Understanding Issue Attention Using Variation in Vocal Pitch*

Bryce J. Dietrich†

Draft: November 10, 2016

Abstract

Do members of Congress remain "on message" when speaking about partisan issues? For legislative scholars, this question is essential to understanding floor speeches in the U.S. House of Representatives. Whether it is the "Democratic Message Group" or "Republican Theme Team" parties have become increasingly interested in organizing speaking efforts on the House floor, but such efforts work against an individual representative's ability to get noticed by the media. By speaking passionately, not only can members of Congress increase their media exposure, but they can also signal to their party the degree to which they are committed to party issues. While there are a number of ways to assess a speaker's "passion," this study utilizes non-verbal expressions, such as raising the tone of one's voice. Using the audio from 74,158 floor speeches, this study shows party members are more likely to become emotionally activated when discussing their party's issues. Ultimately, this creates an environment in which party polarization and legislative gridlock are both more likely.

---

*Comments and suggestions welcome. Thank you to Graeme Hirst and the those who attended the 2016 "New Directions in Text as Data" conference for the helpful feedback.

†Bryce J. Dietrich is an Assistant Professor of Social Science Informatics in the Department of Political Science and Department of Sociology at the University of Iowa.

Introduction

The study of elite polarization is nearly synonymous with the study of legislative behavior. While most agree that members of Congress have become more polarized over time, there is considerable debate over the causes of such polariztion. Even though this study does not aim to say definitively what has caused Democrats and Republicans to become increasingly divided, this study suggests part of this divisivness is grounded in a subconscious emotional attachement to partisan issues. Using the audio from 75,000 floor speeches, this study shows not only do Democrats and Republicans talk more about their party's issues, but when they do they become more emotionally activated. Even though there are a number of ways to measure emotional activation, this study relies on subconcious changes in vocal inflections. When people think of partisanship at the mass-level they often think of the "sentimental citizen" (). This study suggests members of Congress are equally driven by emotion, especially when it comes to their partisanship.

The 2016 election demonstrated voters are becoming increasingly angry (). While some of this anger is directed at the federal government (), most of it is directed at the opposing party (). Among other things, "Democrats are angry that Republicans have thwarted efforts to spend more on social programs and increase taxes on higher income Americans," while "Republicans are angry that Democrats won't cut spending, except growth in planned spending, which is not the same as spending cuts" ().

Unsurprisingly, angry voters produce angry candidates. Again, the 2016 election is a telling example. Both Donald Trump and Bernie Sanders were able to tap into this sentiment among primary voters. For Donald Trump, he was able to voice the frustration and anger felt by a large segment of the Republican base towards President Obama and many of his poliices. For Bernie Sanders, he was able to tap into Democratic anger towards Republican efforts to thwart policies such as raising the minimum wage, addressing climate change, and regulating Wall Street. Even though Bernie Sanders did not win the Democratic nomination, both candidates were able to voice the anger that existed in their respective bases, ultimately helping them mobilize their supporters.

Members of Congress face a similar problem. As voters become more and more ideologically sorted into political parties, Democratic canditates have to become more and more liberal to win Democratic primarites. The same is true for Republicans trying to win over increasingly conservative prirmary voters. However, increasingly this ideological divide has been associated with an equally large emotional divide. For both Democrats and Republicans, it seems like it is no longer enough to simply toe the party line. Instead, you have to strike the right emotional chord with the electorate. However, demonstrating this is easier said than done. Popkin (1994) puts the argument in this way:

> When voters watch a candidate perform on television, making promises and taking hard-line rhetorical positions on issues, they question if there is congruence between avowal and actual feelings-whether the candidate's support for a cause represents a genuine personal commitment or only a campaign tactic... When voters estimate a candidate's preferences, they take account of sincerity-whether the candidate really cares about their concerns (65).

However, since it is difficult for voters to know whether a member of Congress' commitment is sincere, they must "take shortcuts" (Popkin, 1994, p. 65). This study argues vocal inflections are one of these shortcuts.

It seems as though "[w]e are in an angry, angry time..." (). Among other things, "Democrats are angry that Republicans have thwarted efforts to spend more on social programs and increase taxes on higher income Americans," while "Republicans are angry that Democrats won't cut spending, except growth in planned spending, which is not the same as spending cuts" (). This partisan outrage is increasingly being reflected on Capitol Hill.

