Fostering Civil Discourse Online: Linguistic Behavior in Comments of #MeToo Articles Across Political Perspectives

EUGENIA HA RIM RHO*
GLORIA MARK*
MELISSA MAZMANIAN*

*Department of Informatics, School of Information and Computer Sciences, University of California, Irvine

Linguistic style and affect shape how users perceive and assess political content on social media. Using linguistic methods to compare political discourse on far-left, mainstream and alt-right news articles covering the #MeToo movement, we reveal rhetorical similarities and differences in commenting behavior across the political spectrum. We employed natural language processing techniques and qualitative methods on a corpus of approximately 30,000 Facebook comments from three politically distinct news publishers. Our findings show that commenting behavior reflects how social movements are framed and understood within a particular political orientation. Surprisingly, these data reveal that the structural patterns of discourse among commenters from the two alternative news sites are similar in terms of their relationship to those from the mainstream – exhibiting polarization, generalization, and othering of perspectives in political conversation. These data have implications for understanding the possibility for civil discourse in online venues and the role of commenting behavior in polarizing media sources in undermining such discourse. 1

CCS Concepts: • Computing methodologies → Artificial intelligence; Natural language processing; Discourse, dialogue and pragmatics

KEYWORDS
News media; SNS; political discourse; polarization; hashtags; online social movements; embedding; TF-IDF; word2vec; NLP

ACM Reference format:

1 INTRODUCTION

Obtaining political news or information through online news media is commonplace in the current digital age. Many, if not most, online news sites offer functionality for commenting on news items, providing a public sphere where people can engage in modern day political discourse [73, 91, 96, 99]. Most news sites have a presence on social networking sites (SNS) such as Facebook, where discussions on moral and political issues routinely take place [44, 55, 73, 96, 106]. In fact, over the past few years users have been increasingly using SNS to discover news...
users are not only reading, but increasingly sharing and commenting on personal thoughts, asking questions, and engaging in discussion with those connected within their networks [2, 44, 149]. An important aspect of this online political news sphere, documented repeatedly, is that there is a strong partisan divide on what news sources people typically rely on and trust [5, 96, 99, 106, 145]. People tend to habitually consume the news source of their choice based on political beliefs [96, 106, 108, 131, 145] and engage in discussion within circles of similar political affinity [17, 58, 86, 99]. Researchers in Computer-Supported Cooperative Work (CSCW) and Human Computer Interaction (HCI) have repeatedly shown that users engage with more politically homogenous communities along partisan lines [4, 17, 34, 58, 73, 149]. Users curate their social media feeds that are reflective of their political preferences [50, 91, 106] and change their settings to see fewer posts from individuals in their network who hold differing political views [96, 145]. This is not surprising especially in light of the highly polarized political climate after the 2016 presidential election. Based on a Pew Research survey conducted right after the election, nearly two-thirds of Americans reported that online encounters with users on the opposite side of the political spectrum leave them feeling as if they have even less in common than they thought [44]. Furthermore, about 60% of American adults reported social media interactions with other users with opposing political views to be stressful and frustrating while a much smaller proportion (34%) found them interesting or informative [44].

News organizations are indeed aware of this polarized landscape and take it into consideration when trying to drive online traffic to their sites. For example, a group of news sites that target the politically far-left and far-right are both owned by the same company that often post their articles on social media [1, 131, 132]. Hence, even when reporting on the same fact or event, the company uses two different headlines with clear partisan slants [132] for their respective news sites even though the text of the articles below each headline are nearly identical [61]. As a result, coverage of the same set of events or facts are presented with radically different hyper-partisan undertones which can serve as qualitatively distinct fodder for political discourse among audiences depending on which news media they consume. Such practices can not only shift the focus of political discussion around identical events, but also isolate political discourse into echo-chambers [139] created by homogenous communities along partisan lines.

In our work we use the #MeToo movement as the topical lens to examine the nature of political discourse on SNS using the cases of three politically distinct (far-left, mainstream, and alt-right) news publishers on Facebook. Online social movements embody important social issues that generate momentum and political discourse [13, 24, 31, 53, 59, 69, 78]. The linguistic style and affect of online comments are known to influence how users evaluate and react to content on social media [7, 25, 27, 28, 38, 52, 71, 82, 147]. This implies that the way people talk about online social movements could in fact, influence their perception of related topics.

In this work we theorize that linguistic behavior not only shapes perceptions, but can also reflect and even influence various aspects of social identities, including political orientation. The goal of our work is to use NLP techniques and qualitative methods to discover how linguistic practices might differ among people with distinct political views and to understand how such practices may contribute to the formation of political identities and perspectives. Studying and contrasting language patterns of commenters from three politically distinct sites enable us to take a deeper
dive into the nature of the conversation that is happening in these public spheres across users who are polarized and potentially non-overlapping.

Our findings show that linguistic analysis tools are promising methods to inform current understanding of how people react to a political movement based on their commenting behavior. The nature of discourse in both the far-left and alt-right sites is structurally similar to each other in terms of their relationship to the mainstream. Through our work, we show that these structural similarities in commenting behavior might serve to reinforce polarization.

The main contribution of our work lies in the demonstration of how different components of linguistic behavior shape diverse political framings [23, 45, 90] of a social topic based on the political orientation of the commenter’s news source. This work contributes to developing a better understanding of the underlying linguistic factors that hinder as well as encourage healthy political conversation among users of politically diverse websites. We believe such factors are important in online political dialogue to not only understand ways to reduce incivility, but also to widen perspectives among audiences with opposing political viewpoints.

2 BACKGROUND

2.1 #MeToo Movement

On October 15, 2017, American actress Alyssa Milano took to Twitter to encourage women sexually harassed or assaulted in the past to use “MeToo” [107, 123, 159] – a phrase first used by Tarana Burke in this context in 2006 [48] – to demonstrate the pervasiveness of the problem. Within 24 hours, the hashtag was tweeted more than 500,000 times and used by more than 4.7 million Facebook users in approximately 12 million posts [123]. Since then, the hashtag spread across 85 countries in dozens of other languages with local variations of the phrase [107]. Millions of SNS users used “#MeToo” to share experiences of sexual assault [10, 60], prompting a national discussion of sexual misconduct at the workplace in many countries [29, 30, 107, 112, 157]. Events surrounding the #MeToo movement, however, garnered mixed reactions from the public with diverse media portrayals from various news sources across the political spectrum. While many praised the movement for serving as a platform for marginalized voices [29, 60, 157], others criticized the movement as a reflexive social media witch-hunt [72, 121, 158, 159] with an unclear agenda around who is and is not (e.g., sex workers and males) [30, 36] allowed to participate or use the hashtag. Garnering significant amounts of both support [11, 32, 122, 135] and backlash [14, 67, 120, 142, 143, 156], the online social movement generated both massive on- and offline political discourse on SNS [3, 33, 60, 155].

2.2 The Role of Hashtags in Online Social Movements & News Media Practices

Hashtags are powerful tools that focus attention, and in essence ‘brand’ an issue. As a locus of media attention and a shorthand phrase that serves to link numerous texts, hashtags can mobilize people around specific topics and dramatically amplify a message that goes ‘viral’. Hence, in the context of online social movements, hashtags can effectively operate as political framing tools [57, 64, 65], social markers of identity [63, 97, 117], and conversation facilitators [21, 69, 97, 105, 133, 154].

