years' changed narratives complement previous literature that found moral language is used more frequently by media outlets when the opposing party is in power Wang and Inbar (2021) and the role of electoral cycles in increasing individuals' attachment to political parties Michelitch and Utych (2018). The current paper contribute to these findings by analyzing the narratives of news media and their changing strategies in the years of heightened political sensitivity.

Additionally, the variability in language usage patterns across individual publishers, as indicated by the random effects in the mixed linear models, underscores the importance of considering publisher-specific factors in the analysis of media bias. Editorial policies, target audiences, and organizational culture may all contribute to the distinct linguistic profiles of individual publishers, even within the same partisan category. Future research could delve deeper into these publisher-specific characteristics to provide a more comprehensive understanding of the factors that shape language usage in news media.

While this study offers valuable insights, it is important to acknowledge its limitations. The use of a dictionary-based approach for sentiment analysis, while widely accepted, may not capture the full nuance and context-dependence of language. Future studies could employ more advanced natural language processing techniques, such as sentiment analysis using transformer architectures (e.g., BERT), to obtain a more granular understanding of sentiment in news content.

Additionally, the study focuses on a specific set of language dimensions (emotional, moral, intergroup) and sentiment. While these dimensions were selected based on their theoretical relevance to partisan bias and polarization, there may be other linguistic features that contribute to the dynamics of media bias. Future research could expand the scope of analysis to include additional dimensions, such as the use of persuasive language, framing devices, or rhetorical strategies.

Despite these limitations, the present study makes a valuable contribution to the understanding of partisan bias and electoral influences on news media language. By combining computational methods with theoretical insights from political science and psychology, this research provides a nuanced perspective on the complex interplay between language, politics, and media.

6 Conclusion

This study offers a comprehensive examination of the relationship between partisan bias, electoral cycles, and the use of emotional, moral, intergroup language, and sentiment in news media content. Through a combination of ANOVA and mixed linear model analyses, the research reveals significant differences in language usage across the political spectrum and between election and non-election years.

The findings highlight the strategic employment of emotional and moral language by partisan publishers, particularly those on the liberal end of the spectrum, to engage audiences and reinforce partisan narratives. The decreased usage of intergroup language by neutral publishers and during election years suggests a more cautious approach to highlighting group distinctions in response to the political climate.

The sentiment analysis reveals a more negative sentiment expressed by conservative publishers compared to liberal ones, which may contribute to the polarization of public opinion. The examination of language usage in the context of electoral cycles indicates a significant interaction between language dimensions and election years, with publishers adapting their linguistic strategies to navigate the heightened political intensity of these periods.

While the study has some limitations, the findings have important implications for media literacy, as they highlight the subtle linguistic strategies employed by partisan outlets to influence public opinion and reinforce ideological echo chambers.

As the media landscape continues to evolve, it is crucial to develop a nuanced understanding of the factors that shape the language of news content. By shedding light on the complex interplay between partisan bias, electoral cycles, and linguistic choices, this study contributes to the ongoing discourse on the role of media in shaping public opinion and political polarization. The insights gained from this research can inform efforts to promote media literacy, encourage critical consumption of news content, and foster a more inclusive and balanced public dialogue.

In conclusion, this study demonstrates the significant influence of partisan bias and electoral cycles on the use of emotional, moral, intergroup language, and sentiment in news media content. The findings underscore the importance of considering the linguistic strategies employed by media outlets in shaping public opinion and highlight the need for further research to deepen our understanding of the complex dynamics between language, politics, and media in the digital age.

References

Brady, W. J., Gantman, A. P., and Van Bavel, J. J. (2020). Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology. General*, 149(4):746–756.

Brady, W. J., Jackson, J. C., Lindström, B., and Crockett, M. (2023). Algorithm-mediated social learning in online social networks. *Trends in Cognitive Sciences*, 27(10):947–960.

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., and Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, 114(28):7313–7318. Publisher: Proceedings of the National Academy of Sciences.

Frimer, J. (2019). Moral Foundations Dictionary 2.0. Publisher: OSF.

Hamborg, F., Donnay, K., and Gipp, B. (2019). Automated identification of media bias in news articles: an interdisciplinary literature review. *International Journal on Digital Libraries*, 20(4):391–415.

Huszár, F., Ktena, S. I., O'Brien, C., Belli, L., Schlaikjer, A., and Hardt, M. (2022). Algorithmic amplification of politics on Twitter. *Proceedings of the National Academy of Sciences*, 119(1):e2025334119. Publisher: Proceedings of the National Academy of Sciences.

Hutto, C. and 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*, 8(1):216–225. Number: 1.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., and Westwood, S. J. (2019). The Origins and Consequences of Affective Polarization in the United States. *Annual Review of Political Science*, 22(1):129–146.

Lelkes, Y., Sood, G., and Iyengar, S. (2017). The Hostile Audience: The Effect of Access to Broadband Internet on Partisan Affect. *American Journal of Political Science*, 61(1):5–20. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12237.

Ludwig, K., Grote, A., Iana, A., Alam, M., Paulheim, H., Sack, H., Weinhardt, C., and Müller, P. (2023). Divided by the Algorithm? The (Limited) Effects of Content- and Sentiment-Based News Recommendation on Affective, Ideological, and Perceived Polarization. *Social Science Computer Review*, 41(6):2188–2210. Publisher: SAGE Publications Inc.

Michelitch, K. and Utych, S. (2018). Electoral Cycle Fluctuations in Partisanship: Global Evidence from Eighty-Six Countries. *The Journal of Politics*.

Puglisi, R. and Snyder, J. M. (2011). Newspaper Coverage of Political Scandals. *The Journal of Politics*, 73(3):931–950. Publisher: The University of Chicago Press.

Rozado, D., Hughes, R., and Halberstadt, J. (2022). Longitudinal analysis of sentiment and emotion in news media headlines using automated labelling with Transformer language models. *PLOS ONE*, 17(10):e0276367. Publisher: Public Library of Science.

Shaver, P., Schwartz, J., Kirson, D., and O'Connor, C. (1987). Emotion knowledge: further exploration of a prototype approach. *Journal of Personality and Social Psychology*, 52(6):1061–1086.

Wang, S.-Y. N. and Inbar, Y. (2021). Moral-Language Use by U.S. Political Elites. *Psychological Science*, 32(1):14–26. Publisher: SAGE Publications Inc.