Same Words, Different Meanings: Semantic Polarization in Broadcast Media Language Forecasts Polarity in Online Public Discourse

Xiaohan Ding[1], Mike Horning[2], Eugenia H. Rho[1]

[1] Department of Computer Science, Virginia Tech
[2] Department of Communication, Virginia Tech
xiaohan@vt.edu, mhorning@vt.edu, eugenia@vt.edu

Abstract

With the growth of online news over the past decade, empirical studies on political discourse and news consumption have focused on the phenomenon of filter bubbles and echo chambers. Yet recently, scholars have revealed limited evidence around the impact of such phenomenon, leading some to argue that partisan segregation across news audiences cannot be fully explained by online news consumption alone and that the role of traditional legacy media may be as salient in polarizing public discourse around current events. This research expands the scope of analysis by including both online and more traditional media by investigating the relationship between broadcast news media language and social media discourse. By analyzing a decade's worth of closed captions (2.1 million speaker turns) from CNN and Fox News along with topically corresponding discourse from Twitter audiences who follow and engage with the two stations on social media, we provide a novel framework for measuring semantic polarization between America's two major broadcast networks to demonstrate how semantic polarization between these outlets has evolved (Study 1), peaked (Study 2) and influenced partisan discussions on social media (Study 3) across the last decade. Our results demonstrate a sharp increase in polarization in how topically important keywords are discussed between the two channels, especially after 2016, with overall highest peaks occurring in 2020. The two stations discuss identical topics in drastically distinct contexts in 2020, to the extent that there is barely any linguistic overlap in how identical keywords are contextually discussed. Further, we demonstrate how such partisan division in broadcast media language significantly impacts semantic polarity trends on Twitter (and vice-versa), empirically linking for the first time, how online discussions are influenced by televised media. We show how the language characterizing opposing media narratives about similar news events on TV can increase levels of partisan discourse online. To this end, our work has implications for how media polarization on TV plays a significant role in impeding rather than supporting online democratic discourse.

Introduction

Mass media plays a vital role in democratic processes by influencing how institutions operate, political leaders communicate, and most importantly, how citizens engage in politics (McLeod, Scheufele, and Moy 1999). Although it is no

surprise that America's two political divides speak different languages (Westfall et al. 2015), research has also shown that partisan language in news media has sharply increased in recent years, particularly in broadcast news (Horning 2018). This is concerning given that news consumption is critical for helping the public understand the events around them. According to Agenda Setting Theory, the language used by the media to frame and present current events impacts how the public perceives what issues are important (McCombs 1997; Russell Neuman et al. 2014).

While some may have the impression that mainstream legacy media is becoming less relevant amid the explosive growth of online news, American news consumption is still overwhelmingly from television, accounting for nearly five times as much as online news consumption across the public (Allen et al. 2020). Despite the notion that TV news consumption is more "passive" than reading the news, research shows that people tend to recall televised news better than online news (Eveland, Seo, and Marton 2002). Further, a recent study comparing TV vs. internet news consumption found that there are four times as many Americans who are partisan-segregated via TV than via online news (Muise et al. 2022). The study shows that TV news audiences are several times more likely to maintain their partisan news diets over time, and are much narrower in their sources while even partisan online news readers tend to consume from a variety of sources.

Yet studies on media polarization and ensuing public discourse are overwhelmingly based on online content (Garimella et al. 2021). For example, even research that analyzes data from traditional news outlets solely relies on tweets from the official Twitter accounts of newspapers, TV shows, and radio programs rather than the direct transcription of content from these legacy media sources (Recuero, Soares, and Gruzd 2020). This is due to the fact that unlike online information, legacy media data (e.g., closed captions) are harder to collect, exist in formats incompatible for quick pre-processing (e.g., srt files), and scattered across institutions that lack incentives to share data with academics. Hence, much of how mainstream legacy media affects online discourse is unknown.

In that sense, our analysis of a decade's worth of closed captions from 24hr-broadcast TV news programs from America's two major news stations presents a unique op-

portunity to empirically demonstrate how linguistic polarization in broadcast media has evolved over time, and how it has impacted social media discourse. In this work, we examine how semantic differences in broadcast media language have evolved over the last 11 years between CNN and Fox News (Study 1), what words are characteristic of the semantic polarity peaks in broadcast media language (Study 2), whether semantic polarity in TV news language forecasts polarization trends in social media discourse, and how language plays a role in driving relational patterns from one to the other (Study 3).

In Study 1, we leverage techniques in natural language processing (NLP) to develop a method that quantitatively captures how semantic polarization between CNN and Fox News has evolved from 2010 to 2020 by calculating the semantic polarity of how socially important, yet politically divided topics (*racism*, *Black Lives Matter*, *police*, *immigration*, *climate change*, and *health care*) are discussed by the two news channels. We then use a model interpretation technique in deep learning to linguistically unpack what may be driving these spikes by extracting contextual tokens that are most predictive of how each station discusses topical keywords in 2020 (Study 2). To investigate whether partisan trends in broadcast media language influence polarity patterns in social media discourse, we use Granger causality to test whether and how semantic polarization between the two TV news stations forecasts polarization across Twitter audiences replying to @CNN and @FoxNews (Study 3). Finally to understand the language that drives the Granger-causal relations in how semantic polarity trends in televised news affect that across Twitter users (and vice-versa), we identify tokens that are most predictive of how topical keywords are discussed on TV vs. Twitter, separated by lag lengths that correspond to Granger-causality significance. Our contributions are as follows:

- We provide a novel framework for quantifying semantic polarization between two entities by considering the temporality associated with how semantic polarization evolves over time. Prior research that quantifies polarization as an aggregate measure from a single longitudinal data dump often leaves out key temporal dynamics and contexts around how polarity unfolds across time. Our framework incorporates temporal fluctuations by computing diachronic shifts using contextual word embeddings with temporal features.

- In showing how semantic polarization in broadcast media has evolved over the last 11 years, we go beyond providing a mere quantification of polarization as a metric by using Integrated Gradients to identify attributive tokens as a proxy to understand the contextual language that drives the 2020 ascent in semantic polarity between the two news stations.

- We address the question of whether and how polarization in televised news language forecasts semantic polarity trends across Twitter, providing new evidence around how online audiences are shaped in their discourse by TV news language — an important link that has not been empirically established at-scale in prior research.

- Finally, we use model interpretation to extract lexical features from different entities, to show which words drive significant Granger-causal patterns in how broadcast media language shapes Twitter discourse and vice-versa, thereby highlighting the manner in which language plays a key role in driving semantic polarity relations between online discussions and broadcast media language.

Our findings are one of the first to quantify how language characterizing opposing media narratives about similar news events on TV can increase levels of partisan discourse online. Results from this work lend support to recent scholarship in communications research, which theorizes that both media and public agendas can influence each other, and that such dynamics can polarize the manner in which the public engages in discourse, thereby influencing democratic decision-making at-large.

Related Work

Temporal Dynamics in Linguistic Polarization

Scholars have measured linguistic polarization through stance (Dash et al. 2022), toxicity (Sap et al. 2019), sentiment detection (Yang et al. 2017), topic modeling (Ebeling et al. 2022), and lexicon-dictionaries (Polignano et al. 2022). Such methods typically capture an aggregated snapshot of polarization across large textual corpora, providing a static quantification of media bias based on preset criteria. As a result, such approaches seldom capture temporal fluctuations in semantic polarity. Only a few research to date quantify linguistic polarization over time. For example, researchers have measured partisan trends in congressional speech by defining polarization as the expected posterior probability that a neutral observer would infer a speaker's political party from a randomly selected utterance (Gentzkow, Shapiro, and Taddy 2019) across various issues. Demszky et al. use this Bayesian approach to capture how social media discussions on mass shootings become increasingly divisive (2020).