Data and Methods

Data

The data used in this study was compiled from *HouseLive*. Started as part of the "Open Government" initiative, *HouseLive* has videos of House floor proceedings going back to January 6, 2009. Even though `MP4`s are available up to the current week, `MP3`s are only available up until August 4, 2014. Even though one can obtain audio from an `MP4` file, downloading and processing `MP4` files takes considerably longer. Given that, the present study only considers audio from January 6, 2009 to August 4, 2014. Ultimately, the data set is composed of 6,432 hours of audio from 863 legislative debates.

To identify both the speakers and what was being said, I used the closed-captioning information provided by *HouseLive*. Unlike the *Congressional Record*, closed-captioning information has the advantage of reporting what is actually said on the House floor. Even though the *Congressional Record* is very similar to the words spoken on the House floor, members of Congress can change the *Congressional Record* after the fact. Often times members of Congress use this power to add things to the *Congressional Record* that were not actually spoken on the House floor. For example, members of Congress often read things into the *Congressional Record*, meaning instead of reading out loud a letter they received from a concerned constituent they will simply ask the letter be added to the *Congressional Record*. When one reads the *Congressional Record* it is readily apparent when documents are added, but it is difficult to make this determination using automated methods. Closed-captioning information does not suffer from similar limitations since it is a live transcript of what is being said.

With that said, since closed-captions are produced in real-time, things like typographical errors, may be a concern. In email correspondence, the company that performs the closed-captioning service for the House of Representatives asserts their transcribers are generally 95 percent accurate, meaning 95 percent of the time the words that are transcribed are the words actually spoken on the House floor. This assessment is based on yearly evaluations, in which the company randomly selects a certain number of transcripts from each of their transcribers and determines the degree to which those transcripts capture the floor debate for that day. For this study, 100 randomly selected speeches were transcribed. When these speeches were compared to the closed-captioning information, regardless of the similarity measure one used, the closed-captions essentially mirrored the transcripts. Based on these results and my communication with the closed-captioning company, I am confident that the closed-captioning found on *HouseLive* is an accurate reflection of what is said in the U.S. House of Representatives.

Even though this paper is primarily interested in non-verbal expressions, using closed-captioning information is an important contribution to the study of floor behavior in and of itself. Not only does closed-captioning information not suffer from the limitations of the *Congressional Record*, but closed-captioning information can be obtained from a variety of sources. For example, most television streams include closed-captioning information, meaning understanding the textual dynamics found within closed-captioning information has more general applications. Indeed, the company that provides the closed-captioning for the House of Representatives does the same for several state legislatures. Even though the *Congressional Record* is useful for understanding what transpires on the House floor, it is very unique and does not necessarily reflect records kept in other institutions. Closed-captioning information is provided by most legislatures, domestic and international, making it potentially a more universal data source for those interested in understanding floor behavior more broadly.

Emotional Activation

On average, each MP3 was 7 hours and 27 minutes long. Using ffmpeg, these longer audio files were split into individual speeches using the time stamps found in the closed-captioning information. This resulted in 152,117 WAV files. Many of these files were fairly short, so this study only uses floor speeches that had at least 50 words. Ultimately, this resulted in the text and audio of 74,158 speeches. Even though one could use the full corpus, often times the closed-captioning transcriber would enter a separate line for statements like "uh" and "um." These statements are not included in the *Congressional Record* and could be an interesting area of future research, but are not particularly useful for understanding partisan vocal inflections.

Using *Praat*, I extracted the mean fundamental frequency (F_0) from each floor speech. Countless studies "have routinely shown that [pitch] F_0-related measures, such as mean F_0 and over-all pitch contour, are influenced by affect-related arousal," meaning individuals who speak with higher levels of F_0 are generally more emotional than those who do not (Owren and Bachorowski, 2007, p. 240). More formally, Titze (2000) defines F_0 using this equation:

$$F_0 = \frac{1}{2L}\sqrt{\frac{\sigma}{\rho}} \tag{1}$$

where L is the vocal fold length, σ is the longitudinal stress on the vocal folds, and ρ is the vocal fold tissue density. Here, "...voice pitch is inversely proportional to vocal fold length and directly proportional to the square root of tension on the vocal folds" (Puts, Gaulin, and Verdolini, 2006, p. 284). Thus, "[l]onger vocal folds with less tension on them lead to lower voice pitch" (Puts, Gaulin, and Verdolini, 2006, p. 284). This relationship is why men typically talk with lower pitch than women. Specifically, male vocal folds are generally between 17.5 and 25mm long, whereas female vocal folds are typically between 12.5 and 17.5mm long, making the denominator for the first part of Equation 1 smaller for women, increasing their F_0 (Titze, 2000).