The power of hashtags is demonstrated through the intense momentum and political discourse surrounding #Blacklivesmatter, #Ferguson, and more recently with the #TimesUp and #MeToo
movements. Researchers have shown that hashtags can “rapidly elevate discourse beyond specific localities” [65] and raise the overall profile of its intended message as demonstrated in several studies [15, 24, 53, 118, 148]. News media companies also understand such possibilities that hashtags can afford. Hashtags that take off, not only increase audience reach, but also elicit strong emotional responses that play on the social and political identity of the reader based on where he or she may stand in regards to the online social movement [57, 69, 77, 118]. That said, it is not surprising that news media companies often use political hashtags in their headlines as a way to cater content to specific viewership profiles (based on political orientation, age, gender etc.). Therefore, selection of news source may in return, influence the nature of political discourse around online hashtag movements along partisan lines. This could possibly explain why, while reactions around online hashtag movements are often messy, they are known to be strongly divided by political orientation [24, 47, 69, 136]. That said, to our knowledge, our work is the first to examine from a linguistic perspective what and how people talk about the events and issues surrounding an online social movement in relation to the political orientation of the news source they consume on social media.

3 LINGUISTIC BEHAVIOR IN POLITICAL DISCOURSE

Understanding linguistic behavior in political commentary could shed light on factors that might reinforce partisan views. Scholars have shown that linguistic style is associated with interpersonal and social differences [22, 39, 56, 103, 110] (e.g. linguistic behavior differs by gender [16, 85, 126]). In fact, social identity is a product of linguistic practices [22], meaning that “identity is something that is constituted through social action, [especially] through language” [22]. According to sociolinguistic research, “identity emerges from the specific conditions of linguistic interaction” [22] rather than serving as a “pre-existing source of linguistic practice” [22]. Thus, linguistic behavior can potentially reflect deep differences among people involving identity and social differences.

Concerning political views, research has also shown that the way people talk about politics, the language they use, and the topics they focus on affect their evaluation and perception about politics [28, 54, 68, 79, 115, 116, 147, 151]. For example, swear words and vulgarity reduces the perceived persuasiveness of political messages [25]. When individuals receive political information intertwined with affective language, this language “creates a particular mood giving them another consideration that seems relevant” [147]. For example, negative affective language influences how individuals make decisions toward political subjects [146, 147]. Political information communicated with negative affective language persists longer in people’s memory compared to information delivered through neutral language [147]. This in turn can influence certain keywords to be mentioned more frequently compared to others, leading to opinion reinforcement [100, 101, 128] and thereby causing commenters to take more extreme positions [49, 62, 100, 101, 152]. The consistent appearance of important keywords in online discussion, in fact affects selective exposure [49, 91, 152], underpinning information that occurs most frequently as the dominant perspective [91, 152] while marginalizing perspectives that are less frequently mentioned or appear as less important [115, 116]. This can potentially reduce diversity of perspectives in political discourse, a key element that allows society to find common ground on important issues through democratic conversation [37, 115].

That said, differences revealed in linguistic patterns could serve as a building block for understanding polarization and entrenchment in political views expressed on social media. Based
on this premise, our work investigates how linguistic styles differ among commenters from politically distinct news sites, addressing the same topic embodied by an online social movement.

### 3.1 Research Questions

Based on the literature review, we expect the following linguistic attributes and patterns on social media to affect perception of content, especially for “hot-button” issues embodied by the #MeToo movement: 1) structural patterns of linguistic content, 2) semantic contexts, and 3) rhetorical patterns based on important keywords associated with discussion themes. Hence, to examine how these linguistic attributes and patterns might differ among people who produce and consume different political information, we ask the following research questions. All research questions are examined in the context of the #MeToo movement.

#### 3.1.1 RQ1: What linguistic and affective attributes characterize commenting behavior across the three politically distinct news sources covering #MeToo?

Given that the linguistic style of online comments is known to affect how users perceive content, this research question allows us to uncover distinct structural patterns in linguistic behavior across the three publishing sites. Such patterns may underlie factors inducing further polarization of perspectives among commenters.

As discussed above, negative affective language has been shown to have considerably stronger effects on people’s attitudes than positive information, especially when the content is political [95, 113]. Hence, the use of negative affective language, (e.g., anger) can have strong persuasive influences [146, 147] in formation of political opinion. Furthermore, the use of socially offensive language such as swear words or sexually explicit expressions can elicit affective reactions which can lead to more entrenched opinions. Therefore, this research question will identify whether differences exist among commenting behaviors in the three politically distinct sites, whether there are strong trends around affective language patterns, and how these differences might be associated with various viewpoints.

#### 3.1.2 RQ2: What are the differences in the semantic contexts in which #MeToo is framed in the commenting discussion across the three politically distinct news sources?

Understanding how a word is characterized by its nearby words can reveal key linguistic contexts in which it is discussed [46]. People of different political orientations may in fact discuss the same issue with very different contexts. Context, in turn, can influence how arguments are understood. For example, the semantic context through which a topic is conversed has been shown to predict the norms of online group discussion among different weight loss communities on Reddit [26]. Hence, with this research question we investigate the latent contextual cues of words associated with the token, “MeToo” across the three publishers by examining the semantic closeness of nearby words.

#### 3.1.3 RQ3: What kind of rhetorical patterns are observed from the discussion of the most important keywords across commenters from the three politically distinct news sources?

The types of rhetorical engagement [83] that characterize political conversations can potentially reveal how commenters understand and make sense of an issue. It can also inform us of the persuasive tactics used on different sides. This research question is comprised of two parts. First, we identify the tokens that consistently appear in the commenting discussion among the three political news sites. Then we use the token results as anchors from which to explore rhetorical patterns surrounding these important keywords to help understand how the three political commenter groups frame issues and construct arguments.
4 METHODS

4.1 Data Source

After a careful review of approximately 100 news sources from mediabiasfactcheck.com as well as their social media presence on Facebook, we selected three contrasting sites for our data sources that represent three distinct political viewpoints: Democracy Now (DemNow) which represents a far-left viewpoint, the New York Times (NYT) which represents a mainstream viewpoint, and Breitbart which represents a right-wing viewpoint (referred to in popular culture as the alt-right)\(^2\). All three news publishers consistently posted their articles on Facebook. As criteria for our selection, we used the number of Facebook page followers\(^3\), likes\(^4\), and the overall number of #MeToo articles as well as the number of Facebook comments on these articles.

4.2 Data Collection

To gather data for this study, we first selected all article posts that included the phrase “#MeToo” either in the article headline or in the text portion of the Facebook post during the period between October 2017 (when Alyssa Milano first shared her tweet using #MeToo) and March 2018 from all three publishers’ Facebook pages. We used the Facebook API to collect all user comments and replies from the selected article posts\(^5\). Unfortunately, due to Facebook’s API restriction, we were unable to collect anonymized commenter ids\(^6\), which would have provided us valuable information on the unique number of commenters.