Such methods, however, capture linguistic polarization as a general metric across nearly *all* aggregated words within a textual corpus. Given the powerful salience of certain topical keywords that are inseparable from American politics, unlike prior work, we focus on how *specific* keywords polarize over time. We depart from previous approaches by devising a framework that captures semantic polarity as a function of how two different entities semantically diverge across time in their contextual use of identical words that are central to the American public discourse. Specifically, we measure linguistic polarization by capturing the diachronic shifts in how two news stations use the same topical words on their news programs over an 11-year period.

Further, in effort to understand the linguistic drivers behind how and why semantic polarization evolves, we aim to identify the source of diachronic shifts based on how two different entities use identical, politically contentious keywords. Few research has endeavored to detect causal factors underlying how the meaning of words changes over time, which remains an open challenge in NLP scholarship (Kutuzov et al. 2018). Hamilton et al. show how word meanings change between consecutive decades due to cultural

shifts (e.g., change in the meaning of "cell" driven by technological advancement: prison cell vs. cell phone) (2016). Yet such research typically makes a priori assumptions that the language between two textual corpora of comparison are significantly related without statistically demonstrating how so. On the other hand, research that does provide evidence of statistical relations between two separate textual data (Dutta, Ma, and Choudhury 2018), generally does not explain what words drive the direction of semantic influence from one corpus to another. We overcome these limitations: first, we statistically validate the existence of a significant semantic relationship between two textual corpora, TV news language and social media discourse, through Granger-causality (Study 3). Then, we strategically separate our data by lag-times associated with significant Granger-causal relations and apply a deep learning interpretation technique to demonstrate which contextual words from TV news language influences Twitter discussions around specific topical keywords (and vice-versa) over time.

Who Influences Who? Setting a Media Agenda

On any given day, newsrooms select news on a wide range of topics that cover a variety of issues. According to Agenda Setting Theory, this selection process leads audiences to see certain issues as more significant than others (McCombs 1997). Agenda Setting was particularly useful in decades past when media was much more consolidated. In the U.S. for example, three major networks (NBC, ABC, and CBS) provided the bulk of broadcast news content, and so their news agendas had broad influence on what audiences saw as important. However, today's media landscape is much more fragmented. People receive their news from any number of sources including television, newspapers, radio, online websites, and social media. In such an environment, Agenda Setting has certain limitations as the theory assumes that audiences are relatively passive and simply receive information from the media. This was true when news flowed one-way from the media to the mass public.

By contrast nowadays, the two-way form of communication afforded by the internet and social media makes it possible for the public to both influence media agendas and be influenced by them (Papacharissi 2009; Barnard 2018). As a result, new scholarship has argued through Intermedia Agenda Setting Theory that media agendas do not come entirely from within their own organizations, but from two other sources: other media and the mass audience. In the latter case, for example, audiences themselves can influence

	Topic	Topical Keywords
1	Racism	racism, racist
2	Black Lives Matter	blacklivesmatter
3	Police	police
4	Immigration	immigration, immigrants
5	Climate Change	climate change, global warming
6	Health Care	health care

Table 1: Our data consists of speaker turns that contain 9 keywords pertaining to six core topics.

media agendas through online affordances (e.g., retweeting, commenting, sharing), to raise certain issues to prominence online (Rogstad 2016). In this work, we explore these new emerging dynamics with an interest in not only understanding how news agendas are constructed, but how the possible adoption of different agendas might influence agenda dynamics as media interact with the public online.

Study 1: Evolution of Semantic Polarity in Broadcast Media Language (2010-2020)

NLP researchers studying diachronic shifts have been leveraging word embeddings from language models to understand how the meaning of a given word changes over time (Kutuzov et al. 2018). In Study 1, we adapt methodological intuitions from such prior work to calculate **how semantic polarization — *the semantic distance between how two entities contextually use an identical word* — evolves over time**. While scholars have examined online polarization in the context of understanding partisan framing and public discourse around current events, our work is the first to quantify and capture the evolution of semantic polarization in broadcast media language.

Dataset

Our data consists of transcribed closed captions from news shows that were broadcast 24-7 by CNN and Fox News from January 1, 2010 to December 31, 2020, which was provided by the Internet Archive and the Stanford Cable TV News Analyzer. We computationally extracted closed captions pertaining to each televised segment from a total of $181K$ SRT (SubRip Subtitle) files — consisting of 23.8 million speaker turns or 297 billion words spoken across the decade (Figure 1). To focus our analysis on how socially important yet politically controversial topics were discussed between the two major news networks, we narrowed down our sample of speaker turns to those that contained keywords associated with six politically contentious topics: *racism*, *Black Lives Matter*, *police*, *immigration*, *climate change*, and *health care* as shown in Table 1 ("blacklivesmatter" includes both "black lives matter" and its acronym, "BLM"). Keyword selection was based on their relevance to socially important, yet politically divided topics in the U.S. (Doherty et al. 2020; Medzerian 2021). While this selection may not be exhaustive, these keywords are representative of the controversial topics at the center of American political discourse during the study period (Doherty et al. 2021; Morin and Stepler 2016; Doherty, Kiley, and Johnson 2017; Funk and Kennedy 2016; Jones 2021; Allcott et al. 2020). We extracted all sentences containing one of the nine topical keywords, giving us a final corpus of 2.1 million speaker turns or 19.4 billion words. All closed captions pertaining to commercials were removed using a matching search algorithm. (Tuba, Akashe, and Joshi 2020).

Method and Analysis

With more recent advances in deep learning, researchers have leveraged the ability of transformer-based language models to generate contextualized embeddings to detect the
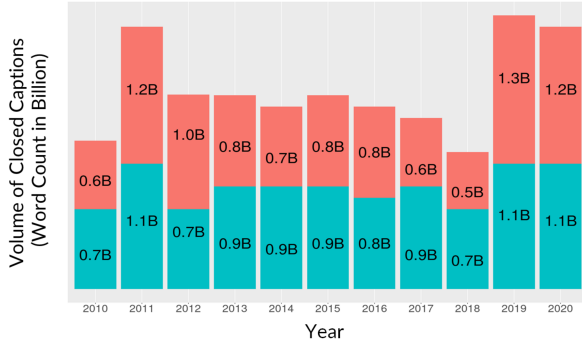
163

Figure 1: Volume of televised closed captions (word count) from CNN and Fox News stations from 2010 to 2020 (red: CNN; blue: FoxNews).

semantic change of words across time (Kutuzov et al. 2018). Similarly, we leverage contextual word representations from BERT (Bidirectional Encoder Representations from Transformers) (Devlin et al. 2019) to calculate the semantic polarity in how CNN and Fox News use topical keywords over time.

First, we use BERT to learn a set of embeddings for each keyword from the hidden layers of the model for every year from 2010 to 2020, to quantify how much the embeddings for a given topical keyword from CNN semantically differ from those of Fox across an 11-year period. We extracted all contextual word representations from our corpus of 2.1 million speaker turns. Then, for every occurrence of the 9 topical keywords in the data, we calculated the semantic distance between each contextual word representation from CNN and Fox using cosine values. The use of cosine values of BERT word embeddings is standard in machine translation and other NLP domains for computing contextual similarities in how tokens are used in sentences (Zhang et al. 2020). Our calculation of semantic polarization between how CNN vs. Fox News uses a topical keyword for a given year is as follows.

