Conversational Inequality Through the Lens of Political Interruption

Clay H. Yoo[1] , **Jiachen Wang**[1] , **Yuxi Luo**[1] , **Kunal Khadilkar**[1] , **Ashiqur R. KhudaBukhsh**[2]
[1]Carnegie Mellon University
[2]Rochester Institute of Technology
{hyungony, jw7, yuxiluo}@andrew.cmu.edu, kkhadilk@alumni.cmu.edu, axkvse@rit.edu

Abstract

We present a novel dataset of dialogues containing interruption with an aim to conduct a large-scale analysis of interruption patterns of people from diverse backgrounds in terms of gender, race/ethnicity, occupation, and political orientation. Our dataset includes 625,409 dialogues containing interruptions found in 275,420 transcripts from CNN, Fox News, and MSNBC spanning between January 2000 and July 2021. From this large, unlabeled pool of interruptions, we release an annotated dataset consisting of 2,000 dialogues with fine-grained interruption labels. We use this dataset to train an interruption classifier and predict the interruption type of a given dialogue. Our results reveal that male speakers (in our collected samples) tend to talk more than female speakers, while female speakers interrupt more. Moreover, people tend to use less intrusive interruptions when talking to others sharing the same political belief. This pattern becomes more pronounced among news media with stronger political bias.

1 Introduction

Over the last few decades, political polarization within the USA has increased at an alarming rate [Iyengar *et al.*, 2012]. Multiple research disciplines have analyzed this growing political polarization [Poole and Rosenthal, 1984; Demszky *et al.*, 2019; KhudaBukhsh *et al.*, 2021] with several studies analyzing its manifestation in mainstream news media [Hyun and Moon, 2016; KhudaBukhsh *et al.*, 2021; KhudaBukhsh *et al.*, 2022]. In these paper, we focus on the three biggest US cable news networks in terms of viewership – Fox News, CNN, and MSNBC [1]. The reported partisanship notwithstanding [2], over the last two decades, these cable news networks have discussed a multitude of hot button issues through debates, panel discussions, and interviews with eminent politicians and political analysts.

[1]https://www.statista.com/statistics/373814/cable-news-network-viewership-usa/

[2]https://www.nytimes.com/2012/08/31/us/politics/msnbc-as-foxs-liberal-evil-twin.html

Kellyanne Conway (KC): He's not exactly a topic of hot, hot, hot topic of conversation.
[Dialogue 1 Starts]
Chris Cuomo (CC): --saying he was a plant, using him as an extension of a Deep State--
KC: No, no, he wasn't. We don't care about his wife. They were--
CC: --and saying the FBI and the DOJ--
KC: --oh, you mean the women who got the--
CC: --were against him.
[Dialogue 1 Ends]
KC: You mean his wife who got $500,000 from a - a Democratic PAC when she was running, think it failed, race for State Office? Nobody really cares about that.
CC: The President did.

Figure 1: A portion of conversation between Kellyane Conway and Chris Cuomo aired on CNN on February 19, 2019. We mark the start and end of an interruption dialogue.

While most of the prior work focuses on quantifying or evaluating this divide, in this paper, our work goes beyond the typical Republican/Democrat divide as we take a closer look at the nature of conversational inequality through the lens of interruptions and analyze how interruptions vary across race, gender, and political leanings of the conversationalists. As shown in Figure 1, an example interruption from our dataset, INTERRUPT, our dataset has the potential to offer critical insights into nuanced conversational inequality along multiple dimensions. Via a substantial corpus of 625,409 interruptions and an annotated data set (dubbed INTERRUPT) consisting of 2,000 instances introduced in this paper[3], we present a comprehensive analysis of interruption along multiple dimensions.

Our contributions are as follows:

• ***Resource****:* We release a dataset consisting of 625,409 interruption dialogues spoken between thousands of unique speakers from a wide range of gender, race/ethnicity, occupation, and political orientation. We provide fine-grained annotation of 2,000 dialogues. Our annotation scheme is grounded in extensive social science literature on interruption.

[3]Data and resources are available at the project page: https://github.com/anton-sturluson/IJCAI2022INTERRUPT

• ***Technical:*** We demonstrate that a robust interruption classifier can be trained using the manual labels and can reveal interruption patterns that are hard to be found from analyzing smaller number of transcripts.
• ***Social:*** We analyze the underlying dynamics of communication and interruptions, and its effects on race, gender, political affiliations.