Table 1. Descriptive Statistics: Number of comments (per article) & comment length from #MeToo coverage posts shared by Democracy Now, New York Times, and Breitbart Facebook pages

<table>
<thead>
<tr>
<th></th>
<th>Democracy Now (21 articles)</th>
<th>NYT (62 articles)</th>
<th>Breitbart (10 articles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of original</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facebook comments per</td>
<td>35 (67)</td>
<td>158 (282)</td>
<td>927 (1,082)</td>
</tr>
<tr>
<td>article (total comments</td>
<td>30 (49)</td>
<td>56 (128)</td>
<td>553 (669)</td>
</tr>
<tr>
<td>in parentheses)</td>
<td>8 (16)</td>
<td>4 (4)</td>
<td>297 (426)</td>
</tr>
<tr>
<td>max</td>
<td>112 (220)</td>
<td>2,003 (2,868)</td>
<td>3,306 (3,455)</td>
</tr>
<tr>
<td>total</td>
<td>743 (1,409)</td>
<td>9,811 (17,491)</td>
<td>9,267 (10,821)</td>
</tr>
<tr>
<td>Number of Facebook</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>replies per article</td>
<td>32</td>
<td>124</td>
<td>155</td>
</tr>
<tr>
<td>mean</td>
<td>20</td>
<td>74</td>
<td>140</td>
</tr>
<tr>
<td>median</td>
<td>5</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td>min</td>
<td>108</td>
<td>865</td>
<td>326</td>
</tr>
<tr>
<td>max</td>
<td>666</td>
<td>7,680</td>
<td>1,554</td>
</tr>
<tr>
<td>Length of Facebook</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>comments (in number of</td>
<td>1st quartile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>characters)</td>
<td>35</td>
<td>35</td>
<td>25</td>
</tr>
<tr>
<td>mean</td>
<td>146</td>
<td>139</td>
<td>85</td>
</tr>
<tr>
<td>median</td>
<td>80</td>
<td>80</td>
<td>52</td>
</tr>
<tr>
<td>3rd quartile</td>
<td>160</td>
<td>170</td>
<td>100</td>
</tr>
<tr>
<td>max</td>
<td>4,972</td>
<td>4,925</td>
<td>7,904</td>
</tr>
</tbody>
</table>

\(^2\) DemNow and Breitbart are a close mirroring of each other in terms of political extremity \[87, 88\]. We take into consideration that NYT is considered left-center \[89\].

\(^3\) Number of Facebook page followers at the time of data collection: DemNow=1.2M, NYT=15.1M, Breitbart=3.8M

\(^4\) Number of Facebook page likes at the time of data collection: DemNow=1.2M, NYT=15.5M, Breitbart=3.9M

\(^5\) Comments were not curated by the publishers and were publicly available on Facebook through the publishers’ Facebook page.

\(^6\) Only the owners of the publisher’s Facebook pages have API access to anonymized versions of the commenters’ user ids.
During this six-month period, there were a total of 21 published article posts from DemNow, 10 from Breitbart, and 62 from NYT Facebook pages. Such difference in the number of articles is a potential limitation, which we discuss further in later sections. For our analyses, we used both original comments and replies to these comments from all the #MeToo articles posted on the three publishers’ Facebook pages, giving us a total of 17,491 user comments from NYT, 10,821 user comments from Breitbart, and 1,409 user comments from DemNow posts as shown in Table 1. While the number of average comments per article was highest in Breitbart news posts, the average comment length in character count was the lowest compared to those of DemNow and NYT.

4.3 Analysis

4.3.1 Linguistic Style & Affect (RQ1). In order to investigate the linguistic and affective traits characterizing commenting behavior, we used the 2015 Linguistic Inquiry and Word Count (LIWC). LIWC is a well-validated psycholinguistic lexicon [109] widely used by HCI and CSCW scholars to understand online discourse [20, 28, 39, 40]. We focused on six LIWC categories to measure affect (positive emotion, negative emotion, and anger) and linguistic style (swear words, informal language, and sexual words) based on prior findings on linguistic behavior in political discourse as discussed in the literature review. We added the sexual word category given the topical nature of the data as well as the fact that a large proportion of the swear words in the data were related to sex.

4.3.2 Semantic Proximity to “MeToo” (RQ2). Unlike traditional topic modeling, word embeddings, implemented through neural network architecture such as word2vec [93, 94] and GloVe [111] reveal latent contextual cues of tokens by capturing the co-occurrence of terms with an associated word. Hence, in order to answer RQ2, we used word2vec to quantify the semantic proximity using cosine similarity distance between the token “MeToo” and nearby words that co-occur with it. Using this unsupervised learning algorithm, we built embedding models for each publisher with a minimum count of 50 for all words based on a conservative window of 12 words. Our total vocabulary size was 6,302, 23,655, and 12,507 unique tokens for DemNow, NYT, and Breitbart, respectively. Before vectorizing our tokens, we lowercased all text and removed stop words (e.g., functional words, such as “the”, “is”, “at”, etc.). We also customized our stop words to ensure that words prefaced with the hashtag symbol, ‘#’ were not removed given the nature of our data.

4.3.3 Rhetorical Patterns in Discussion of Important Keywords (RQ3). First, to investigate the important keywords frequently discussed among commenters across the three different news sites, we used term frequency-inverse document frequency (TF-IDF) analysis. TF-IDF is a ranking function widely used in information retrieval and text-mining to investigate how important a word is within a corpus of text data [134, 153]. The importance of a word, or its TF-IDF weight increases in proportion to the number of times the word appears in a document (e.g. one comment from NYT), but is offset by the frequency of the word in the entire corpus (e.g. all comments from NYT articles). To examine the important keywords discussed in the alternative news sites (Breitbart and Democracy Now) in relation to those that appear in the more mainstream news source (New York Times), we conducted two TF-IDF analyses – first on the

---

7 As most comments contained 12-14 words on average ($\bar{x}$ =12 for Breitbart, $\bar{x}$ =13 for DemNow, and $\bar{x}$ =14 for NYT)
combined corpus of Breitbart and NYT comments and second, on the corpus of DemNow and NYT comments.

Next, we used discourse analysis [51] to analyze all comments that included the top tokens with the highest TF-IDF weights to explore the rhetorical patterns used by commenters in their discussion of important keywords. Discourse analysis is a form of qualitative analysis that involves identifying patterns, relationships, and values in text data [51], and has been used in several HCI and CSCW works examining online comments [114, 150]. First, we thematically grouped all 1302 comments with the top TF-IDF tokens (“sex”, “color”, “experience” “flirting”) and used axial coding [138] to delve into each comment in greater depth. We performed discourse analysis on the comments to identify consistent patterns of rhetoric across commenters. All three authors iteratively engaged in thorough discussion and interpretation of comments throughout the analysis.

5 RESULTS

Our findings show strong differences in linguistic style and affect across the comments from the three politically distinct news publishers. Commenters from both the far-left and alt-right sites use a significantly greater proportion of negative affective words and informal speech compared to those who engage in discussion on mainstream news media (RQ1). Furthermore, the top tokens with the highest semantic proximity to “MeToo” are thematically different across comments from DemNow, NYT, and Breitbart, suggesting that the framing of discussion around the MeToo hashtag is different based on the political orientation of the news source (RQ2).

Partisan identity is also reflected in the important keywords commenters discuss based on our TF-IDF results (RQ3). By taking a closer examination of how the top tokens with the highest TF-IDF weights are discussed among commenters, we examine distinct patterns of rhetoric and discourse across the three groups in our data. We discuss our results in greater detail below.

5.1 RQ1: What linguistic and affective attributes characterize commenting behavior across the three politically distinct news sources covering #MeToo?

Table 2 summarizes the six linguistic measures of affect and style on user comments on DemNow, NYT and Breitbart #MeToo article posts shared by the three publishers’ Facebook pages. In the summary table, we show the mean proportion of affective words considered swear, informal, or sexual language. Given that the LIWC values for each comment across the three news sources were not normally distributed, we used the Kruskal-Wallis test, a non-parametric alternative of analysis of variance (ANOVA), to examine whether the differences in the proportion of LIWC words were statistically significant across the three groups. The test indicated that the proportion of words corresponding to each of the six LIWC categories was indeed significantly different among the three corpora of comments from each publisher.