Given the set of word embeddings for a keyword in a given year t from CNN:

$$C_t = \{w_1^c, w_2^c, w_3^c, \ldots\} \quad (1)$$

and the set of word embeddings for a keyword in a given year t from Fox News:

$$F_t = \left\{w_1^f, w_2^f, w_3^f, \ldots\right\} \quad (2)$$

we define the semantic polarity (SP) between CNN and Fox News as:

$$\text{SP}(C_t, F_t) = \frac{\sum_{w_i^c \in C_t, w_j^f \in F_t} 1 - \cos\left(w_i^c, w_j^f\right)}{n_1 n_2} \quad (3)$$

where:
C_t : set of word embeddings from CNN
F_t : set of word embeddings from Fox News
w_i^c : word embedding of a keyword from CNN
w_j^f : word embedding of a keyword from Fox News
n_1 : number of word embeddings from CNN
n_2 : number of word embeddings from Fox News

In Figure 2, we visualize our methodological framework to illustrate how we derive the semantic polarization (SP) score in how CNN vs. Fox News uses the keyword, "racist" in 2010. First, we extract all contextual word embeddings (each of which is a 768×1 tensor) associated with every occurrence of the word "racist" in our corpus of speaker turns from Fox and CNN in 2010. Next, we calculate the cosine distance between the tensors associated with the word representation "racist" in the CNN corpus and those from the Fox News corpus. Finally, we use Equation 3 to calculate the average of all cosine distance values to obtain the semantic polarization score or the semantic distance between how Fox News and CNN uses the word "racist" in 2010.

Results

Between 2010 and 2020, semantic polarity between America's two major TV news stations increased with a significant upward trend starting in 2015 and 2016 (Figure 3). Over the past decade, CNN and Fox News became increasingly semantically divergent in how they used the topical keywords "*racist*", "*racism*", "*police*", "*blacklivesmatter*", "*immigrant*", "*immigration*", "*climate change*","*global warming*", and "*health care*" in their broadcasting language. Though there is a general dip in semantic polarity around 2013-14 across all keywords with the exception of "*health care*", semantic polarity between the two stations starts to sharply increase every year on average by 112% across all keywords starting in 2016. These results corroborate findings from a recent study analyzing the airtime of partisan actors on CNN and Fox, which shows that the two stations became more polarized in terms of visibility bias particularly following the 2016 election (Kim, Lelkes, and McCrain 2022).

Further, the highest peaks generally occur in 2020 for the terms "*racism*", "*racist*", "*blacklivesmatter*", "*police*", along with "*immigrant*", and "*global warming*" (Table 2). Specifically, semantic polarity is highest for the keyword "*police*" in 2020, denoting that across the nine keywords, CNN and Fox were most polarized in how they semantically used the word "*police*" in their news programs in 2020.

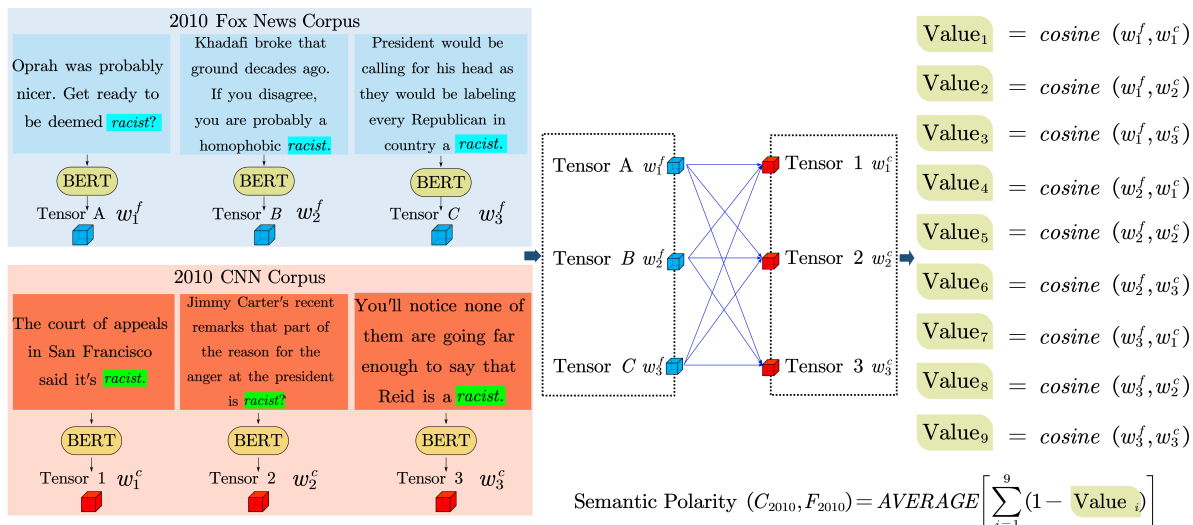| Keywords	Min SP	Max SP
1. racism	0.095(2014)	0.289(**2020**)
2. racist	0.100 (2014)	0.341(**2020**)
3. blacklivesmatter	0.148(2015)	0.320(**2020**)
4. police	0.119(2014)	0.574(**2020**)
5. immigration	0.103(2013)	0.362(2016)
6. immigrant	0.157(2013)	0.383(**2020**)
7. climate change	0.144(2013)	0.334(2018)
8. global warming	0.175(2014)	0.401(**2020**)
9. health care	0.076(2010)	0.496(2018)

Table 2: Range of 2020 semantic polarity scores.

Figure 2: Methodological framework of deriving semantic polarization (how CNN & Fox News use the keyword *racist* in 2010).
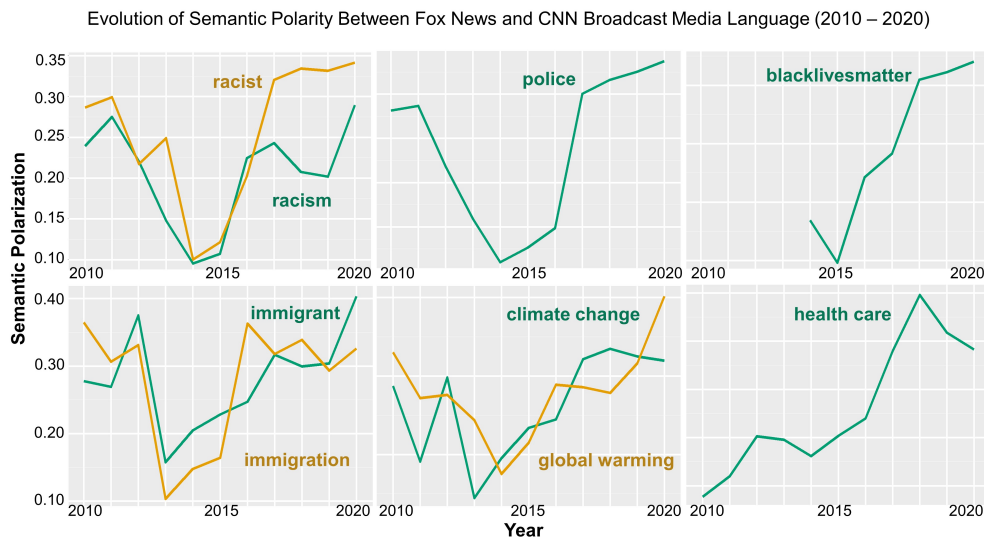
$$\text{Value}_1 = cosine\ (w_1^f, w_1^c)$$
$$\text{Value}_2 = cosine\ (w_1^f, w_2^c)$$
$$\text{Value}_3 = cosine\ (w_1^f, w_3^c)$$
$$\text{Value}_4 = cosine\ (w_2^f, w_1^c)$$
$$\text{Value}_5 = cosine\ (w_2^f, w_2^c)$$
$$\text{Value}_6 = cosine\ (w_2^f, w_3^c)$$
$$\text{Value}_7 = cosine\ (w_3^f, w_1^c)$$
$$\text{Value}_8 = cosine\ (w_3^f, w_2^c)$$
$$\text{Value}_9 = cosine\ (w_3^f, w_3^c)$$

$$\text{Semantic Polarity}\ (C_{2010}, F_{2010}) = AVERAGE\left[\sum_{i=1}^{9}(1 - \text{Value}_i)\right]$$

Figure 3: Evolution of semantic polarization in how CNN and Fox News use topically contentious keywords from 2010 to 2020.