2 Related Work

Sacks *et al.* [1978] proposed viewing a conversation as a turn-taking system, where speaker transition often recurs, one party talks at a time for most of the time, and occurrences of more than one speaker are common yet brief, among many rules involved. A deviation or violation from these turn-taking rules is often considered as interruption [Zimmermann and West, 1996; Beattie, 1982; Murata, 1994].

There exists substantial social science literature focused on studying gender and conversational interactions [Tannen, 1994]. Some of these studies present contrasting findings about interruptions. In what follows, we present a brief summary.

Interruption is often considered a sign of dominance and its use between speakers of different genders has been actively studied. Zimmermann and West [1996] showed the most pronounced result out of all and showed that most interruptions are initiated by male speakers. They studied transcripts of daily conversations recorded in coffee shops, drug stores, and other public places, where most of the participants shared similar backgrounds (white, middle-class, and from twenty to thirty-five years of age) with a wide range of relationships between speakers. Similar results were found among preschoolers aged between three and four years old [Aries, 1996; Esposito, 1979].

Another line of research reveals that interruption happens in the opposite direction, however. Kennedy and Camden [1983] recorded group conversations between graduate students and showed that female participants interrupted significantly more than male participants. When conversations between random pairs of 24 undergraduates aged between 18 and 23 with an even female/male ratio were recorded, female participants also showed to interrupt more [Nohara, 1992]. The literature indicates mixed results in showing interruption patterns between female and male participants [James and Clarke, 1993].

In addition to gender-specific analysis, interruptions have been studied through cultural and societal lens. Murata [1994] showed the cross-cultural dependence of interruption habit by analyzing the difference between native Japanese and English speakers. The study revealed that while Japanese speakers refuse to use intrusive interruptions when speaking with another Japanese speaker, they tend to use more interruptions when speaking with a native English speaker. Beattie [1982] related interruption styles of U.K. leading politicians, Margaret Thatcher and Jim Callaghan, to their public perception. The author explained that the public perceives Margaret Thatcher as domineering because she tends to win the battle for the floor when interrupted, not because she interrupts more. Grebelsky-Lichtman and Katz [2019] used verbal and non-verbal communication methods characterized as either feminine or masculine to study the conversation patterns of two previous US presidential candidates, Donald Trump and Hillary Clinton, during the 2016 presidential debates.

Not all interruptions are the same, and there have been attempts to formalize different types of interruptions. Beattie [1982] laid out a technical definition of interruptions, where a speech can be classified based on the success of attempted speaker switch, presence of simultaneous speech, and the first (inward) speaker finishing her utterance or not. Murata [1994] and Kennedy and Camden [1983] suggested more functional definitions of interruptions. Murata [1994] sub-divided interruption as co-operative and intrusive, whereas Kennedy and Camden [1983] categorized interruptions as clarification, agreement, disagreement, tangentialization, and subject change based on the intention of the speaker and how the interruptions are being used within the context. Generally speaking, overlap interruptions are widely considered as unintentional infringement happening near the transition-relevant places in conversations [Beattie, 1982; Murata, 1994; Zimmermann and West, 1996].

3 Dataset

Previous works relied on manual annotation of interruption in conversations using recorded audio and transcribed text, which is a time-consuming task hindering large-scale experiments [Anderson and Leaper, 1998]. As we present a dataset consisted of over 600K interruption dialogues aired on CNN, Fox News and MSNBC, our dataset holds promise to overcome this barrier and enable large-scale analysis of interruptions in political conversations. We manually label 2,000 dialogues into one of five interruptions and others - *Floor-taking Interruption*, *Overlap*, *Disagreement Interruption*, *Co-operative Interruption*, *Topic-changing Interruption* and *Others* - following the definitions laid out by Murata [1994]. We further group overlap and co-operative interruption as *benign interruption* and floor-taking/disagreement/topic-changing interruption as *intrusive interruption*. We provide an example dialogue in Figure 1, and the details of collection, processing and annotation of the dataset can be found in the project page. Toward the growing consensus of responsible use of data in AI, we follow Gebru *et al.* [2021] and present a datasheet for datasets.

To study interruption patterns arising from the distinct speakers involved, we further annotate 1,666 speakers according to their gender, race/ethnicity, occupation and political occupation. We selected most frequently appearing speakers from each news media to ensure the most coverage of dialogues. Each speaker is annotated into the following categories.