To understand the differences in more depth, we conducted Wilcoxon signed rank pairwise tests for the three pairs of groups across the six LIWC categories. As shown in Table 3, pairwise tests indicated that the proportion of lexical content based on affect and linguistic style were significantly different (p-value <0.001, adjusted for a Bonferroni correction) for certain pairs. We describe the differences in more detail in the following sections.
5.1.1 Affect. There is a significantly higher proportion of both positive and negative emotional words across Breitbart comments compared to DemNow and NYT comments. However, there was no significant difference in the proportion of affective words between DemNow and NYT comments. Breitbart comments also have the highest, and NYT comments the lowest proportion of anger-related words among the three groups.

5.1.2 Linguistic Style. Comments from the three publishers also differed in terms of linguistic style. There was a significantly higher proportion of sexual words across Breitbart comments compared to those of DemNow and NYT comments. Breitbart comments contained nearly three times the proportion of sexual words compared to NYT comments. DemNow comments contained the second highest proportion of sexual words, with nearly twice the proportion found in NYT comments. Furthermore, Facebook comments from Breitbart posts contained a significantly greater proportion of swear words and informal speech compared to those written on NYT article posts. In fact, the average proportion of swear words used in Breitbart comments is five times higher, and the use of informal speech is nearly twice as greater than those used in NYT comments. By contrast, the difference in the proportion of swear words and informal speech between Breitbart and DemNow was not significant. Overall, NYT comments contained the least proportion of swear, informal, and sexual words among the three groups of comments.
Thus, in examining the linguistic attributes of comments in RQ1, we discovered that the comments from far-left and alt-right sites exhibit greater use of informal language, profanity, and words related to anger & sex, compared to those from the mainstream publisher. What is interesting is that there are no differences in the proportions of swear words and informal speech between the far-left and alt-right sites. However, comments from the Breitbart site (the alt-right) are the angriest, most emotional (positive and negative) and sexual compared to the comments from the other two publishers.

5.2 RQ2: What are the differences in the semantic contexts in which #MeToo is framed across discussion from the three politically distinct news sources?

In this section, we present results from the word embedding analyses using word2vec to analyze semantic context. Our findings show that the top words most semantically associated with “MeToo” is noticeably different across the three news sources. In Table 4, we show the 25 tokens semantically closest to the word “MeToo” based on cosine similarity values from the three embedding models.

<table>
<thead>
<tr>
<th>word</th>
<th>Democracy Now cosine similarity</th>
<th>word</th>
<th>New York Times cosine similarity</th>
<th>word</th>
<th>Breitbart cosine similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>metoo</td>
<td>1.000</td>
<td>metoo</td>
<td>1.000</td>
<td>metoo</td>
<td>1.000</td>
</tr>
<tr>
<td>fame</td>
<td>0.887</td>
<td>hashtag</td>
<td>0.906</td>
<td>whore</td>
<td>0.942</td>
</tr>
<tr>
<td>movements</td>
<td>0.874</td>
<td>group</td>
<td>0.838</td>
<td>always</td>
<td>0.940</td>
</tr>
<tr>
<td>society</td>
<td>0.845</td>
<td>campaign</td>
<td>0.819</td>
<td>joke</td>
<td>0.937</td>
</tr>
<tr>
<td>money</td>
<td>0.834</td>
<td>silence</td>
<td>0.807</td>
<td>paid</td>
<td>0.934</td>
</tr>
<tr>
<td>ignored</td>
<td>0.798</td>
<td>courage</td>
<td>0.807</td>
<td>tell</td>
<td>0.933</td>
</tr>
<tr>
<td>predators</td>
<td>0.795</td>
<td>ones</td>
<td>0.798</td>
<td>stupid</td>
<td>0.925</td>
</tr>
<tr>
<td>famous</td>
<td>0.792</td>
<td>giving</td>
<td>0.795</td>
<td>must</td>
<td>0.925</td>
</tr>
<tr>
<td>hollywood</td>
<td>0.768</td>
<td>names</td>
<td>0.780</td>
<td>list</td>
<td>0.923</td>
</tr>
<tr>
<td>rich</td>
<td>0.767</td>
<td>takes</td>
<td>0.775</td>
<td>expect</td>
<td>0.922</td>
</tr>
<tr>
<td>standing</td>
<td>0.766</td>
<td>read</td>
<td>0.750</td>
<td>morals</td>
<td>0.917</td>
</tr>
<tr>
<td>thank</td>
<td>0.766</td>
<td>respect</td>
<td>0.743</td>
<td>day</td>
<td>0.917</td>
</tr>
<tr>
<td>times</td>
<td>0.758</td>
<td>past</td>
<td>0.739</td>
<td>boy</td>
<td>0.913</td>
</tr>
<tr>
<td>agenda</td>
<td>0.748</td>
<td>breakers</td>
<td>0.734</td>
<td>back</td>
<td>0.911</td>
</tr>
<tr>
<td>choice</td>
<td>0.742</td>
<td>share</td>
<td>0.724</td>
<td>movement</td>
<td>0.907</td>
</tr>
<tr>
<td>cultural</td>
<td>0.737</td>
<td>instead</td>
<td>0.717</td>
<td>makes</td>
<td>0.904</td>
</tr>
<tr>
<td>seriously</td>
<td>0.736</td>
<td>stories</td>
<td>0.712</td>
<td>ok</td>
<td>0.904</td>
</tr>
<tr>
<td>least</td>
<td>0.736</td>
<td>sharing</td>
<td>0.712</td>
<td>guys</td>
<td>0.903</td>
</tr>
<tr>
<td>profile</td>
<td>0.733</td>
<td>choice</td>
<td>0.699</td>
<td>office</td>
<td>0.901</td>
</tr>
<tr>
<td>assholes</td>
<td>0.732</td>
<td>name</td>
<td>0.696</td>
<td>actually</td>
<td>0.901</td>
</tr>
<tr>
<td>act</td>
<td>0.732</td>
<td>magazine</td>
<td>0.693</td>
<td>hypocrite</td>
<td>0.899</td>
</tr>
<tr>
<td>sorry</td>
<td>0.730</td>
<td>white</td>
<td>0.692</td>
<td>best</td>
<td>0.898</td>
</tr>
<tr>
<td>prostitution</td>
<td>0.728</td>
<td>help</td>
<td>0.685</td>
<td>anything</td>
<td>0.896</td>
</tr>
<tr>
<td>conversation</td>
<td>0.724</td>
<td>call</td>
<td>0.676</td>
<td>figure</td>
<td>0.896</td>
</tr>
<tr>
<td>knowing</td>
<td>0.718</td>
<td>brave</td>
<td>0.657</td>
<td>people</td>
<td>0.889</td>
</tr>
</tbody>
</table>

Each embedding model was separately built from the corpora of DemNow, NYT and Breitbart comments. The cosine similarity values listed in descending order in Table 4 indicate the
similarity of angles between each vectorized token and the word “MeToo”, ranging from 1 (identical) to -1 (absolute opposites).