Study 2: Words that Characterize Semantic Polarity between Fox News & CNN in 2020

Method and Analysis

Semantic polarity between Fox and CNN is predominantly high in 2020, echoing the series of highly publicized events in the media (e.g., COVID-19, George Floyd, and BLM demonstrations) that reified a widening ideological gap among the American public on a number of issues. Hence, in Study 2, we aim to go beyond providing a mere quantification of polarization by examining how the language surrounding the use of contentious keywords by CNN and Fox contributes to the 2020 peak in semantic polarity. In other words, **how does CNN and Fox News *contextually* differ in their use of topical keywords on their news pro-**grams in 2020? To answer this question, we identified attributive words or contextual tokens that are most predictive of whether a speaker turn containing a topical keyword was from CNN or Fox in 2020. This task entailed a two-step process. First, we trained a BERT-based classifier using the 2020 corpus of closed captions from both stations to predict whether a speaker turn was from CNN or Fox. Then, we used a model interpretation technique called Integrated Gradients (Sundararajan, Taly, and Yan 2017) to identify which words (except for the topical keyword) in each speaker turn contributed most to the classification decision.

Classifying 2020 Speaker Turns. To identify whether a speaker turn was broadcast from Fox or CNN in 2020, we built six BERT-based classifiers pertaining to each topic.

First, we merged the 2020 corpus of speaker turns, each containing one of the nine topical keywords by topic (e.g., speaker turns containing "racism" or "racist" were merged into one topical corpus).

For each of the six topical corpora, we fine-tuned BERT (`BERT-Base-Uncased`) to predict whether a speaker turn was from CNN or Fox News. For this step, we split the data into 80% training, 10% validation, and 10% test sets. We set batch sizes to 16, training epochs to 6 with 3-fold cross-validation, and used AdamW for optimization with a learning rate of 2e−05. We used PyTorch for all model implementations.

Model Interpretation. To identify the contextual language or the attributive words that contributed most to the 2020 semantic polarity between CNN and Fox News, we used Integrated Gradients (IG) to interpret the classification results from our BERT-classifier in the previous step. IG is an attribution method that informs how much an input feature contributes to the output result of a deep neural network (Sundararajan, Taly, and Yan 2017). This post-hoc interpretability technique combines the input's gradients by interpolating in short increments along a straight line between a baseline x (usually a vector with all zeros) and the input x'. Formally, if $F : R_n \rightarrow [0, 1]$ represents BERT, then the integrated gradient (IG) of the ith dimension of x is:

$$IG_i(x) = (x_i - x_i') \times \int_{\alpha=0}^{1} \frac{\partial F\left(x' + \alpha\left(x - x'\right)\right)}{\partial x_i} d\alpha \quad (4)$$

Large pre-trained transformer-based language models, such as BERT may have strong predictive powers, but their multiple nonlinearities prevent users from directly attributing model outputs to various input features. To overcome this lack of interpretive insight, we leverage IG to identify words that are most attributable to each predicted class. Thus, we identify which contextual words within a speaker turn (input features of a classifier) contribute most to, or are most predictive of the output decision, namely, whether the speaker turn is from CNN or Fox News in 2020.

Results

Our BERT-classifier distinguishes 2020 speaker turns from CNN and Fox News with a mean accuracy of 75.3% across all six topics (Table 3). Model performance is highest for speaker turns topically focused on the police, corresponding to Study 1 findings demonstrating highest SP scores for "police" across all keywords in 2020.

Topic	Acc.	F1	Prec.	Recall
Racism	0.710	0.709	0.707	0.712
BLM	0.724	0.726	0.710	0.743
Police	0.826	0.829	0.815	0.845
Immigration	0.773	0.780	0.762	0.800
Climate change	0.737	0.710	0.750	0.675
Health care	0.750	0.752	0.748	0.757

Table 3: Model performance for station classification.

In Table 4, we show the top 10 words most predictive of how CNN and Fox News discuss the six topics on their news programs in 2020. To avoid noise caused by low-frequency words, we limited our token selection to words with a frequency above the 95th percentile. We categorized the tokens with the highest and lowest attribution scores ranging from positive (attributed to CNN) to negative (attributed to Fox News) values for each topic. As shown, across all topics, none of the top tokens overlap between the two stations, indicating the drastically distinct contexts in which the two channels discuss key topics in 2020. For example, words like "*illegal*", "*enforcement*", and "*order*" contextualize immigration discussions in a rather legalistic manner on Fox News whereas, "*parents*", "*family*", "*children*", "*daughter*", and "*communities*" - humanizing words connotative of family relations best predict how CNN discusses the same topic. Both stations may be discussing the same topic with identical keywords, but the contextual language surrounding their discussion is strikingly different. Such divergence sheds light on the underlying differences in the linguistic signature of televised media language that help explain the 2020 peak in semantic polarity between the two stations.

Study 3: How Semantic Polarization in Broadcast Media Language Forecasts Semantic Polarity in Social Media Discourse

Study 1 shows that semantic polarization in broadcast media language has been increasing in recent years and overall peaking closer to 2020 during which the two major stations use starkly contrasting language in their discussion of identical words (Study 2). Such partisan divide in televised news language is similarly observed online where public discourse is becoming increasingly fraught with political polarization (Chinn, Hart, and Soroka 2020). To what degree is broadcast media one of the causes of online polarization in today's agora of public discourse on social media? **Do the temporal changes in semantic polarity in broadcast media significantly predict ensuing temporal changes in the semantic polarization across social media discourse?**

To answer this, we used Granger causality to test whether semantic polarity between CNN and Fox News broadcast media language over the last decade Granger-causes the semantic polarity between Twitter users who follow and mention @CNN or @FoxNews in their tweets that contain one of the nine keywords ("*racist*", "*racism*", "*police*", "*blackbivesmatter*", "*immigrant*", "*immigration*", "*climate change*", "*global warming*", and "*health care*") over the same time period.

Twitter Dataset

For our social media corpus, we collected all tweets between 2010 and 2020 based on the following criteria: (1) written by users who follow both @CNN and @FoxNews; (2) mention or are replying to either @CNN or @FoxNews; (3) contain one of the nine keywords. This criterion allows us to capture a sample that is likely to be exposed to how the two stations frame and discuss the 9 keywords on social media. Our final corpus amounted to 131, 627 tweets spanning over a decade. The volume of tweets and mean word count per tweet are shown in Table 5.

Table 4 (top half):

| Racism | | | | Black Lives Matter | | | | Climate Change | | | |
| CNN | | Fox News | | CNN | | Fox News | | CNN | | Fox News | |
Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.
president	0.457	called	-0.482	confederate	0.717	party	-0.505	human	0.483	gun	-0.440
whitehouse	0.348	anti	-0.442	people	0.350	video	-0.475	science	0.435	left	-0.419
senator	0.224	crime	-0.413	we	0.285	shooting	-0.454	scientists	0.289	democrats	-0.385
there	0.155	left	-0.398	symbol	0.281	mayor	-0.440	states	0.259	democratic	-0.369
protests	0.153	accused	-0.380	saw	0.261	anti	-0.400	today	0.244	agenda	-0.352
comments	0.146	school	-0.376	thousands	0.255	weekend	-0.394	climate	0.236	party	-0.349
he	0.141	criminal	-0.330	came	0.249	night	-0.351	california	0.207	progressive	-0.324
conversation	0.128	democratic	-0.325	players	0.247	democratic	-0.349	believe	0.207	right	-0.276
brutality	0.111	government	-0.319	know	0.240	police	-0.340	issues	0.203	voters	-0.264
ago	0.102	bad	-0.302	they	0.240	democrats	-0.335	important	0.199	Obama	-0.264

| Police | | | | Immigration | | | | Health Care | | | |
| CNN | | Fox News | | CNN | | Fox News | | CNN | | Fox News | |
Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.
different	0.414	seattle	-0.641	racist	0.370	illegal	-0.787	vaccines	0.735	illegal	-0.647
they	0.351	city	-0.537	Trump	0.298	laws	-0.422	vaccine	0.525	businesses	-0.568
certainly	0.346	portland	-0.512	parents	0.254	enforcement	-0.367	va	0.453	free	-0.477
racial	0.295	minnesota	-0.489	obama	0.252	police	-0.308	masks	0.386	business	-0.459
actually	0.285	chicago	-0.446	pan	0.212	legal	-0.303	deaths	0.336	open	-0.361
understand	0.285	left	-0.414	family	0.210	order	-0.288	virus	0.323	pay	-0.352
know	0.285	street	-0.406	vice	0.208	million	-0.278	faucci	0.295	economy	-0.338
shows	0.278	riots	-0.401	daughter	0.186	protect	-0.261	climate	0.283	jobs	-0.328
hands	0.269	democratic	-0.397	children	0.167	wants	-0.222	weeks	0.273	safe	-0.318
we	0.267	crime	-0.395	communities	0.167	law	-0.213	Trump	0.259	city	-0.311