- Gender: Female, Male
- Race/ethnicity: White, Black/African American, Hispanic/Latino, Asian, American Indian/Alaskan Native, and Native Hawaiian/other Pacific Islanders (following U.S. Equal Employment Opportunity Commission[4])

[4]https://www.eeoc.gov/statistics/introduction-race-and-ethnic-hispanic-origin-data-census-2000-special-eeo-file

- Second race/ethnicity: Optional second race/ethnicity
- Occupation: Journalist, Politician, Political appointees, Politically-related Others, Others
- Political orientation: Democratic and Republican

• ***Political orientation***: We take a rather conservative approach in marking political orientation and only consider it as a ground truth if it shows in the "Personal details" or introduction section in Wikipedia[5]. If a speaker switched a party since 2000, we selected the party with the longest overlapping years from 2000. In this way, we identified 272 Republicans and 229 Democrats.

• ***Gender and race/ethnicity*** were decided based on place of birth, heritage, self-identified information, and photos. In particular, Wikipedia categories (e.g., "Hispanic and Latino American journalists") and ethnicity provided in Wiki[6] were used to confirm the speaker's race/ethnicity. For the individuals identified as more than one race/ethnicity, we marked both and included such speakers in both groups for analysis.

• ***Occupation***: The occupation was selected as a career since 2000 as this is the starting year of our dataset. For the individuals qualifying for multiple political occupations, we followed the ordering of politician $\rightarrow$ political appointee $\rightarrow$ politically related others. For example, if a person was both a politician and a political appointee (e.g., Mike Pence), we marked her as a politician. If a person was both a political appointee and politically related others (e.g., Kellyanne Conway), we marked her as a political appointee. If a person had a political occupation as well as a non-political one, we prioritized the political occupation. These rules were imposed for simplicity and clarity in the annotation.

Each speaker was annotated by at least two coders with disagreements resolved through adjudication. Despite our endeavor to prevent bias, our annotation may not align with the individual's preference or self-identification since it was done based on publicly available databases, such as Wikipedia, Wiki, Everipedia[7], and official websites of news media.

Definitions. We use the following terms in describing the dataset and experiments (Figure 1). *Line* refers to a block of an utterance spoken by each speaker until the speaker transition happens. *Interruption tokens* refer to four special marks located at the end of each line: "...", "--", "- -", and "-". The presence of interruption tokens indicates that the speaker associated with this line could not finish her speech. We define an *interruption dialogue* as contiguous lines containing interruption tokens at the end except for the last line, which ensures uninterrupted speaker transition at the end of each dialogue. We made this choice to leave more context when labeling the dialogues since interruptions often occur contiguously, i.e., after the outward speaker interrupts, the inward speaker shortly interrupts back. Throughout the paper, we use dialogue and interruption dialogue interchangeably.

4 Experimental Setup

We discard dialogues labeled as others since they lack context and/or can be classified as more than one label. Next, we use 80/10/10 split for train/validation/test datasets. While fixing the test set, we randomly shuffle and split train and validation sets over five seeds. To find the best performing models, we use an automatic hyperparameter-tuning framework Ray Tune [Liaw *et al.*, 2018].

For each (dialogue, label) tuple, dialogue is split into a sequence of lines. We use `[CLS]` token obtained from pretrained `BERT` [Devlin *et al.*, 2019] (`bert-base-uncased`) as line embeddings. For starting lines exceeding 512 token length, we truncate from the beginning since interruptions always happen at the end of the line. A sequence of line embeddings is then fed into a locked dropout layer [Gal and Ghahramani, 2016] and unidirectional GRU layer [Cho *et al.*, 2014]. The last hidden vector of the sequence is projected to a linear layer to output un-normalized probability distribution. Adam optimizer [Kingma and Ba, 2015] with the default hyperparameters in Hugging Face [Wolf *et al.*, 2020] and linear scheduler with warm-up steps are used for training. Because the size of our dataset is small, we kept hidden dimensions and number of layers small.

For each of the five seeds, we sample 20 configurations from the following list for hyperparameter tuning. All but learning rate were chosen uniformly; learning rate was chosen log-uniformly. Weighted F-1 score on the validation set is used for early stopping.