The embedding results show thematic differences in the linguistic context of how #MeToo is described and understood among commenters from the three publishers. Among NYT comments, “MeToo” is semantically closest with tokens such as “hashtag”, “group”, and “campaign” – words that are somewhat intuitively characteristic of #MeToo as an online social movement or a hashtag campaign. Tokens such as “courage”, “respect”, and “brave” are also semantically associated with the word “MeToo”, suggestive of the emotive context in which NYT commenters describe content related to the hashtag. Action words such as “giving”, “take”, “read”, “share”, “help”, and “call” are also often used in relation to “MeToo” in the corpus of NYT comments. By contrast, the embedding tokens from Breitbart and DemNow comments that are semantically closest to “MeToo” are somewhat more thematically dispersed. However, it is noticeable that some of these top tokens, such as “whore”, “stupid”, and “hypocrite” from Breitbart and “assholes” from DemNow have clear negative connotations. Top DemNow tokens, such as “fame”, “money”, “famous” and “rich” and perhaps “hollywood” are also linguistically related to socioeconomic status.

To more deeply examine the differences in the semantic context of these tokens in relation to the keyword, “MeToo”, we present a discussion of selected quotes from the comments. We present examples focusing on the top tokens that reflect clear thematic patterns and nuances around how #MeToo is discussed among commenters based on its larger semantic framing across the three news sites.

5.2.1 New York Times – “hashtag”, “group”. Among NYT comments, the token “MeToo” is often referred to as a hashtag. As it is the case with online hashtag movements, the brevity of a hashtag phrase that goes viral is indeed effective, but can also be confusing in meaning, due to its pervasive use across social media users in multiple contexts that can change and evolve over time. Compared to Breitbart and DemNow commenters there is a greater number of NYT commenters who hint confusion, but also greater willingness to understand and engage in discussion as to what the hashtag represents. In one of the NYT article posts, one commenter, Jim, expressed confusion towards what the “MeToo” hashtag stood for:

“Not trying to start fights, but metoo is hard to understand...To my understanding, sexual harassment can be something as simple as a catcall on the street. So I guess my question is are all instances of misconduct, both verbal and physical a part of metoo, or is it women who have been assaulted and abused?”

In response to Jim, another commenter replied:

“Jim, the #metoo hashtag is intended to capture all of those instances. However, I think you will find that stories like the ones these women are telling are not as uncommon as most people would like to believe...When I finally started talking to other women about my rape, I was shocked and dismayed to find how many of them (my friends, women I had known for years) had very similar stories to tell. #MeToo”

Here, the commenter discusses “MeToo” in the context of a hashtag by explaining the purpose of its use and the spectrum of what the hashtag encompasses (“to capture all of those instances” of what #MeToo stands for). The commenter also refers to the functional aspects of the hashtag by hinting its affordance to highlight the pervasiveness of an issue (“stories...these women are telling are not as uncommon”). Furthermore, the commenter also uses the MeToo hashtag as it
has been used by others by signing off her comment with “#MeToo” after mentioning her own personal experience (“my rape”).

The token “group” is the second closest word semantically associated with “MeToo” in the context of this NYT use-case as demonstrated in the following example. Throughout our analysis of NYT comments, there were numerous commenters, like the example from above, who used the comment threads below the article to briefly mention or share personal experiences of sexual harassment and assault with others. However, unlike personal Facebook posts where one can adjust the privacy settings to customize audience visibility, article posts shared by the publishers are all public, including the comment threads where any Facebook user can see or reply to what others wrote. Hence, in order to steer personal experiences into a less visible and more private space, some commenters invite others into “MeToo” groups. Thus, the context of the token “group” here refers to persons taking action related to the “MeToo” topic:

“#MeToo i have too many stories to write. I started a #MeToo group for all survivors.”

5.2.2 Breitbart – “joke”, “whore”, “always”. The context in which “MeToo” is discussed among Breitbart commenters strongly contrasts with the tone in which NYT commenters talk about the survivors or those associated with the movement. For example, the token “MeToo” is often referred to as a “joke”, the second highest embedding token from the model, as this quote makes clear:

"MeToo is a huge joke, so is Holly weird and mainstream media – no use for any of them."

In response to a Breitbart article, in which the sub-headline read, “The #MeToo movement sought to remove those in Hollywood who have been accused or found guilty of rape and sexual assault. Instead, they just gave one of them an Oscar”, one commenter responded:

"Always a pleasure to hear a lecture from someone, that literally bought and paid his way out of a rape charge, lecture others. #MeToo what a joke."

Most Breitbart coverage of the #MeToo movement shared through its Facebook page contained headlines and post texts with strong insinuations that seem to undermine the general purpose of the #MeToo movement. Unsurprisingly, this caters and feeds well to its Facebook audience who also commonly refer to #MeToo participants as “whores”:

"METOO has no meaning. Finding out they all are a bunch of hussie whores looking for cash. They all need to get out on the street corner and get their cash. Maybe 10% have a legitimate complaint."

Furthermore, Breitbart commenters often use linguistic absolutes, such as “always” in their comments:

"Because a lot of #metoo is about fake idiot liars looking for profit and control, hence it’s always the ‘powerful’ males they name... Naming little schleppy nobody will get them nothing. But when it’s profitable to suck up to them or on them they are still willing to do that too. Whatever happened...she was a profit seeking slutty little whore. Can’t believe he fell for it.”

Here, the commenter generalizes how it is “always” the powerful that #MeToo participants accuse rather than those who are less well-known. This shifts the context of the #MeToo movement from a campaign of sharing experiences to one purposed to name accusers.

5.2.3 Democracy Now – “fame”, “money”, “rich”. On DemNow comment threads, “MeToo” is often semantically associated with words like “fame”, “money”, and “rich” where commenters often discuss the #MeToo movement in relation to race and socioeconomic status:

"I was more than insulted listening online this morning to Oprah compare the violent and brutal pain of rapes and even murders that Black women endured by racist white men during Jim Crow
to that of rich white women in Hollywood and business. The brutal gang rape of Recy Taylor by six white men in Alabama is not comparable to the alleged sexual assaults that rich white women (often times purposely endured for fame and money) are fighting in their new #TimesUp and #MeToo movements. To compare the savagery and racism that fueled the many rapes and abuse that Black women had to endure by the hands of racist white men to that of rich white women’s new fight for dominance and power is a shameful erasure, even for Oprah. These rich women weren’t raped, beaten, bloodied and left to die because of hate and white power, these rich women chose silence out of fear of their careers and wealth, Black women who chose silence during Jim Crow etc., chose silence out of fear for their very lives and that of their families.”

Here, the commenter takes a strong partial stance by using language that villainizes the experiences of one group (rich white women) over those of others (Black women). The commenter makes generalized assumptions on race and socioeconomic identities, claiming that “alleged sexual assaults that rich white women” experience are often “purposely endured for fame and money”. This kind of discourse with strong partial references to race and socioeconomic identities is heavily echoed throughout DemNow comments as it will be further shown in the findings.

In summary, the top tokens that are semantically associated with “MeToo” show noticeable thematic differences. This suggests that the framing of discussion around #MeToo is different as shown through the different words that are semantically closest to the hashtag. That said, commenters can understandably react to #MeToo or related-topics in different ways depending on how the hashtag is presented across the three publishers. In other words, potential different framing biases [42, 43, 75] around the MeToo hashtag can elicit different cognitive prejudices related to the movement, manifest in the discussion among commenters of these politically distinct news sites.

5.3 RQ3: What kind of rhetorical patterns are observed from the discussion of the most important keywords across the three politically distinct news sources?