Table 4: Top 10 tokens with the highest (attributed to CNN) and lowest (attributed to Fox News) attribution scores as a proxy for understanding words that linguistically characterize the 2020 semantic polarity between Fox News and CNN across six topics.

Hypothesis Testing with Granger Causality

To test whether temporal patterns in semantic polarity in TV news language significantly predict semantic polarity trends across Twitter users, we conducted a Granger causality test. Granger causality is a statistical test to determine whether a time series X is meaningful in forecasting another time series Y (Granger 1980). Simply, for two aligned time series X and Y, it can be said that X Granger-causes Y if past values $X_{t-l} \in X$ lead to better predictions of the current $Y_t \in Y$ than do the past values $Y_{t-l} \in Y$ alone, where t is the time point and l is the lag time or the time interval unit in which changes in X are observed in Y. Using this logic, we tested the following hypotheses:

H1. Semantic polarization on TV news significantly Granger-causes semantic polarization on Twitter.

H2. Semantic polarization on Twitter significantly Granger-causes semantic polarization on TV.

Method and Analysis

First, we computed the monthly semantic polarity (SP) scores spanning from 2010 to 2020 for a total of 132 time points (12 months $\times$ 11 years) for each topical keyword for both the televised closed captions and Twitter data. We then built 18 time series using the monthly SP values for each of the 9 keywords across 11 years derived from the closed captions and the tweets. We then paired each of the TV news language time series with the Twitter time series by keyword, giving us a total of 9 causality tests per hypothesis. As a prerequisite for Granger causality testing, we conducted

| | Tweet Volume | | Avg. Word Count (per Tweet) | |
Keyword	FoxNews	CNN	FoxNews	CNN
racism	8,787	8,944	24	22
racist	7,887	8,795	27	22
blacklivesmatter	4,863	4,528	21	23
police	8,731	8,670	24	22
immigration	8,163	7,883	19	21
immigrant	8,239	7,721	20	28
climate change	7,536	7,140	21	22
global warming	5,042	5,140	18	17
health care	7,612	6,943	20	19

Table 5: Volume of tweets and mean word account.

the Augmented Dickey-Fuller (ADF) test (Cheung and Lai 1995) to ensure that the value of the time series was not merely a function of time (see Appendix). We used the serial difference method (Cheung and Lai 1995) to obtain stationarity for ΔTwS_{racism} and ΔTvS_{BLM}, which were the only two time series with an ADF test value greater than the 5% threshold. In total, we conducted 18 time series analyses to test whether there was a Granger-causal relationship between the monthly SP values of broadcast media language (SP_{TV}) and that of Twitter ($SP_{Twitter}$) by each of the 9 keywords to determine the extent to which semantic polarization on TV drives semantic polarization on Twitter (*H1*) and vice versa (*H2*). For both hypotheses, we used a time-lag ranging from 1-12 months.

Results

We observe that changes in semantic polarity patterns on Twitter can be significantly predicted by semantic polarity trends in broadcast media language and vice-versa. Table 6, shows Granger causality results for hypotheses 1 and 2 with corresponding significant lag lengths in months ($p < 0.05$).

Hypothesis 1. As shown in Table 6, semantic polarity trends in broadcast media's discussion on racism (keywords: "**racism**", "**racist**") and immigration (keywords: "**immigration**" and "**immigrant**") significantly forecasts semantic polarity shifts in how people discuss these topics on Twitter with a time lag of 2 and 3 months, respectively. In other words, it takes about 2 months for the influence of semantic polarization in broadcast media language to manifest across how Twitter users semantically diverge in their discussion on racism, and about 3 months in how they talk about keywords linked to immigration.

Hypothesis 2. To test whether the directionality of influence is not only such that semantic polarization in televised news shapes polarity across Twitter conversations, but also vice-versa, we tested *H2*. Our results show that for "**climate change**" and "**global warming**", semantic polarization across how people discuss these keywords on Twitter Granger-causes how CNN and Fox News semantically diverge in their use of these words on their shows with a lag of 3 months. This may be due to the fact that discussions on climate change and global warming have become increasingly political and less scientific on Twitter (Jang and Hart 2015; Chinn, Hart, and Soroka 2020), therefore raising online prominence of the topic through controversial political sound bites that may in return, draw attention from mainstream media outlets as a source for news reporting.

Keyword	Hypos	Lag	F-value	p-value
racism	H1	2	1.70	**0.0386**
	H2	2	0.19	0.4526
racist	H1	2	3.17	**0.0101**
	H2	5	0.36	0.8730
blacklivesmatter	H1	3	2.90	**0.0376**
	H2	3	3.25	**0.0241**
police	H1	3	0.75	0.5189
	H2	5	0.50	0.7747
immigration	H1	3	2.13	**0.0333**
	H2	1	1.11	0.2929
immigrant	H1	3	2.19	**0.0325**
	H2	1	1.11	0.2929
climate change	H1	7	2.00	0.0502
	H2	3	3.18	**0.0262**
global warming	H1	8	0.96	0.4691
	H2	4	2.15	**0.0213**
heath care	H1	4	0.58	0.6738
	H2	3	0.72	0.5377

Table 6: H1 and H2 results using Granger causality tests. H1: $SP_{TV} \not\to SP_{Twitter}$; H2: $SP_{Twitter} \not\to SP_{TV}$.

Bidirectional Causality. For Black Lives Matter, there is bidirectional causality, meaning semantic polarity in how the keyword "**blacklivesmatter**" is discussed on televised news programs shapes semantic polarity trends in how users talk about the topic and vice-versa. This may be due to the fact that BLM is an online social movement, largely powered by discursive networks on social media. Hence, while users may draw upon the narratives generated by traditional media outlets in their Twitter discussions, journalists and broadcast media companies too, may be attuned to how online users discuss BLM issues on social media.

Words That Characterize the Relational Trends in Semantic Polarization Between Broadcast Media and Twitter Discourse

To understand the manner in which language plays a role in how semantic polarity trends in broadcast media forecasts polarization patterns on Twitter, for topical keywords with significant Granger-causality results, we identified contextual tokens that are most characteristic of how each topical keyword is discussed on TV vs. Twitter separated by the corresponding lag lengths based on results shown in Table 6.

Hence, given the 132 time points (12 months × 11 years) represented by t_{ij}, where i represents the month ($i \in [1, 12]$), j is the year ($j \in [1, 11]$), and l is the value of the lag length in months corresponding to significant Granger coefficients ($l \in [1, 8]$), we first re-organized our TV news data with $t_{i'j}$ time points, where i' represents the value of the maximum month minus l with the range: ($i' \in [1, 12 - l]$) and our Twitter data with $t_{i''j}$ time points, where i'' represents the value of the minimum month plus l with the range: ($i'' \in [1 + l, 12]$). For topical keywords where semantic polarity in Twitter Granger-causes semantic polarity trends in broadcast media language, we interchange the values of i' and i''. Next, following a similar approach to Study 2, we separately fed the TV news and Twitter corpora into a BERT classifier and used Integrated Gradients to identify the top 10 most predictive tokens for each topical keyword.