- GRU hidden dimension: $\{64, 128, 256\}$
- GRU number of layers: $\{1, 2\}$
- Batch size: $\{8, 16, 32\}$
- Warm-up steps: $\{0, 10\}$
- Learning rate: $[1e-5, 1e-3]$

5 Results

We train two models, one with fine-grained five labels ($\mathcal{M}_1$) and the other with binary labels ($\mathcal{M}_2$), and report their performance in Table 2. $\mathcal{M}_1$ shows an average weighted F-1 score of 74.7%, but we observe high bias and noise when predicting disagreement, topic-changing, and co-operative interruptions. $\mathcal{M}_2$ shows an average weighted F-1 score of 82.4% while the performance between benign (75.2%) and intrusive (86.1%) interruption is more balanced. Acknowledging that we are far from accomplishing competitive performance with the fine-grained labels, we use $\mathcal{M}_2$ for further analysis. Out of the five seeds we trained our models with, we choose the seed that yielded the highest test performance. Out of 625,409 interruption dialogues, 262,851 are labeled as benign and 362,558 are labeled as intrusive interruptions.

Figure 2 shows the monthly trend of benign and intrusive interruptions since January 2000. Monthly interruption rate is calculated as $r_k = \frac{|I_k|}{|L_k|}$, where I_k is the set of benign or intrusive interruption dialogues in k-th month and L_k is the set of lines from all transcripts in k-th month. A higher interruption rate suggests that there were more interruptions in that month for every line spoken by the participants. We also mark eight important political events with grey vertical lines. In most of the neighborhoods of these events, we can observe a local surge in interruptions. Most noticeably, the month before the 2016 US presidential election yields the highest intrusive interruption rate. Intrusive and benign interruptions

[5] https://www.wikipedia.org/
[6] https://www.wiki.ng/en/
[7] https://everipedia.org

Occupation	Gender	Race/Ethnicity	$r_{\text{benign}}/r_{\text{intrusive}}/r_{\text{all}}$ (%)	Avg.Words per Line	Count
Political Occupation	Female	Asian	24.4/21.4/45.8	55.0	3
		Black/African American	16.0/20.5/36.5	72.2	16
		Hispanic/Latino	23.4/28.3/51.7	51.4	3
		White	18.0/20.5/38.5	58.1	51
	Male	Asian	16.1/13.2/29.3	79.5	7
		Black/African American	13.9/21.4/35.3	66.1	26
		Hispanic/Latino	16.7/15.8/32.5	65.6	5
		White	12.7/20.5/33.2	61.3	252
Journalist	Female	Asian	18.2/23.2/41.4	90.0	20
		Black/African American	19.2/21.3/40.5	82.2	57
		Hispanic/Latino	18.3/20.4/38.7	86.1	10
		White	18.4/22.4/40.8	87.8	296
	Male	Asian	14.4/14.1/28.5	114.1	12
		Black/African American	19.4/18.2/37.6	83.5	51
		Hispanic/Latino	19.4/18.2/37.6	92.5	11
		White	16.6/21.4/38.0	88.7	402

Table 1: Average benign/intrusive/all interruption rate (in percentage), average number of words per line, and number of speakers (Count) in each group related to gender, race/ethnicity, and occupation. Political occupation combines politician, political appointee, and politically related others. Speakers appearing at least 10 times are included in the analysis.

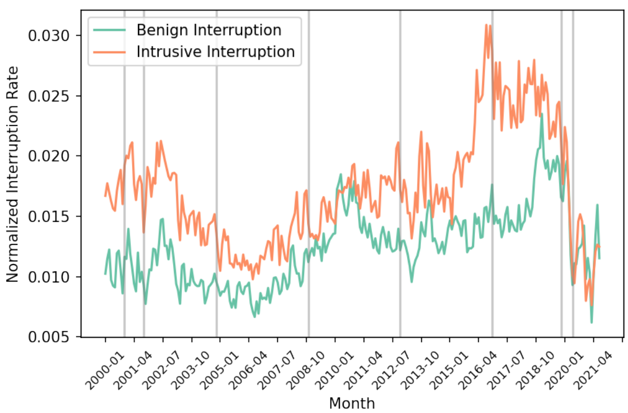

Figure 2: Monthly interruption rate per label between 01 January 2000, and 30 July 2021. The interruption rate is calculated by dividing the number of interruptions by the total number of lines in monthly transcripts. The grey vertical lines represent the following events from left to right order: 1) 2000 US presidential election; 2) 9/11; 3) the 2004 US presidential election; 4) the 2008 US presidential election; 5) the 2012 US presidential election; 6) the 2016 US presidential election; 7) the 2020 US presidential election; 8) the beginning of protests following George Floyd's murder.