To analyze rhetorical patterns based on important keywords, we first needed to identify the key tokens discussed among commenters from DemNow, NYT, and Breitbart Facebook article posts on #MeToo. We conducted two TF-IDF analyses – first on the combined corpus of Breitbart and NYT comments and second, on the corpus of DemNow and NYT comments. Our TF-IDF results show that commenters from DemNow and Breitbart tend to focus on sexual subject matters while NYT commenters are more topically focused on the nuance of events related to the online movement as well as experiences shared through the #MeToo movement.

5.3.1 Important keywords across the three politically distinct news sites. In Figures 3 and 4 we show the top 40 linguistic tokens in descending order of their TF-IDF weights from Breitbart and NYT, and from DemNow and NYT analyses, respectively. We first describe the keyword results before presenting findings from the discourse analysis.

Comments from Breitbart & DemNow are topically focused on sex. Of the top 40 words with the highest TF-IDF weights for Breitbart in Figure 3, nearly three-fourths of the tokens are either swear- or sexual words, consistent with the results from the LIWC analysis. Similar to the list of Breitbart tokens, almost half of the top 40 DemNow tokens shown in Figure 4 are also either expletives or words related to sexual profanity. In fact, the 12 overlapping tokens between Breitbart and DemNow TF-IDF are: “sex”, “whores”, “ass”, “pussy”, “bullshit”, “fuck”, “fucking”, “asshole”, “assholes”, “dick”, “bitch, and “penis”. Tokens from Breitbart comments contain
degenerative expressions such as “libtard”, “douche”, “prick”, “jackass” as well as those more specific to women (“whores”, “slut”, “bitches”, “prostitute”, “skank”). Furthermore, Breitbart and DemNow tokens suggest that comments from both the alt-right and far-left coverage on #MeToo tend to be topically focused on aspects related to sexual profanity based on the frequent references to both male and female body parts (“ass”, “asshole”, “dick”, “penis”, “balls”, “vaginas”).

![Graph of Top 40 Tokens in TF-IDF Weights](image)

**Fig. 3. Top 40 tokens in descending order of their TF-IDF weights based on a TF-IDF analysis of Breitbart & New York Times comments on #MeToo articles.**

**Comments from NYT focus on experience & nuance.** By contrast, NYT tokens are more semantically nuanced in reference to sex: “flirting”, “advances”, “dating”, “relationships”, “touched”, “knee” as shown in Figure 4. In fact, each list of the top 40 NYT tokens from both TF-IDF results do not contain any profanity or sexual references to women. Instead, the linguistic tokens are suggestive of aspects related to people’s experience, context, or discourse around the topic of #MeToo as well as the various perspectives and quality of perspectives surrounding issues related to the topic as shown in Table 5.

While all articles from the three publishers are topically focused on events surrounding the #MeToo movement, the TF-IDF results clearly resonate strong differences in important keywords discussed among commenters. This suggests that it may be challenging for political discourse surrounding important social topics (as embodied by the #MeToo movement) to topically...
converge among those who consume news sources that are highly distinct in political orientation. In the next section, we present the context in which the top TF-IDF tokens are discussed by the commenters in greater depth.

Fig. 4. Top 40 tokens in descending order of their TF-IDF weights based on a TF-IDF analysis of Democracy Now & New York Times comments on #MeToo articles.

Table 5. Categorization of NYT tokens from the two TF-IDF results (mutual tokens are bolded).

<table>
<thead>
<tr>
<th>Experience</th>
<th>Context</th>
<th>Discourse</th>
<th>Perspectives</th>
<th>Quality of Perspectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;experiences&quot;</td>
<td>&quot;context&quot;</td>
<td>&quot;discussed&quot;</td>
<td>&quot;norms&quot;</td>
<td>&quot;societal&quot;</td>
</tr>
<tr>
<td>&quot;survivor&quot;</td>
<td>&quot;nuance&quot;</td>
<td>&quot;justify&quot;</td>
<td>&quot;attitudes&quot;</td>
<td>&quot;generational&quot;</td>
</tr>
<tr>
<td>&quot;occurred&quot;</td>
<td>&quot;language&quot;</td>
<td>&quot;suggesting&quot;</td>
<td>&quot;views&quot;</td>
<td>&quot;pervasive&quot;</td>
</tr>
<tr>
<td>&quot;degrees&quot;</td>
<td>&quot;expression&quot;</td>
<td>&quot;assumption&quot;</td>
<td>&quot;mentalit&quot;</td>
<td>&quot;internalized&quot;</td>
</tr>
<tr>
<td>&quot;spectrum&quot;</td>
<td>&quot;judgement&quot;</td>
<td>&quot;convinc&quot;</td>
<td>&quot;recognized&quot;/&quot;recognize&quot;</td>
<td>&quot;compassion&quot;</td>
</tr>
</tbody>
</table>

5.3.2 Rhetorical patterns using important keywords. To investigate the various rhetorical patterns observed in discussion among commenters, we draw examples from the data in which the top TF-IDF tokens are expressed in the comments. Findings from our discourse analyses, as
shown in the selected examples from our data, demonstrate similarities in rhetorical style between DemNow and Breitbart commenters. Commenters on the far-left and alt-right sites both exhibit discourse behavior that subverts a particular social group and fragments solidarity of the movement based on racial and socioeconomic statuses.

**Breitbart comments tend to dehumanize one group.** Among the top 40 Breitbart tokens, the word “sex” has the highest TF-IDF weight. Breitbart commenters use the token most commonly to describe how the #MeToo movement is about women using sex to advance their careers:

“Isn’t this how the whole #MeToo movement got started? Have sex in exchange for what you want.”

The token “sex” is often used when #MeToo participants are framed in the context of participating in or encouraging their own sexual abuse, or what the commenter below would describe as engaging in “prostitution” to “further their career”:

“I have ZERO respect for women who participated in the sexual abuse in exchange to further their career and now are crying me too. That is not sex abuse, that is called prostitution.”

Breitbart commenters also often bring up the notion of whether #MeToo experiences truly entailed “forcible sex”, often defining those that can be considered legitimate “MeToo” incidents as “rape” or life-threatening situations where victims would have had to understandably choose forced sex in order to have saved their lives at knife- or gunpoint:

“So you sleep around...and then 20 or 30 years later decide to say you were harassed or made to have sex. Did any claim rape? Did they have a gun or knife held to their throats? How does one claim they were forced to have sex? I'd just like to know.”

In these examples, it is evident that Breitbart commenters are bringing in strong prior beliefs about the legitimacy of victim status in sexual encounters. These data reveal a prevailing attitude that #MeToo stories are illegitimate statements of harassment. Further, there is a sense that commenters see themselves as exposing hypocrisy in the movement by suggesting those who have come forward are weak, manipulative, and untrustworthy individuals who are, across the board, using harassment complaints for personal gain. These commenters mostly criticize the #MeToo movement and its participants for encouraging women to fabricate their narratives of sexual abuse. The criticisms are usually accompanied by the commenters’ own assumptions around what they consider legitimate sexual abuse or harassment.

**DemNow comments tend to promote social fragmentation based on race.** Among the top 40 DemNow tokens, the word “color” has the highest TF-IDF weight. “Color” is an important keyword frequently mentioned among commenters who often argue over which race is excluded or included in the #MeToo narrative:

“That's true! It's largely a bunch of fed up white women in #MeToo.”

One commenter echoes agreement, arguing the #MeToo conversation is primarily focused on those who are racially and socioeconomically privileged:

“Yes...women of color and working-class women are especially vulnerable, but the #MeToo conversation is primarily with regard to upper class white women.”