Results are shown in Tables 8 - 11 (Appendix) where we demonstrate the top 10 contextual tokens that are most attributive of how each topical keyword is used either by a speaker from CNN or Fox News stations or by a user responding to @CNN or @FoxNews in their tweet during time periods separated by monthly lag lengths corresponding to significant Granger-causal relationships.

As shown by the bolded tokens, a fair proportion of the top 10 words most characteristic of how televised news programs (CNN vs. Fox News) and Twitter users (replying to @CNN vs. @FoxNews) discuss topical keywords overlap quite a bit, suggesting that these tokens are semantically indicative of how linguistic polarity in one media manifests in another over time. For example, six of the top 10 tokens predictive of how CNN broadcast news contextually uses the keyword "racism" are identical to that of how users on Twitter, replying to @CNN, use "racism" in their tweets: "believe", "today", "president", "Trump", "campaign", "history", "issue".

For "blacklivesmatter" where the Granger-causality relationship is bidirectional, words most predictive of semantic

polarity in broadcast news that drive polarized discourse on Twitter are different from the tokens on Twitter that drive polarity in televised news language. For example, the set of tokens ("violence", "matter", "white", "support", "blacks") most predictive of how TV news programs on Fox News talk about BLM that appear three months later in tweets replying to @FoxNews, are strikingly different from the set of words ("blacks", "cities", "democrats", "lives", "moral", "organization", and "shoot") that best characterize the tweets replying to @FoxNews, which appear three months later on the Fox News channel.

Discussion

The rising levels of semantic polarity between the two major broadcast news organizations, as demonstrated in our work may render people's ability to reach across partisan divides and to perceive and solve issues democratically more difficult. The way CNN and Fox News discuss identical keywords on their programs is remarkably distinct in 2020 where semantic polarity between the two stations reaches its peak over an 11-year period, corroborating the widening partisan media gap highlighted by recent scholarship. Framing Theory argues that even subtle changes in the wording around how an issue is described by the media can significantly influence how audiences interpret the issue (Scheufele 2000; Entman 2003). Yet, our findings show that the contextual language driving semantic polarization in broadcast media is not nuanced at all. The surrounding words that characterize how each station discusses topically important keywords are drastically different, to the extent that identical words seem to reflect different meanings altogether. This is a crucial point of consideration, as linguistic cues in the media can play a powerful role in selectively reifying certain aspects of the perceived reality of an issue while downplaying others. Our findings suggest that TV news language does not only shape how online audiences perceive issues, but also how they talk about them.

In linguistics, collocation is the property of language where two or more words appear in each other's company with greater than random chance (Hoey 2005), such as "illegal" and "immigrants" in Fox News and "climate change" and "science" from CNN, as shown in our findings. Repeated encounters with collocated words drive what linguists call lexical priming (Hoey 2005), in which a priming word (e.g., "blacklivesmatter") provokes a particular target word ("protests" for CNN and "violence" for Fox) more strongly and quickly than disparate words encountered more rarely. In this vein, online audiences who consume very different perceived realities from the media may be lexically primed through repeated exposure to collocated words that frame identical issues in acutely contrasting contexts. This may help theorize why TV news language can shape how online audiences engage in public discourse as demonstrated in this work. Semantic polarization in televised media not only forecasts semantic polarity trends on Twitter, but the *words* that characterize broadcast media polarization re-appear across Twitter discussions, separated by significant lag months. Our results demonstrate that the language characterizing opposing media narratives around topically

similar news events on TV can be linguistically mimicked in how social media users are polarized in their discourse around important social issues.

Limitations. Differences in cosine values between contextual word representations stem not only from semantic, but also from positional differences of words in sentences. Although word embeddings generally embody contextual nuances, we acknowledge that our work predominantly considers semantic rather than syntactic aspects, and that to some extent our calculation of semantic polarity could embody syntactic differences in how keywords are discussed. Despite its strengths, as with any classical vector autoregressive models, Granger-causality assumes linear interactions, which might lead to inconsistent estimations when there are nonlinear dependencies (Tank et al. 2022). We recognize this as a potential limitation. Further, our findings pertain to American political contexts, which might not be generalizable to foreign public discourse and media language. For future work, we aim to apply our model to relevant textual corpora in other languages by incorporating cross-lingual transfer learning methods to better understand the generalizability of our framework by comparing distributional differences with cross-lingual data.

Appendix

Time Series	ADF	Threshold (1%)	Threshold (5%)	Conclusion
TwS_1	-8.52	-3.48	-2.88	stat
TwS_2	-1.86	-3.45	-2.88	non-stat
ΔTwS_2	-5.28	-2.69	-1.85	stat
TwS_3	-4.12	-3.43	-2.82	stat
TwS_4	-7.43	-3.48	-2.87	stat
TwS_5	-8.01	-3.48	-2.85	stat
TwS_6	-6.80	-3.47	-2.88	stat
TwS_7	-4.12	-3.48	-2.82	stat
TwS_8	-7.78	-3.46	-2.84	stat
TwS_9	-5.78	-3.48	-2.82	stat
TvS_1	-4.48	-3.48	-2.88	stat
TvS_2	-7.87	-3.48	-2.87	stat
TvS_3	-2.30	-3.21	-2.88	non-stat
ΔTvS_3	-8.53	-2.35	-1.74	stat
TvS_4	-3.59	-3.43	-2.83	stat
TvS_5	-8.85	-3.48	-2.83	stat
TvS_6	-3.79	-3.41	-2.86	stat
TwS_7	-5.22	-3.42	-2.81	stat
TwS_8	-7.17	-3.42	-2.89	stat
TwS_9	-4.45	-3.49	-2.88	stat

Table 7: ADF test results showing whether the Twitter and TV news time series data are stationary (stat) or non-stationary (non-stat). 1-9 refers to keyword number (see Table 2).

Table 8 header: Keyword "racism" — H1: $SP_{TV} \rightarrow SP_{Twitter}$ (lag: 2 months) | Keyword "racist" — H1: $SP_{TV} \rightarrow SP_{Twitter}$ (lag: 2 months)

CNN TV Token	Attr.	CNN Twitter (+2) Token	Attr.	FoxNews TV Token	Attr.	FoxNews Twitter (+2) Token	Attr.	CNN TV Token	Attr.	CNN Twitter (+2) Token	Attr.	FoxNews TV Token	Attr.	FoxNews Twitter (+2) Token	Attr.
believe	0.25	**believe**	0.18	**right**	-0.30	system	-0.17	systemic	0.27	**Trump**	0.20	**crime**	-0.37	**Africa**	-0.22
men	0.25	issues	0.18	**party**	-0.29	basis	-0.13	**president**	0.25	terrorism	0.19	states	-0.36	**politics**	-0.20
Americans	0.24	history	0.18	police	-0.29	**party**	-0.12	**hate**	0.22	children	0.19	**politics**	-0.31	crime	-0.20
today	0.24	he	0.17	marginalized	-0.29	**media**	-0.12	country	0.20	**president**	0.17	**Africa**	-0.30	standard	-0.15
president	0.22	**presidents**	0.16	**black**	-0.29	Africa	-0.12	think	0.18	**hate**	0.14	black	-0.29	Africans	-0.14
Trump	0.22	we	0.16	**media**	-0.28	diversity	-0.12	trying	0.17	**party**	0.13	Americans	-0.22	parents	-0.13
understand	0.21	hate	0.15	Obama	-0.26	guilty	-0.11	**Trump**	0.17	white	0.12	right	-0.21	share	-0.12
campaign	0.17	**today**	0.14	anti	-0.21	**right**	-0.10	**years**	0.16	bias	0.12	white	-0.20	please	-0.11
history	0.17	**campaign**	0.13	law	-0.21	suicide	-0.11	history	0.16	**years**	0.11	police	-0.16	**surveillance**	-0.08
issue	0.13	**Trump**	0.11	problem	-0.20	**black**	-0.09	**party**	0.15	presents	0.10	**surveillance**	-0.15	reforms	-0.08

Table 8: Top 10 tokens most predictive of how CNN and Fox News TV stations and Twitter users replying to @CNN and @FoxNews use keywords topically related to *racism*.