Label	Precision	Recall	F-1
$\mathcal{M}_1$			
Floor-taking Interruption	$81.5_{1.0}$	$84.6_{3.4}$	$83.0_{2.1}$
Overlap	$75.8_{4.1}$	$78.7_{4.0}$	$77.2_{3.7}$
Disagreement Interruption	$59.7_{5.7}$	$40.0_{3.1}$	$47.8_{3.1}$
Co-operative Interruption	$53.3_{5.2}$	$50.5_{7.1}$	$51.8_{5.9}$
Topic-changing Interruption	$0.0_{0.0}$	$0.0_{0.0}$	$0.0_{0.0}$
Combined	$74.6_{2.1}$	$75.2_{2.4}$	$74.7_{2.2}$
$\mathcal{M}_2$			
Benign Interruption	$71.5_{2.1}$	$79.4_{4.4}$	$75.2_{2.7}$
Intrusive Interruption	$88.8_{2.1}$	$83.7_{1.6}$	$86.1_{1.3}$
Combined	$82.9_{1.9}$	$82.2_{1.7}$	$82.4_{1.7}$

Table 2: Mean and standard deviation (in subscript) of precision, recall and F-1 score of individual labels and the weighted score of the whole dataset using two types of labels.

show Pearson's correlation of $\rho = 0.72$.

To study the interruption style of individual speakers, we compute three interruption rates: benign interruption rate (r_{benign}), intrusive interruption rate ($r_{\text{intrusive}}$), and total interruption rate (r_{all}) for each speaker (s).

$$r_{\text{benign}}(s) = \frac{|\mathbb{1}\{\text{outward\_speaker = s, benign interruption}\}|}{|\mathbb{1}\{\text{outward\_speaker} = s\}|}$$

$$r_{\text{intrusive}}(s) = \frac{|\mathbb{1}\{\text{outward\_speaker = s, intrusive interruption}\}|}{|\mathbb{1}\{\text{outward\_speaker} = s\}|}$$

$$r_{\text{all}}(s) = r_{\text{benign}}(s) + r_{\text{intrusive}}(s)$$

$r_{\text{all}}(s)$ then becomes a measure of how much a speaker interrupts. $r_{\text{all}}(s) = 0.5$ indicates the speaker interrupts and gets interrupted an equal number of times, 0 indicates the speaker only gets interrupted while never interrupting others, and 1 indicates the speaker always interrupts without getting interrupted. We compute the average interruption rates of grouped speakers according to their occupation, gender, and race/ethnicity and report the statistics in Table 1. Speakers with others as their occupation are excluded from the analysis, and speakers appearing in at least ten dialogues are used to compute the statistics. Speakers who are either politician, political appointee, or politically related others are grouped as political occupation. We point out that the number of speakers involved in certain Asian or Hispanic/Latino-related groups is small, so the computed statistics may be skewed

Occupation	r_{benign}	$r_{\text{intrusive}}$	r_{all}	$\frac{r_{\text{benign}}}{r_{\text{intrusive}}}$
Journalist	17.5%	21.5%	39.0%	81.4%
Political Occupation	13.9%	20.3%	34.2%	68.4%
Gender				
Female	18.2%	21.3%	39.4%	85.5%
Male	15.5%	20.5%	35.9%	75.4%

Table 3: Interruption rates (in percentage) by occupation and gender groups.

Inward - Outward Speaker	Benign	Intrusive	Count
F-F	48.3%	51.7%	29,896
M-F	47.2%	52.8%	64,462
F-M	45.7%	54.3%	61,249
M-M	38.7%	61.3%	98,016
			253,623

Table 4: Proportion of benign and intrusive interruptions (in percentage) and count of each pair. F stands for female and M stands for male.

towards the individuals involved.

Overall, journalists interrupt more ($r_{\text{all}} = 39.0\%$) than speakers with political occupations (34.2%) (Table 3). Since journalists include talk show hosts and news anchors, it is understandable that they interrupt more to mediate and control the flow of the show/interview. When we compute benign/intrusive interruption ratios ($\frac{r_{\text{benign}}}{r_{\text{intrusive}}}$) for each occupation, journalists show higher ratio (81.4%) compared to politically related speakers (68.4%), indicating more frequent use of overlaps or co-operative interruptions during conversations. Similarly, although female speakers tend to use more interruptions (39.4%) than male counterparts (35.9%), they show higher benign to intrusive ratio (85.5% v.s. 75.4%).