Here, the commenter may be making a valid point on the importance of recognizing intersectional identities. However, the comment also serves to ‘other’ those who have come forth by suggesting that some women are garnering attention based on race and socioeconomic status.

In response to a DemNow interview article on Tarana Burke, one Facebook commenter criticizes the #MeToo founder for “making it about racism”:
"OMG, Stop already. Nobody is being excluded. Since when did using the word "women" mean just white women or whatever perceived persecution she sees? I thought we were talking about ALL women...Did someone go out there and say, 'this conversation is strictly limited to white women'? She's CHANGING the discourse now and making it about racism and exclusion of people of color because SHE'S doing that...The Time cover I'm looking at has a black woman squarely in the FRONT of the group...So where is the exclusion?"

Ironically, the DemNow commenter criticizes the focus on race by using race as rhetoric: “The Time cover...has a black woman squarely in the FRONT...So where is the exclusion?”. The comment exhibits a form of discourse that emphasizes social fragmentation based on race, stripping away empathy towards understanding others’ experiences. While not blaming the victim based on the same criteria as those posting comments on Breitbart stories, such generalizing and othering emerges as a striking rhetorical similarity in how each of the polarizing venues differs from the mainstream.

**NYT comments encourage perspective-taking, informing, & educating others.** The word with the highest TF-IDF weight in the list of NYT tokens from the NYT-Breitbart TF-IDF result was “experiences”. While Breitbart comments largely expressed judgments about others who had come forth with #MeToo stories, NYT comments often contain descriptions of the commenters’ own experiences. Taking the position of one’s own experience, these comments are written in style of personal disclosure and empathy. Such comments are framed to inform, encouraging others to take victims seriously and withhold generalized perspectives. An NYT article headlined, “Catherine Deneuve and Others Denounce the #MeToo Movement” was posted by the NYT Facebook page with the following Facebook post text: “Catherine Deneuve and others disapprove of #MeToo for punishing men when ‘the only thing they did wrong was touching a knee, trying to steal a kiss, or speaking about ‘intimate’ things at a work dinner or sending messages with sexual connotations.” In response, one commenter wrote:

> “It’s not just a knee touch or sexual comments made in the work place! It’s the entire experience of discomfort or guilt that you may have unknowingly somehow led that male on. It happened to me. My coworker and I became close friends, but he took it further, thinking I was flirting and hinting I would welcome his advances. I spent MONTHS being uncomfortable in a job that was my dream job otherwise. And when I finally DID address it, I was let go and he got to stay, because he had tenure, and no one witnessed his behavior (we often worked alone in a lab together). It’s not a witch hunt. Each situation should be heard and judged independently, and NOT solely on your own harassment experiences-or lack of them.”

Here, the commenter shares a personal experience in an attempt to provide context around why certain incidents that some might consider harmless can be an injuring experience with significant consequences for others. The focus of this comment highlights the fact that incidents of sexual harassment are more complicated than the description of concrete acts (“just a knee touch or sexual comments”) that have occurred. Rather than defining whether a certain act is and is not legitimate sexual harassment, the commenter here focuses on how the experience made her feel, encouraging others that “each situation should be heard and judged independently” rather than generalized.

The token “experiences” is also used in the context of informing and educating others on the topic of sexual abuse and harassment:

> “Keep in mind that many who have been assaulted began their experiences as children, or preteens/teens, setting them on a path to later engage in relationships that would further subject them to more assaults. The words "me, too" often do not equate to one single event of sexual
The discussion prompted by the articles encourages some NYT commenters to remind readers that the point of the #MeToo movement is about “Listening...to put aside your prejudices and just be present to others in pain” of their experiences. Some NYT commenters warn others not to “categorically deny other women’s experiences” or put the burden on #MeToo participants to “explain away their personal experiences in some broad manner”. As such, whether the token “experience” is used in the context to describe one’s own or others’, many NYT comments focus on informing and sharing the experience of sexual harassment and abuse rather than generalizing and demonizing those who have come forth.

From the TF-IDF analysis of DemNOW and NYT comments, the word “flirting” had the highest TF-IDF weight in the list of NYT tokens. Among NYT commenters, there was a lot of discussion over the nuances that separated flirting from sexual harassment. For example, one commenter makes a distinction by providing situational context to her perspective:

“Touching my knee, going in for a kiss, and flirting are NOT harassment or assault. They might be inappropriate/harassment depending on context (i.e. one person is a teacher and the other is a student). I don’t mind a man doing any of those things I listed, even if I’m not interested. The problem is if he continues when I’ve made it clear it’s unwanted attention.”

Another NYT commenter offers a discussion around how definitions need to be re-evaluated as they can evolve with time and encourages perspective-taking across generations:

“The world is changing. What was ‘flirting’ and ‘just having a little fun’ a decade or so ago is being re-evaluated. The definitions of sexual assault have changed - they’re tighter, now. Society evolves, and #MeToo has kindled a flare-up of awareness and of reaction. It isn't simply my parents' generation that has to rethink some aspects of casual behavior, it's mine, too.”

NYT commenters also express the need for perspective-taking across culture when talking about what is considered harmless flirting. In response to a lengthy discussion among several NYT commenters on a French actress’s infamous critique of the #MeToo movement, one commenter argues that the critique warrants a deeper cultural understanding even though she personally disagrees with the actress:

"American living in France here. I have come across many French women who feel this way, and it does not surprise me. I feel like it’s a very big part of the macho culture here. When I first moved here I was shocked by things said in commercials on regular TV channels during the middle of the day that kids can see, the manner men talk to women and the racist jokes that are just culturally acceptable. A lot of the older women and younger women I speak to are scared of losing this macho man culture, men open doors for then, whistling at them, [because some] enjoy it. Whereas others including myself, am not flattered by a cat call. But this is why it is up to us to speak up in the moment and say we are uncomfortable with what is happening. Some women want and enjoy using their femininity as power and more power to them! Others do not. And that does not make one wrong or right. As mentioned above by [previous commenter’s name] the difference is someone in power flirting/theft kisses/placing hands on knees and using this power over the person and it’s up to us the person receiving these gestures to openly question this behavior and speak up if we are not comfortable with it.”

Here, the commenter uses expressions such as “I feel like” and “I was shocked” in sharing her own experience of living in France as a way to provide other commenters, a better contextualization to the actress’s critique in the original article shared on the NYT Facebook page. In addition to sharing her own experience, the commenter emphasizes reflection and
acknowledgement of other people’s experiences, noting that the different perspectives people have therefore “does not make one wrong or right”.

In summary, findings from our discourse analysis demonstrate strong stylistic differences as well as similarities in rhetorical engagement based on the three politically distinct groups embodied by the DemNow, NYT, and Breitbart comments. Research has shown that online users tend to converge to one another’s communicative behavior [39]; in other words, commenters in the same discussion group are likely to parrot each other’s rhetorical manner in online conversation. This could in turn, amplify both positive (informing, educating, and perspective-taking) and negative (generalization, dehumanization, and social fragmentation) impacts as shown through our analysis. We discuss further implications in greater depth in the next section.