Table 9 header: Keyword "immigrant" — H1: $SP_{TV} \rightarrow SP_{Twitter}$ (lag: 3 months) | Keyword "immigration" — H1: $SP_{TV} \rightarrow SP_{Twitter}$ (lag: 3 months)

CNN TV Token	Attr.	CNN Twitter (+3) Token	Attr.	FoxNews TV Token	Attr.	FoxNews Twitter (+3) Token	Attr.	CNN TV Token	Attr.	CNN Twitter (+3) Token	Attr.	FoxNews TV Token	Attr.	FoxNews Twitter (+3) Token	Attr.
tax	0.26	charged	0.20	bid	-0.26	killer	-0.22	come	0.25	**white**	0.22	law	-0.26	**legally**	-0.24
united	0.26	make	0.20	**health**	-0.25	**right**	-0.21	family	0.24	law	0.21	people	-0.24	scouts	-0.20
Trump	0.26	**united**	0.19	**immigration**	-0.25	**immigration**	-0.21	**united**	0.23	**united**	0.19	**policies**	-0.22	**policies**	-0.17
country	0.25	parents	0.17	**right**	-0.24	healthcare	-0.18	**rights**	0.22	**Russia**	0.18	**public**	-0.20	international	-0.17
care	0.24	family	0.16	state	-0.20	**American**	-0.16	**country**	0.22	**country**	0.17	president	-0.18	safety	-0.16
know	0.24	temporarily	0.15	people	-0.19	peaceful	-0.14	**security**	0.21	visa	0.16	**legally**	-0.17	**public**	-0.15
think	0.21	**country**	0.14	**president**	-0.15	**health**	-0.13	Trump	0.21	**security**	0.15	federal	-0.16	**illegal**	-0.14
white	0.20	**immigrants**	0.14	**American**	-0.15	president	-0.12	**white**	0.20	court	0.12	system	-0.15	election	-0.14
immigrants	0.19	prison	0.13	law	-0.11	liberal	-0.12	**Russia**	0.19	police	0.12	order	-0.15	praises	-0.11
lawyer	0.18	obviously	0.11	new	-0.10	Trump	-0.12	way	0.19	**rights**	0.10	**illegal**	-0.15	Trump	-0.10

Table 9: Top 10 tokens most predictive of how CNN and Fox News TV stations and Twitter users replying to @CNN and @FoxNews use keywords topically related to *immigration*.

Table 10 header: Keyword "climate change" — H2: $SP_{Twitter} \rightarrow SP_{TV}$ (lag: 3 months) | Keyword "global warming" — H2: $SP_{Twitter} \rightarrow SP_{TV}$ (lag: 4 months)

CNN TV (+3) Token	Attr.	CNN Twitter Token	Attr.	FoxNews TV (+3) Token	Attr.	FoxNews Twitter Token	Attr.	CNN TV (+4) Token	Attr.	CNN Twitter Token	Attr.	FoxNews TV (+4) Token	Attr.	FoxNews Twitter Token	Attr.
science	0.28	**years**	0.24	president	-0.26	**policy**	-0.23	**important**	0.24	human	0.23	**Obama**	-0.23	**Obama**	-0.24
years	0.24	threatened	0.23	**right**	-0.25	job	-0.22	**water**	0.23	drugs	0.23	future	-0.23	hurricanes	-0.24
primarily	0.24	**science**	0.21	**democrats**	-0.22	struggle	-0.22	**science**	0.23	government	0.22	national	-0.22	protect	-0.23
time	0.24	national	0.21	**economy**	-0.22	presidential	-0.21	lives	0.22	**water**	0.20	rights	-0.22	communities	-0.22
people	0.23	geography	0.21	jobs	-0.22	**right**	-0.20	**health**	0.22	**condition**	0.19	**Trump**	-0.21	**Trump**	-0.22
health	0.23	diseases	0.19	**policy**	-0.22	violence	-0.20	**condition**	0.20	teenager	0.18	**gas**	-0.20	**forest**	-0.21
world	0.23	**ocean**	0.18	care	-0.21	diet	-0.19	world	0.19	**health**	0.17	**economy**	-0.20	**economy**	-0.17
think	0.22	**issues**	0.13	energy	-0.20	**economy**	-0.19	fossil	0.19	marketing	0.16	energy	-0.20	they	-0.17
ocean	0.22	**world**	0.11	**Trump**	-0.18	**democrats**	-0.19	issue	0.19	**science**	0.16	immigration	-0.19	USA	-0.17
issues	0.21	predator	0.10	global	-0.16	**Trump**	-0.17	environment	0.17	**important**	0.14	green	-0.17	**gas**	-0.15

Table 10: Top 10 tokens most predictive of how CNN and Fox News TV stations and Twitter users replying to @CNN and @FoxNews use keywords topically related to *climate change*.

Keyword: "BlackLivesMatter"								Keyword: "BlackLivesMatter"							
H1: $SP_{TV} \rightarrow SP_{Twitter}$ (lag: 3 months)								H2: $SP_{Twitter} \rightarrow SP_{TV}$ (lag: 3 months)							
CNN				FoxNews				CNN				FoxNews			
TV		Twitter (+3)		TV		Twitter (+3)		TV (+3)		Twitter		TV (+3)		Twitter	
Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.	Token	Attr.
know	0.33	**protests**	0.24	**right**	-0.24	Baptist	-0.24	**Americans**	0.25	**Americans**	0.21	**cities**	-0.22	**blacks**	-0.18
going	0.32	**people**	0.24	country	-0.24	moral	-0.22	**movement**	0.20	church	0.21	**democrats**	-0.20	**cities**	-0.18
people	0.31	technology	0.23	lot	-0.21	**right**	-0.21	plaza	0.21	citizens	0.18	**lives**	-0.20	**democrats**	-0.20
plaza	0.28	county	0.21	**violence**	-0.20	toxic	-0.18	**president**	0.23	**movement**	0.23	mayor	-0.20	federal	-0.21
lives	0.27	**movement**	0.21	way	-0.20	praising	-0.17	**states**	0.21	powerful	0.17	**moral**	-0.23	Jackson	-0.24
that	0.27	**lives**	0.20	**white**	-0.19	**blacks**	-0.17	that	0.21	**president**	0.14	officers	-0.21	liberals	-0.19
movement	0.24	family	0.20	**support**	-0.18	**violence**	-0.16	they	0.21	protests	0.20	**organization**	-0.19	**lives**	-0.24
president	0.22	**protesters**	0.20	blacks	-0.18	**support**	-0.15	today	0.20	**states**	0.24	party	-0.24	**moral**	-0.20
protesters	0.18	virus	0.20	want	-0.18	white	-0.15	**world**	0.21	**world**	0.13	**shoot**	-0.18	**organization**	-0.18
protests	0.18	**president**	0.19	homicides	-0.17	matters	-0.13	**year**	0.24	**year**	0.17	**cities**	-0.22	**shoot**	-0.21

Table 11: Top 10 tokens most predictive of how CNN and Fox News TV stations and Twitter users replying to @CNN and @FoxNews use keywords topically related to *Black Lives Matter*.

Ethics Statement

No personal information of Twitter users was collected nor compromised throughout our project. All data in this work are securely stored on servers only accessible by the authors. Our semantic polarization framework is made publicly available on the authors' GitHub page and can be applied on a variety of textual corpora with diachronic and contrastive properties.