For the most part, individual variations in vocal fold length (L) and density (ρ) are determined by genetics (e.g. Przybyla, Horii, and Crawford, 1992; Debruyne et al., 2002). For example, vocal fold length is positively correlated with size (both height and weight), which is one of the reasons why a boy's voice drops once he hits puberty (Fitch and Giedd, 1999). Conversely, variations in longitudinal stress (σ) are almost entirely determined by the speaker. For example, "[e]motional activation raises F_0

by increasing tension on the focal fold mucosa (σ, in Eq. (1)), mainly via contraction of the circothyroid muscles and consequent lengthening of the vocal folds" (Puts, Gaulin, and Verdolini, 2006, p. 285).

Thus, higher F_0 is associated with "...high-activation emotions such as hot-anger, elation, and panic fear, whereas lowered F_0 is associated with sadness, boredom, and contempt (low-activation emotions)" (Puts, Gaulin, and Verdolini, 2006, p. 285). In one very influential study Banse and Scherer (1996) found using F_0-related measures "judges are able to accurately recognize virtually all of the large set of emotions..." (624). These results and others (for review see Bachorowski and Owren, 2003; Calvo and D'Mello, 2010; Cowie et al., 2001; Owren and Bachorowski, 2007; Zeng et al., 2009) are why "F0 and F0 variability" are the "acoustic properties typically emphasized in searching for cues to vocalizer emotional state..." (Bachorowski and Owren, 2003, p. 246).

Beginning with Scherer and Oshinsky (1977), scholars have also suggested speaking rate, or the number of words per second, may also be associated with emotional arousal (Apple, Streeter, and Krauss, 1979; Breitenstein, van Lancker, and Daum, 2001; Davitz, 1964; Kehrein, 2002; Laukka, Juslin, and Bresin, 2005). This mostly has to do with the known association with speaking rate and blood pressure. More specifically, when an individual's blood pressure rises they tend to speak at a higher rate (Friedmann et al., 1982; Siegman, Anderson, and Berger, 1990). Even though the literature documenting the relationship between speaking rate and emotional activation is considerably less, speaking rate is used in the present study to validate mean vocal pitch. If one finds both are correlated, then it provides some evidence that mean vocal pitch is capturing the degree to which a speaker is emotionally activated. Similarly, if party members increase both their speaking rate *and* vocal pitch when speaking about their party's issues, then it provides additional evidence that partisan issues produce a more animated reaction on the House floor. The same can be said for both variables being predictive of polarization. In the literature on vocal communication of emotion "the most reliable finding is that pitch appears to be an index into arousal," which is why the primary variable of interest in this study is mean vocal pitch (Calvo and D'Mello, 2010, p. 24).

Structural Topic Model

With these measures in hand, the next step is to identify whether the speaker is delivering a speech on an issue owned by either the Democratic or Republican party. To achieve this end, this study uses a modified version of Latent Dirichlet Allocation (LDA) topic modeling (Blei, Ng, and Jordan, 2003).

Generally speaking, topic modeling is a computer-assisted process in which a corpus of documents are grouped into substantively meaningful "topics." Like other unsupervised topic models, the researcher does not provide the LDA topic model any information about the topics that exist within the corpus. Instead, the model measures the patterns of co-occurrence of words in individual texts across the entire corpus. While researchers could use a supervised approach in which a training set is used to classify the rest of the corpus, when there are a large number of documents often times such approaches are not tractable. Even in a world where researchers have the ability to code the entirety of the corpus, some topics may be too subtle for the coding protocol to detect, especially when those topics are unknown a priori. When it comes to floor speeches, one generally knows the types of topics members of Congress discuss, but the nuance that exists both between and within those topics is difficult to predict. Recent research has shown that unsupervised topic models, like the one employed in this study, have been performed very well in a variety of applications (e.g., Farrell, 2016; Genovese, 2015) making them particularly useful for understanding what topics are discussed on Capitol Hill.