6 DISCUSSION

6.1 Affective Language & Heuristic Processing

Heuristic processing, which involves making quick judgments about the information one comes across rather than engaging in deep reflection or discussion with others before coming to a conclusion [127], is particularly salient to social media consumption [70, 92, 130]. Content related to politics is not an exception [91]. Researchers have shown that users often rely on cognitive shortcuts and heuristic cues to evaluate political comments on social media [74]. Moreover, research has shown that when people engage in heuristic processing of political information online, negative emotional words induce more negative conclusions on the topic of discussion [74, 91, 147]. Hence, a deluge of comments charged with negative emotions (negative affect, swear words, anger, sexual profanity) can create a strong negativity bias [6, 25, 137, 146, 147] towards the subject of discourse. Our work supports these findings while also suggesting that different venues appear to play a key role in the form of the emotional weight and rhetorical style in commenting behavior.

Research has also shown that when political information is communicated using negative emotional words, people are able to recall that information more easily [147]. In other words, using language laden with negative affect makes a longer-lasting impression on people’s memory, potentially allowing certain topics to be discussed more disproportionately often among commenters. This in return, can exacerbate the problems of political echo chambers on SNS [50, 139], biasing the nature of discussion around online news content among commenters.

Our analyses reveal that, compared to NYT commenters, Breitbart and DemNow commenters tend to use substantially more sexual profanity and curse words that are dehumanizing towards women. Studies have shown that dehumanizing language leads to negative emotional responses and attitudes towards the dehumanized group [146]. This prevents openness of understanding other people’s experiences – a key component of constructive democratic discourse [37].

6.2 Framing Effects and Generalized Perspectives

Framing effects [42, 75] have been a useful concept in understanding the complex facets around social movements [8, 9, 84], news coverage [84, 124, 125, 129] as well as the formation of political opinion [42, 90, 102]. A framing effect is a form of cognitive bias, in which people react to a given choice in different ways depending on how it is presented [42, 75]. Results from our embedding analyses showed noticeable differences in the top tokens most commonly associated with
“MeToo” across comments from the three publishers. Such differences imply potential framing effects [42, 75] around how #MeToo is discussed across commenters consuming news sources with clear differences in political orientation – at least in our three samples. Furthermore, discourse analysis and embedding results on Breitbart comments demonstrate that the token “MeToo” is semantically associated with absolutist expressions, such as “always”. Absolutist rhetoric is harmful for democratic discourse [12], especially on moral and political topics that entail a wide spectrum of opinions. Research has shown that generalizing expressions tend to brush off contrasting perspectives, arguments [80], as well as important facts [80, 81] without seriously engaging with them – elements characteristic of fake news dissemination [35, 66, 119]. Furthermore, Breitbart comments are often laden with sweeping assumptions about the definition of sexual harassment and the character of people who have come forth while DemNow commenters make blanket statements around which group deserves more sympathy in the #MeToo narrative. Generalized expressions promote dominant viewpoints while marginalizing minority perspectives and stifling discussion [12, 144], which discourages empathy and sharing [68, 80, 81, 144]. This may account for why there was a much smaller proportion of Breitbart and DemNow commenters disclosing personal stories of sexual harassment or #MeToo experiences compared to NYT commenters in the data.

6.3 In-Group and Out-Group Dynamics

Our analyses reveal strong in-group and out-group dynamics [115, 140, 141] among Breitbart and DemNow commenters. Breitbart comments portray #MeToo participants as “whores” and “sluts” who have manipulative sex to advance their careers. This form of out-group derogation [18, 76, 141], or the tendency to have negative views about people not part of one’s own group, villainizes all survivors of sexual harassment who are using the #MeToo movement as a platform to share personal stories.

DemNow commenters, on the other hand, exhibit strong in-group favoritism [76, 98, 141] based on racial and socioeconomic groupings. Commenters frequently argue over how women of color deserve more attention within the #MeToo movement or that rich, white women are unfairly hijacking the #MeToo narrative from black women. Such in-group/out-group biases in discussion could polarize the #MeToo movement and undermine the initial solidarity of those using the movement to speak out on difficult experiences.

6.4 Summary of Implications

The most surprising insight generated by these data was the structural and rhetorical alignment between commenters in our selected cases of the far-left and alt-right news sites in relation to the comments on more mainstream media. The crassness, emotional weight, generalizing, and othering of these comments were noticeably similar. Media polarization research suggests that there is likely little to no overlap in the population of individuals consuming the articles in these different media venues. However, distance from the mainstream appears to foster a particular interaction pattern in relation to current events. Across the board, the comments in both of these venues tended to be more absolutist and judgmental in striking contrast to the emotionally subdued, inclusive, personal, and empathetic comments that proliferated in the mainstream media. These data suggest that if online media has any chance of fostering democratic dialogue – a discourse in which viewpoints are shared reasonably, assumptions are challenged productively,
and personal experience is used to promote inclusivity – it would not happen at either end of the political spectrum. This is a hypothesis that can be explored with further research.

6.5 Limitations

Our study has several limitations. First, for our data source, we chose only three sites, which may not have been representative of the three disparate political views. However, we feel that these three cases provide a basis for future investigation, in which we will employ a larger sample of news sources so as to increase generalizability. Furthermore, the number of articles among the three publishers were different. Breitbart contained the fewest, and New York Times the highest, number of articles posted on Facebook. While this sheds interesting light in terms of publication behavior towards #MeToo coverage between the two politically distinct news media, such difference in the number of articles could have influenced the topical diversity of discussion among commenters. Another potential limitation is that the tone of the article could have influenced the tone of the comments. While analyzing the alignment of article and comment discourse is beyond the scope of the current paper, this can be addressed in future research.

Furthermore, reader demographics could explain the differences among the sites’ comments as education levels are likely to affect linguistic style. Based on traffic statistics provided by Alexa, a data analytics company owned by Amazon, readers who went to college and graduate schools are overrepresented on NYT and DemNow websites [41, 104]. For breitbart.com, the proportion of those who went to college among its audience are similar to the general Internet population while those who are graduate degree holders are underrepresented [19]. Furthermore, compared to the general online population, males are overrepresented at breitbart.com while underrepresented on nytimes.com and democracynow.com [19, 41, 104].

7 CONCLUSION

In this paper, we use linguistic tools to compare commenting behavior across three politically distinct news sources covering the #MeToo movement. Through our work, we show that commenting behavior reflects how topics surrounding an online social movement are framed and conversed about within a particular political orientation. Our work shows that linguistic style and affect, as well as rhetorical patterns, can shed light on the underlying factors that influence civil discourse on social media. Comments from the far-left and alt-right sites that we analyzed exhibited structural similarities in rhetorical and linguistic patterns that could promote polarizing viewpoints, while comments from the mainstream site we analyzed tended to encourage contextual understanding through empathetic discussion. While in our study we examined three sites of disparate political orientations, we feel that analyses of linguistic patterns can be applied more broadly to examine a range of different sites and topics to more deeply understand polarization and entrenchment of views.

REFERENCES


[55] Grieco, E. 2017. More Americans are turning to multiple social media sites for news.


[67] Kahn, C. 2018. Mexico’s #MeToo Faces Backlash After Celebrities Air Accusations Of Rape And Assault.


[78] Loza, S. Hashtag Feminism, #SolidarityIsForWhiteWomen, and the Other #FemFuture. Ada: A Journal of Gender, New Media, and Technology. 5. DOI: https://doi.org/10.7264/N337770V.
Linguistic Behavior in Comments of #MeToo Articles Across Political Perspectives


[107] Park, A. 2017. #MeToo reaches 85 countries with 1.7M tweets.


Linguistic Behavior in Comments of #MeToo Articles Across Political Perspectives


[155] 2017. #MeToo: Sexual harassment stories sweep social media after... Reuters.

Received April 2018; revised July 2018; accepted September 2018.