6 Discussion

Gender and interruption. To study if there exists any conversational inequality between genders, we compute the average number of words per line spoken by annotated speakers from all collected transcripts (Table 1), where speakers are grouped according to their occupation, gender, and race/ethnicity. Comparing male and female speakers, we find that male speakers have a higher average word per line across all but one pair (Black/African American in political occupation) grouped by occupation and race/ethnicity.

Our result indicates that female speakers interrupt more than male speakers, however. When the overall interruption rates are computed for two gender groups, the female group shows higher interruption rate (39.4%) compared to male interruption rate (35.9%) (Table 3). This interruption pattern is consistent when compared across different occupation and race/ethnicity groups (Table 1). This result is in line with Kennedy and Camden [1983] and Nohara [1992], where the authors found higher female interruption rates among graduate and undergraduate participants. Education and social status can influence the interruption patterns of the chosen speakers [Kennedy and Camden, 1983]. A big part of our annotated speakers are journalists or politically related, receiving college education from a range of academic disciplines. Since we selected the most frequently appearing speakers to annotate, there may be a selection bias in effect as well. Female speakers in our pool may be professionally experienced enough to deny getting subjugated to male dominance as described in Zimmerman and West [1996].

Interestingly, Table 4 shows that male speakers tend to use more intrusive interruption when talking to male speakers (61.3%) than talking to female speakers (54.3%). Whereas female speakers tend to use more intrusive interruption when talking to male speakers (52.8%) than talking to female speakers (51.7%).

Inward - Outward Speaker	Benign	Intrusive	Count
D-D	34.4%	65.6%	2,411
R-D	25.2%	74.8%	4,889
D-R	29.5%	70.5%	3,935
R-R	38.3%	61.7%	3,892
			15,127
D-CNN	30.6%	69.4%	14,331
R-CNN	27.5%	72.5%	19,241
D-Fox	34.1%	65.9%	4,135
R-Fox	41.6%	58.4%	5,625
D-MSNBC	32.3%	67.7%	2,863
R-MSNBC	22.9%	77.1%	2,496
			48,691

Table 5: Proportion of benign and intrusive interruptions (in percentage) and count of each pair. D stands for Democrat and R stands for Republican. We include only journalists when outward speaker is marked as one of the news media.

Political orientation and interruption. To study if one's political orientation can have an impact on interruption patterns, we compute benign and intrusive interruption rates among 272 Republican and 229 Democratic speakers (Table 5). Our analysis reveals that speakers use more intrusive interruptions (hence less benign interruptions) when talking to speakers with different political orientations. Democrats employ intrusive interruption 65.6% of the time when talking to fellow Democrats while this number is 74.8% against Republicans. Similarly, Republicans show 61.7% intrusive interruption rate when talking to Republicans and 70.5% when talking to Democrats.

Next, we examine whether the political bias of the news media influences the interruption pattern by measuring how the interruption styles of journalists representing each news media change depending on who they are talking to. Table 5 shows benign and intrusive interruption rates when journalists speak with people identified as Democratic/Republican. In all comparisons, we observe that journalists in CNN or MSNBC use more intrusive interruptions when speaking with Republicans (69.4% (D) v.s. 72.5% (R) for CNN, 67.7% (D) v.s. 77.1% (R) for MSNBC), while journalists in Fox News use more intrusive interruptions when speaking with Democrats (65.9% (D) v.s. 58.4% (R)).

Politically least biased news media will show the minimum difference in interruption pattern when talking to speakers with different political orientations. The difference in intrusive interruption rates (D - R) is -3.1% for CNN, -9.4% for MSNBC, and 7.5% for Fox News. We hypothesize the magnitude of the difference is related to how much political bias each news media has. The fact-checking website Adfontesmedia.com (measured on 10 May 2022) assigned -8.59 bias points to CNN, -14.17 to MSNBC, and 14.07 to Fox News. The negative number indicates the news media leaning towards the left while the positive indicates to the right. The Spearman's correlation between intrusive interruption rate and political bias of the three news media is 1.

7 Limitations

Selection bias. The subjects included in our analysis are narrowly scoped; all speakers are English speakers appearing in one of the three major US cable news networks. The majority of them have a career in either journalism or politics, and it is likely that they received a high level of education and accomplished a strong social status as the most frequently appearing speakers are chosen for labeling. Therefore, the selected speakers may show particular resistance to sexual and racial biases that are present in other situations. Moreover, culture plays an important role in how interruption is employed in conversations [Murata, 1994]. The reported patterns in interruption may well be specific to the US.