More specifically, this study uses a recently developed approach to topic modeling called "Structural Topic Modeling" (STM) (Roberts et al., 2013; Roberts, Stewart, and Tingley, 2014; Roberts et al., 2014). Unlike standard LDA models, the STM model allows the researcher to incorporate metadata, such as when a floor speech was given or whether the floor speech was delivered by a liberal or conservative member of Congress, in order to improve the estimated LDA. Similar to LDA, each document is modeled as a mixture of K topics. Unlike standard LDA, in STM, topic proportions (θ) can be correlated and the likelihood of those topics can be influenced by some set of covariates X which are modeled using a logistic normal regression. For each word (w) within a given document, a topic (z) is drawn from the document-specific distribution and conditioned on that topic. The word itself is chosen from a multinomial distribution which is parameterized by β and represents the deviations from the baseline word frequencies (m) in log space. This distribution can also vary by some set of covariates U. These additional covariates "structure" the topic model's prior distributions, allowing researchers to use their substantive knowledge of the corpus to facilitate topic identification (for additional details, see Roberts et al., 2014, p. 4).

To use an STM, the researcher has to make a number of choices, most notably the covariates used in order to fit the topic model and the number of topics. For this study, I included two covariates: the speaker's ideology and the date of the speech. The former was measured using DW-Nominate scores which range from -1 ("liberal") to 1 ("conservative"). These scores are calibrated using floor votes and have been consistently used as measures of congressional ideology (Poole and Rosenthal, 2001). I assume that representatives who are on the same side of the ideological spectrum are likely to speak about similar issues, meaning the words used in those speeches are likely to be clustered. I also included the date of the speech as a covariate. This variable was measured in days since the first date in the data set – January 1, 2009. I assume that each legislative day is restricted to a handful of topics, meaning representatives are likely to deliver similar speeches on the same day. If this is the case, then one would expect words appearing on the same day are more likely to be associated with one another.

Unfortunately, in both instances, it is unclear the exact relationship between each covariate and the topics being discussed. Given that, I fitted a b-spline (or basis spline) to both the speaker's ideology and the date of the speech. These smoothed covariates were the ones ultimately used to estimate the STM. Unlike other smoothing functions that do not allow the curve to change locally without causing changes to the full length of the curve, a b-spline is fairly flexible, allowing both local cusps and additional points to be added without increasing the degree of the curve. For these reasons, the b-spline is considered to be "more numerically stable than the cubic spline" (Keele, 2008, p. 59). It is unclear whether smoothing is necessary to fit the model, but b-splines have been used in other STM applications (e.g., Roberts, Stewart, and Airoldi, 2016), making this choice consistent with previous work.

Similar to other unsupervised models, the most important decision the researcher has to make is the number of topics (k). Roberts, Stewart, and Tingley (2014) provide two measures that help researchers with this choice: *semantic coherence* and *exclusivity*. A topic is coherent when the average of pairwise word similarities formed by the topic's top words is high. The average exclusivity of a topic is the degree to which the topic's words appear in that particular topic to the exclusion of others. A "good" topic is one that maximizes both measures, meaning the topic has words that are similar to one another and are unique to that topic. Using these two measures as guidelines, I estimated nineteen structural topic models, with k ranging from 10 to 100, increasing in increments of 5. Ultimately, $k = 30$ and $k = 35$ yielded very similar results. To break the tie, I used the held-out likelihood which estimates the probability of words appearing within a document when those words have been removed prior to estimation (Blei, Ng, and Jordan, 2003). When this was done, a 30-topic model performed slightly better which is why I ultimately chose $k = 30$.