References

Allcott, H.; Boxell, L.; Conway, J.; Gentzkow, M.; Thaler, M.; and Yang, D. 2020. Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics*, 191: 104254.

Allen, J.; Howland, B.; Mobius, M.; Rothschild, D.; and Watts, D. J. 2020. Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14): eaay3539.

Barnard, S. R. 2018. Citizens at the Gates. *Twitter, Networked Publics and The Transformation of American Journalism. Cham, Sveitsi: Palgrave Macmillan.*

Cheung, Y.-W.; and Lai, K. S. 1995. Lag Order and Critical Values of the Augmented Dickey–Fuller Test. *Journal of Business & Economic Statistics*, 13(3): 277–280.

Chinn, S.; Hart, P. S.; and Soroka, S. 2020. Politicization and Polarization in Climate Change News Content, 1985-2017. *Science Communication*, 42(1): 112–129.

Dash, S.; Mishra, D.; Shekhawat, G.; and Pal, J. 2022. Divided We Rule: Influencer Polarization on Twitter during Political Crises in India. *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 16: 135–146.

Demszky, D.; Movshovitz-Attias, D.; Ko, J.; Cowen, A.; Nemade, G.; and Ravi, S. 2020. GoEmotions: A dataset of fine-grained emotions. *In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4040–4054.

Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, 4171–4186.

Doherty, C.; Kiley, J.; Asheer, N.; and Jordan, C. 2020. Important issues in the 2020 election, Pew Research Center, August 2020.

Doherty, C.; Kiley, J.; Asheer, N.; and Jordan, C. 2021. Deep Divisions in Americans' Views of Nation's Racial History – and How To Address It, Pew Research Center, August 2021.

Doherty, C.; Kiley, J.; and Johnson, B. 2017. Partisans Differ Widely in Views of Police Officers, College Professors, Pew Research Center, September 2017.

Dutta, S.; Ma, J.; and Choudhury, M. D. 2018. Measuring the Impact of Anxiety on Online Social Interactions. *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 12(1): 584–587.

Ebeling, R.; Saenz, C. A. C.; Nobre, J. C.; and Becker, K. 2022. Analysis of the Influence of Political Polarization in the Vaccination Stance: The Brazilian COVID-19 Scenario. *Proceedings of the International AAAI Conference on Web and Social Media*, 16: 159–170.

Entman, R. M. 2003. Cascading activation: Contesting the White House's frame after 9/11. *Political Communication,*, 20(4): 415–432.

Eveland, W. P.; Seo, M.; and Marton, K. 2002. Learning From the News in Campaign 2000: An Experimental Comparison of TV News, Newspapers, and Online News. *Media Psychology*, 4(4): 353–378.

Funk, C.; and Kennedy, B. 2016. The Politics of Climate, Pew Research Center, October 2016.

Garimella, K.; Smith, T.; Weiss, R.; and West, R. 2021. Political Polarization in Online News Consumption. *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 15: 152–162.

Gentzkow, M.; Shapiro, J. M.; and Taddy, M. 2019. Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech. *Econometrica*, 87(4): 1307–1340.

Granger, C. W. J. 1980. Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control*, 2: 329–352.

Hamilton, W. L.; Leskovec, J.; and Jurafsky, D. 2016. Cultural Shift or Linguistic Drift? Comparing Two Computational Measures of Semantic Change. *Conference on Empirical Methods in Natural Language Processing.*, 2116–2121.

Hoey, M. 2005. *Lexical Priming: A new theory of words and language*. London: Routledge.

Horning, M. 2018. The pundit problem: A look at bias and negativity in cable news coverage as the 2016 election came to a close. *The 2016 American Presidential Campaign and the News*, 77–99.

Jang, S. M.; and Hart, P. S. 2015. Polarized frames on "climate change" and "global warming" across countries and states: Evidence from Twitter big data. *Global Environmental Change*, 32: 11–17.

Jones, J. 2021. Americans Remain Divided on Preferred Immigration Levels, Gallup Organization. Section: Politics, July 23.

Kim, E.; Lelkes, Y.; and McCrain, J. 2022. Measuring dynamic media bias. *Proceedings of the National Academy of Sciences*, 119(32): e2202197119.

Kutuzov, A.; Øvrelid, L.; Szymanski, T.; and Velldal, E. 2018. Diachronic word embeddings and semantic shifts: a survey. In *Proceedings of the 27th International Conference on Computational Linguistics*, 1384–1397.

McCombs, M. 1997. Building Consensus: The News Media's Agenda-Setting Roles. *Political Communication*, 14(4): 433–443.

McLeod, J. M.; Scheufele, D. A.; and Moy, P. 1999. Community, Communication, and Participation: The Role of Mass Media and Interpersonal Discussion in Local Political Participation. *Political Communication*, 16(3): 315–336.

Medzerian, D. 2021. Political divide remains critically high and immigration is most divisive issue, according to new USC polarization index, USC News, Nov.4.

Morin, R.; and Stepler, R. 2016. The Racial Confidence Gap in Police Performance, Pew Research Center, September 2016.

Muise, D.; Hosseinmardi, H.; Howland, B.; Mobius, M.; Rothschild, D.; and Watts, D. J. 2022. Quantifying partisan news diets in Web and TV audiences. *Science Advances*, 8(28): eabn0083.

Papacharissi, Z. 2009. *Journalism and citizenship*. Routledge London.

Polignano, M.; Basile, V.; Basile, P.; Gabrieli, G.; Vassallo, M.; and Bosco, C. 2022. A hybrid lexicon-based and neural approach for explainable polarity detection. *Information Processing & Management*, 59(5): 103058.

Recuero, R.; Soares, F. B.; and Gruzd, A. 2020. Hyperpartisanship, Disinformation and Political Conversations on Twitter: The Brazilian Presidential Election of 2018. *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 14: 569–578.

Rogstad, I. 2016. Is Twitter just rehashing? Intermedia agenda setting between Twitter and mainstream media. *Journal of Information Technology & Politics*, 13(2): 142–158.

Russell Neuman, W.; Guggenheim, L.; Mo Jang, S.; and Bae, S. Y. 2014. The Dynamics of Public Attention: Agenda-Setting Theory Meets Big Data. *Journal of Communication*, 64(2): 193–214.

Sap, M.; Card, D.; Gabriel, S.; Choi, Y.; and Smith, N. A. 2019. The Risk of Racial Bias in Hate Speech Detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 1668–1678.

Scheufele, D. A. 2000. Agenda-setting, priming, and framing revisited: Another look at cognitive effects of political communication. *Mass communication & society*, 3(2-3): 297–316.

Sundararajan, M.; Taly, A.; and Yan, Q. 2017. Axiomatic Attribution for Deep Networks. In *Proceedings of the 34th International Conference on Machine Learning*, 3319–3328.

Tank, A.; Covert, I.; Foti, N.; Shojaie, A.; and Fox, E. B. 2022. Neural Granger Causality. *IEEE transactions on pattern analysis and machine intelligence*, 44(8): 4267–4279.

Tuba, M.; Akashe, S.; and Joshi, A., eds. 2020. *Information and Communication Technology for Sustainable Development*, volume 933. Advances in Intelligent Systems and Computing.

Westfall, J.; Van Boven, L.; Chambers, J. R.; and Judd, C. M. 2015. Perceiving Political Polarization in the United States: Party Identity Strength and Attitude Extremity Exacerbate the Perceived Partisan Divide. *Perspectives on Psychological Science*, 10(2): 145–158.

Yang, M.; Wen, X.; Lin, Y.-R.; and Deng, L. 2017. Quantifying Content Polarization on Twitter. In *2017 IEEE 3rd International Conference on Collaboration and Internet Computing (CIC)*, 299–308.

Zhang, T.; Kishore, V.; Wu, F.; Weinberger, K. Q.; and Artzi, Y. 2020. BERTScore: Evaluating Text Generation with BERT. arXiv:1904.09675.