The number of speakers in certain sub-groups (e.g., Asian and Hispanic/Latino) is relatively lower than the other groups since the transcripts are collected from American news networks. For these groups, the interruption pattern could be skewed towards the disposition of the selected few.

Types of interruption and dominance. Care should be taken when generalizing our findings to overall interruption patterns. Our analysis grouped co-operative interruption and overlap as benign interruptions and floor-taking, topic-changing, and disagreement interruptions as intrusive interruptions. Even though we consider overlap and co-operative interruption as equally benign in terms of their functionality, this may not be true concerning the level of power dominance. Ferguson [1977] suggests overlap as the most reliable index of dominance as the individuals who brought in more overlaps tended to rate themselves as highly dominant. Hence, our definition of intrusive interruption does not fully capture the level of power dominance played by certain speakers.

If a certain type of interruption is more closely related to the projection of dominance, a study of different types of interruption may shed more light on understanding who interrupts whom and when. Table 2 indicates that fine-grained interruption classification has considerable room for improvement ($\mathcal{M}_1$). Building a more robust fine-grained interruption classifier is a worthy future research challenge.

In addition, several works associated successful interruptions more closely with dominant behavior than unsuccessful ones, though the results seem ambiguous [James and Clarke, 1993]. This hypothesis has an intuitive appeal as dominance can be more associated with the inward speaker yielding the floor (hence, a successful interruption). Although our analysis doesn't distinguish between successful and unsuccessful interruptions, we have an abundance of both interruptions in our dataset. Future work may confirm if yielding one's floor has any material impact on identifying asymmetrical dominance behavior between conversational participants.

Non-content cue. Non-content cues, such as intonation, body motion, pitch, loudness, and drawl, are known to serve as an important signal in a turn-switching system [Duncan, 1972]. For example, drawling the last syllable, a drop in paralinguistic pitch or loudness, and relaxation of tensed hand position are considered turn-yielding cues. Since our analysis focuses on the textual transcription of conversations, we are limited to using content cues. However, our annotation result could have changed if the annotators had access to non-content cues as a context of the dialogue. In particular, it was difficult for annotators to capture the use of sarcasm or joke by the interrupter, and we believe many such dialogues were marked as intrusive interruption.

8 Conclusion and Future Scope

With the analysis and release of a large scale interruption dataset, this study provides a methodology to understand interruption patterns among speakers with diverse backgrounds in terms of gender, race/ethnicity, occupation, and political orientation. Some of the key findings highlighted in our paper are the conversational imbalance when it comes to male and female speakers, and moreover the dynamics of political affiliations which impact the ability to interrupt a conversation. We highlight two encouraging directions to extend our research. Our large scale dataset can be used to better understand Question Answering systems from a real-world perspective, and predict when an interruption could occur in a conversation between a group of individuals. Due to the unique nature of our dataset, covering USA politics, this resource can be further used to understand the sociopolitical happenings in the country and raise awareness about the gender and racial disparity observed in popular news networks.

Ethical Statement

We emphasize that our goal is not to draw broad conclusions about interruptions concerning gender, race, ethnicity, occupation, and political belief but to study existing patterns within the specific group of collected dialogues and speakers involved. Therefore, whenever we say "speakers", it necessarily implies "speakers in our collected samples", acknowledging that our results are limited by the narrow scope of speakers included in our dataset.

Acknowledgements

We thank Justine Cassell, Lipika Mazumdar, and Tom M. Mitchell for their valuable input.

References

[Anderson and Leaper, 1998] Kristin J. Anderson and Campbell Leaper. Meta-analyses of gender effects on conversational interruption: Who, what, when, where, and how. *Sex roles*, 39(3):225–252, 1998.

[Aries, 1996] Elizabeth Aries. *Men and women in interaction: Reconsidering the differences*. Oxford University Press, 1996.

[Beattie, 1982] Geoffrey W. Beattie. Turn-taking and interruption in political interviews: Margaret Thatcher and Jim Callaghan compared and contrasted. *Semiotica*, 39:93–114, 1982.

[Cho *et al.*, 2014] Kyunghyun Cho, Bart van Merrienboer, Çaglar Gülçehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder–decoder for statistical machine translation. In *EMNLP*, 2014.

[Demszky *et al.*, 2019] Dorottya Demszky, Nikhil Garg, Rob Voigt, James Zou, Jesse Shapiro, Matthew Gentzkow, and Dan Jurafsky. Analyzing polarization in social media: Method and application to tweets on 21 mass shootings. In *NAACL-HLT*, pages 2970–3005, 2019.