Table 1: Topic Labels

Topic	Word 1	Word 2	Word 3	Word 4	Word 5	Label	Proportion
1	court	case	justic	judg	law	judicial	0.01
2	right	peopl	constitut	american	freedom	rights	0.03
3	colleagu	support	today	like	new	collegiality 1	0.05
4	work	make	need	peopl	can	values	0.07
5	war	militari	afghanistan	forc	defens	middle east 1	0.02
6	school	educ	student	colleg	communiti	education	0.02
7	republican	american	democrat	will	pass	party	0.04
8	busi	small	regul	cost	will	business	0.02
9	budget	spend	cut	year	debt	spending cut	0.04
10	secur	nation	inform	protect	agenc	security	0.03
11	energi	oil	gas	will	price	energy	0.02
12	state	unit	texa	border	come	immigration 2	0.02
13	care	health	insur	will	cost	health care	0.04
14	women	children	famili	live	life	children	0.03
15	nuclear	israel	iran	world	peac	middle east 2	0.02
16	job	economi	creat	american	econom	jobs	0.04
17	peopl	get	thing	talk	got	discursive 1	0.05
18	honor	year	great	serv	first	collegiality 2	0.04
19	transport	build	new	system	air	transportation	0.02
20	financi	credit	loan	bank	street	financial	0.02
21	will	side	pass	floor	debat	procedural 1	0.05
22	water	land	area	communiti	nation	land	0.03
23	law	immigr	enforc	victim	crime	immigration	0.02
24	say	think	know	want	one	discursive 2	0.08
25	fund	program	million	provid	billion	spending	0.04
26	tax	govern	pay	feder	american	tax	0.03
27	administr	quot	report	obama	public	administration	0.03
28	act	requir	author	law	provis	procedural 2	0.05
29	famili	food	benefit	million	cut	welfare	0.02
30	servic	veteran	nation	serv	support	veterans	0.03

Even though each of these choices are important, they are not the same thing as choosing regression parameters. When describing LDA, DiMaggio, Nag, and Blei (2013) put the argument this way:

> Think of the model as a lens for viewing a corpus of documents. Finding the right lens is different than evaluating a statistical model based on a population sample. The point is not to estimate population parameters correctly, but to identify the lens through which one can see the data most clearly. Just as different lenses may be more appropriate for long-distance or middle-range vision, different models may be more appropriate depending on the analyst's substantive focus (582).

Table 2: "Security" and "Social Welfare" Topics

(a) Security Topics

Topic	Word 1	Word 2	Word 3	Word 4	Word 5	Label	Proportion
5	war	militari	afghanistan	forc	defens	middle east 1	0.02
10	secur	nation	inform	protect	agenc	security	0.03
15	nuclear	israel	iran	world	peac	middle east 2	0.02
30	servic	veteran	nation	serv	support	veterans	0.03

(b) Social Welfare Topics

Topic	Word 1	Word 2	Word 3	Word 4	Word 5	Label	Proportion
6	school	educ	student	colleg	communiti	education	0.02
13	care	health	insur	will	cost	health care	0.04
14	women	children	famili	live	life	children	0.03
25	fund	program	million	provid	billion	spending	0.04
29	famili	food	benefit	million	cut	welfare	0.02

In this paper, the substantive focus is to identify whether members of Congress become more emotionally activated when talking about their party's issues. Table 1 shows the top 5 words appearing in each topic, as well as the average proportion of each speech dedicated to that topic. The labels were added post-hoc and hopefully help the reader understand the substantive importance of each topic. While all the topics are undoubtedly interesting, Petrocik (1996) argues Republicans are more committed to "security" issues and Democrats are more committed to "social welfare" issues. After looking through several example documents, topics 5, 10, 15, and 30 were defined as "security" topics. Similarly, topics 6, 13, 14, 25, and 29 were defined as "social welfare" topics. Table 2 shows the topics included in the "security" and "social welfare" topics as well as the top 5 words appearing in each. In the analysis below, the estimated proportion of each speech dedicated to each of these broader topics is captured in the variables `Security Topic` and `Social Welfare Topic`, respectively.

Other Variables

To test whether party members become more emotionally activated when talking about their party's issues, this study is primarily interested in the interaction between a speaker's party identification and the topic being discussed. For example, if Democrats tend to become more emotional when talking about Democratic issues, then the interaction between `Democrat` and `Social Welfare Topic` should be positive and statistically significant. Whether the speaker was a `Democrat` or `Republican` was determined using *Voteview*. Even though the expected relationship is grounded in partisanship, it could be the case that Democrats become more excited about social welfare issues because certain demographic groups, such as women and minorities, are more likely to be members of the Democratic party.

For example, Hall (1998) found that female members of Congress were more active on the House floor during the the Job Training Partnership Act of 1982 and on the Older Americans Act, both of which were supported by the Congressional Caucus for Women's Issues. Gerrity, Osborn, and Mendez (2007) found no difference between male and female legislators when it comes to the frequency of speeches addressing women's issues. With that said, Osborn and Mendez (2010) did find evidence that female Senators were more likely to speak about health and family issues. Along these lines, Pearson and Dancey (2011a, 2011b) found that female members of Congress were not only more likely to speak on the House floor, but when they did they were more likely to reference women, suggesting floor speeches are used for descriptive purposes.

Unfortunately, very few studies have looked at how floor speeches are used by other descriptive groups. One notable exception is Dietrich et al. (2016). In this study, the authors find African-American members of Congress are much more likely to speak about African-American issues, like civil rights. In a companion piece, Hayes and Dietrich (2016) find that white members of Congress are less likely to mention African-Americans at both the federal and state-level, suggesting floor speeches are used by black representatives to not only raise issues important to the African-American community, but they are more likely to relate all issues to African-Americans.

Thus, to isolate the effect of partisanship, each representative's race and gender was obtained from *GovTrack*. I also included controls to capture each speaker's institutional position. Party members who hold less institutional power may be more willing to speak passionately about party issues in order to gain favor with party leaders. Maltzman and Sigelman (1996), Morris (2001) and Harris (2005, 2013) all make some variation of this argument, so a speaker's seniority and whether he or she was a committee chair were both included as controls. The former was also obtained from *GovTrack*. The latter was obtained from (Stewart and Woon, 2016). Along these lines, Maltzman and Sigelman (1996), Morris (2001) or Harris (2005, 2013) all suggest floor speeches are influenced by electoral incentives. Thus, I included a dummy variable for whether the speech was delivered during an election year.

Results

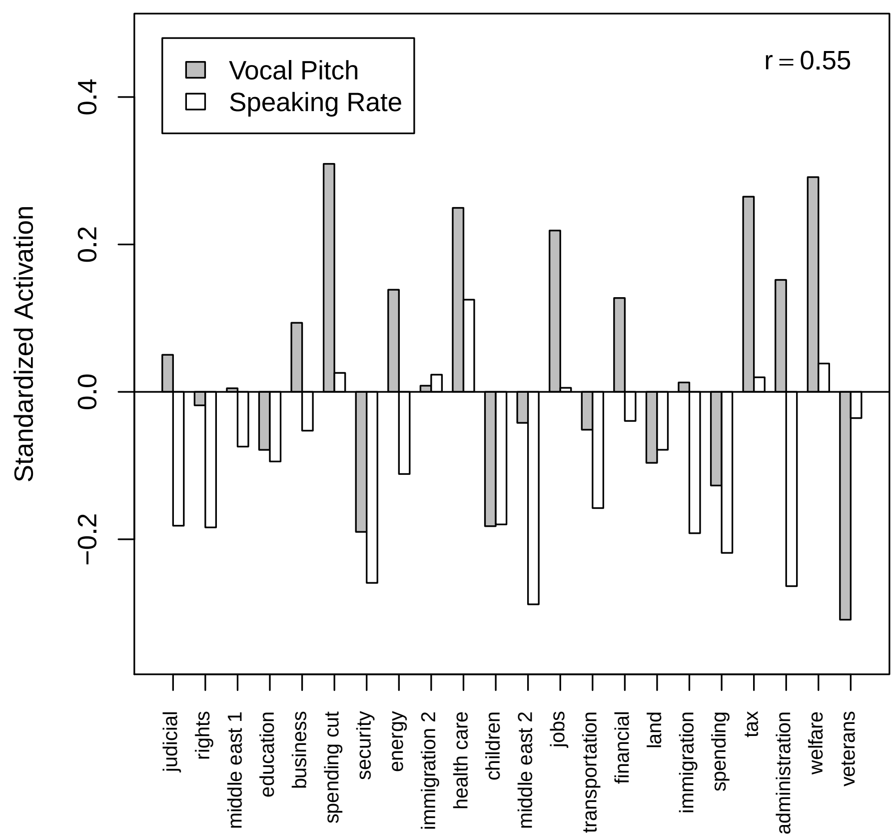
Do Party Members Become More Emotional When Discussing Party Issues?