[Devlin *et al.*, 2019] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *NAACL-HLT*, 2019.

[Duncan, 1972] Starkey Duncan. Some signals and rules for taking speaking turns in conversations. *Journal of personality and social psychology*, 23(2):283, 1972.

[Esposito, 1979] Anita Esposito. Sex differences in children's conversation. *Language and Speech*, 22:213 – 220, 1979.

[Ferguson, 1977] Nicola Ferguson. Simultaneous speech, interruptions and dominance. *British Journal of social and clinical Psychology*, 16(4):295–302, 1977.

[Gal and Ghahramani, 2016] Yarin Gal and Zoubin Ghahramani. A theoretically grounded application of dropout in recurrent neural networks. In *Advances in neural information processing systems*, 2016.

[Gebru *et al.*, 2021] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna M. Wallach, Hal Daumé III, and Kate Crawford. Datasheets for datasets. *Commun. ACM*, 64(12):86–92, 2021.

[Grebelsky-Lichtman and Katz, 2019] Tsfira Grebelsky-Lichtman and Roy Katz. When a man debates a woman: Trump vs. Clinton in the first mixed gender presidential debates. *Journal of Gender Studies*, 28:699 – 719, 2019.

[Hyun and Moon, 2016] Ki Deuk Hyun and Soo Jung Moon. Agenda setting in the partisan TV news context: Attribute agenda setting and polarized evaluation of presidential candidates among viewers of NBC, CNN, and Fox News. *Journalism & Mass Communication Quarterly*, 93(3):509–529, 2016.

[Iyengar *et al.*, 2012] Shanto Iyengar, Gaurav Sood, and Yphtach Lelkes. Affect, not ideologya social identity perspective on polarization. *Public opinion quarterly*, 76(3):405–431, 2012.

[James and Clarke, 1993] Deborah James and Sandra Clarke. Women, men, and interruptions: A critical review. *Gender and conversational interaction*, pages 231 – 280, 1993.

[Kennedy and Camden, 1983] Carol W. Kennedy and Carl T. Camden. A new look at interruptions. *Western Journal of Communication (includes Communication Reports)*, 47(1):45–58, 1983.

[KhudaBukhsh *et al.*, 2021] Ashiqur R. KhudaBukhsh, Rupak Sarkar, Mark S. Kamlet, and Tom M. Mitchell. We don't speak the same language: Interpreting polarization through machine translation. In *AAAI 2021*, pages 14893–14901, 2021.

[KhudaBukhsh *et al.*, 2022] Ashiqur R. KhudaBukhsh, Rupak Sarkar, Mark S. Kamlet, and Tom M. Mitchell. Fringe news networks: Dynamics of US news viewership following the 2020 presidential election. In *14th ACM Web Science Conference 2022*, pages 269–278. ACM, 2022.

[Kingma and Ba, 2015] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2015.

[Liaw *et al.*, 2018] Richard Liaw, Eric Liang, Robert Nishihara, Philipp Moritz, Joseph E Gonzalez, and Ion Stoica. Tune: A research platform for distributed model selection and training. *arXiv preprint arXiv:1807.05118*, 2018.

[Murata, 1994] Kumiko Murata. Intrusive or co-operative? a cross-cultural study of interruption. *Journal of Pragmatics*, 21:385–400, 1994.

[Nohara, 1992] Michiko Nohara. Sex differences in interruption: An experimental reevaluation. *Journal of Psycholinguistic Research*, 21:127–146, 1992.

[Poole and Rosenthal, 1984] Keith T. Poole and Howard Rosenthal. The polarization of american politics. *The journal of politics*, 46(4):1061–1079, 1984.

[Sacks *et al.*, 1978] Harvey Sacks, Emanuel A. Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction*, pages 7–55. Elsevier, 1978.

[Tannen, 1994] Deborah Tannen. *Gender and discourse*. Oxford University Press, 1994.

[Wolf *et al.*, 2020] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. Transformers: State-of-the-art natural language processing. In *EMNLP:System Demonstrations*, pages 38–45, 2020.

[Zimmermann and West, 1996] Don H. Zimmermann and Candace West. Sex roles, interruptions and silences in conversation. In *Amsterdam studies in the theory and history of linguistic science series 4*, pages 211–236. John Benjamins, 1996.