This study expects to see two relationships. First, when party members talk about party issues they are more likely to display emotion. For Democrats and Republicans, this is operationalized as social welfare and security issues, respectively, meaning Democrats are expected to display more emotion when discussing social welfare issues and Republicans are expected to display more emotion when discussing security issues. Second, when emotion is displayed polarization is more likely. There are a variety of ways to measure polarization, this study uses party votes. A party vote is a vote in which the majority of one party votes against the majority of another. Party votes are expected to be more likely when emotion is displayed, especially when those votes deal with issues important to either the Democratic or Republican parties.

Emotion can be expressed in a variety of ways, but this study suggests non-verbal expressions are particularly important to understanding partisan floor speeches. To isolate this effect, I included the proportion of each speech dedicated to positive and negative emotional words. Both of these words were defined using the Linguistic Inquiry Word Count (LIWC) dictionary, with the former including 407 words like "acceptable," "benevolent," "charming," "devoted," and "elegant." and the latter including 500 words like "abusive," "brutal," "contempt," "destructive," and "envious." Even though there are a number of dictionaries one could use, these categories have been used in a variety of applications (for review, see Tausczik and Pennebaker, 2010), including many within political science (e.g., Owens and Wedeking, 2011; Owens and Wedeking, 2012).

The average vocal pitch and speaking rate associated with the substantive topics is plotted in Figure 1. As explained above, each speaker has their own baseline vocal pitch and speaking rate. For example, women naturally speak at a higher pitch than men because their vocal cords tend to be shorter. Given that, vocal pitch and speaking rate were scaled to standard deviations above and below each representative's baseline. Positive values imply the representative was speaking at a mean vocal pitch above his or her baseline, which suggests he or she was emotionally activated. Negative values

Figure 1: Which Topics Are the Most Emotional?



imply the representative was speaking at a mean vocal pitch below his or her baseline, which suggests he or she was not emotionally activated. To be consistent, the number of positive and negative emotion words were also scaled to standard deviations above or below the speaker's baseline.

Both vocal pitch and speaking activation have been shown to be associated with emotional activation, meaning they should be correlated. Not only does Figure 1 show these variables are highly correlated (0.55), but the correlation is also statistically significant at the .01-level ($t = 2.92, df = 20, p = 0.01$). Speakers also seem to be more emotionally activated when talking about some topics versus others. For example, the topic labeled "spending cut" produces the highest increase in vocal pitch, followed by the topics labeled "health care," "jobs," "tax," and "welfare." The "spending cut" topic contains words like "budget," "spend," "cut," and "debt." The "jobs" and "tax" topics contain words like "job," "economic," "tax," and "pay." These generally may capture economic concerns or may be more narrowly focused on government spending. Why these topics produce such a strong reaction is unclear, but the fact that "health care" produces a similar reaction is not surprising given the contentious debate surrounding the Affordable Care Act.

Similar results are found for social welfare topics. 17.97 percent of Republican speeches (or 6,504) were dedicated to social welfare, which is significantly lower ($\chi^2 = 322.98, df = 1, p < 0.01$) than the 22.39 percent of Democratic speeches (or 8,103) dedicated to the same topic. This suggests Democrats dedicated more time to talking about social welfare issues as compared to Republicans. In this instance, both non-verbal variables are consistent, with Democrats not only speaking at a significantly ($t = 2.97, df = 13,796, p < 0.01$) higher vocal pitch (0.10) than Republicans (0.05), but also speaking at a significantly higher rate. Even though the difference is not as stark between Democrats (-0.01) and Republicans (-0.05) when it comes to speaking rate, it is still statistically significant at the .02-level ($t = 2.33, df = 14,604, p < 0.02$). Even though in this instance the number of speeches was consistent with the non-verbal measures, the latter still yields different insights. For example, Democrats delivered 1.31 times as many social welfare speeches than Republicans, but their vocal pitch and speaking rate were 1.95 and 4.45 times higher. This is not to say that verbal expressions provide no insights into partisan rhetoric, but to suggest that non-verbal expressions may provide different insights.

When speaking about security issues, Republicans speak at a higher vocal pitch and speaking rate than Democrats. The inverse is true for social welfare issues, where Democrats speak at a higher vocal pitch and speaking rate as compared to Republicans. Using a multilevel linear model, Table 3 finds

Table 3: Multilevel Linear Models Predicting Emotional Activation

	Vocal Pitch	Speaking Rate
(Intercept)	0.09***	0.09***
	(0.02)	(0.01)
Republican	−0.01	−0.01
	(0.01)	(0.01)
Security	−3.40***	−2.58***
	(0.16)	(0.15)
Election Year	−0.05***	−0.04***
	(0.01)	(0.01)
Male	−0.01	−0.005
	(0.01)	(0.01)
White	0.02	0.01
	(0.01)	(0.01)
Seniority	0.001***	−0.0003
	(0.0004)	(0.0004)
Committee Chair	−0.06***	−0.01
	(0.01)	(0.01)
Positive Emotion Words	−0.04***	−0.02***
	(0.005)	(0.005)
Negative Emotion Words	0.14***	−0.06***
	(0.005)	(0.005)
Republican × Security	0.53**	0.32
	(0.21)	(0.21)
N	69,251	74,105
Log Likelihood	−97,135.28	−104,284.80
AIC	194,296.50	208,595.60
BIC	194,415.40	208,715.40

*p < .1; **p < .05; ***p < .01

(a) Security Topic

	Vocal Pitch	Speaking Rate
(Intercept)	−0.03*	0.07***
	(0.02)	(0.02)
Democrat	−0.05***	−0.02
	(0.01)	(0.01)
Social Welfare	1.09***	−1.40***
	(0.19)	(0.18)
Election Year	−0.06***	−0.06***
	(0.01)	(0.01)
Male	0.01	−0.01
	(0.01)	(0.01)
White	0.004	0.002
	(0.01)	(0.01)
Seniority	0.001***	−0.001*
	(0.0004)	(0.0004)
Committee Chair	−0.06***	−0.02
	(0.01)	(0.01)
Positive Emotion Words	−0.05***	−0.02***
	(0.005)	(0.005)
Negative Emotion Words	0.13***	−0.08***
	(0.005)	(0.005)
Democrat × Social Welfare	0.99***	0.89***
	(0.25)	(0.24)
N	69,251	74,105
Log Likelihood	−97,457.54	−104,521.40
AIC	194,941.10	209,068.90
BIC	195,060.00	209,188.70

*p < .1; **p < .05; ***p < .01

(b) Social Welfare Topic

similar results when additional covariates are added. In each model, the unit of analysis is an individual speech (Level 1). Each speech is treated as a separate measure of a latent trait that exists within each speaker (Level 2). In panels (a) and (b), emotional activation is the dependent variable and the topics are security and social welfare, respectively. In columns labeled "Vocal Pitch" emotional activation is operationalized using the speaker's average vocal pitch (scaled) for that speech. In columns labeled "Speaking Rate" emotional activation is operationalized using the number of words per second (scaled) for that speech. In both panels, the topic of the speech (Level 1) is interacted with the speaker's party identification (Level 2). In panel (a), a positive and statistically significant interaction between `Republican` and `Security` is expected. In panel (b), a positive and statistically significant interaction between `Democrat` and `Social Welfare` is expected.

Whether one is talking about security or social welfare topics, the models reflect the predicted relationship. First, Republicans tend to speak at a higher vocal pitch than Democrats when they are speaking about security issues. Not only is this relationship statistically significant at the .05-level, but it holds even when the number of positive and negative emotion words are included as predictors. Even though Republicans do not seem to speak at a significantly ($p > .05$) higher rate than Democrats when discussing security issues, the coefficients are in the expected direction. For Democrats, similar results are found. Not only do Democrats speak at a significantly ($p \leq .05$) higher vocal pitch than Republicans when discussing social welfare issues, but they also speak at a significantly ($p \leq .05$) higher rate. As before, these results hold even when the number of positive and negative emotion words are included, suggesting partisanship and topic have an effect on vocal inflections that is distinct from the words that are used.

Figures 2 and 3 plot the predicted probabilities from the multilevel models found in Table 3. In each model, the proportion of the speech dedicated to the each topic was allowed to range from the minimum (0) to the maximum (0.25) while holding all other variables constant at their mean and modal values. Even though the model has a randomly varying intercept, the plots only show the predicted probabilities associated with the fixed effects. Beginning with Figure 2, it is apparent that both Democrats and Republicans lower their vocal pitch and speaking rate as the proportion of their speech dedicated to security issues increases. With that said, Democrats seem to lower their vocal pitch and speaking rate at a slightly faster pace, suggesting Republicans are more resistant to this overall downward trend.

Figure 2: Predicted Probabilities (Security Topic)

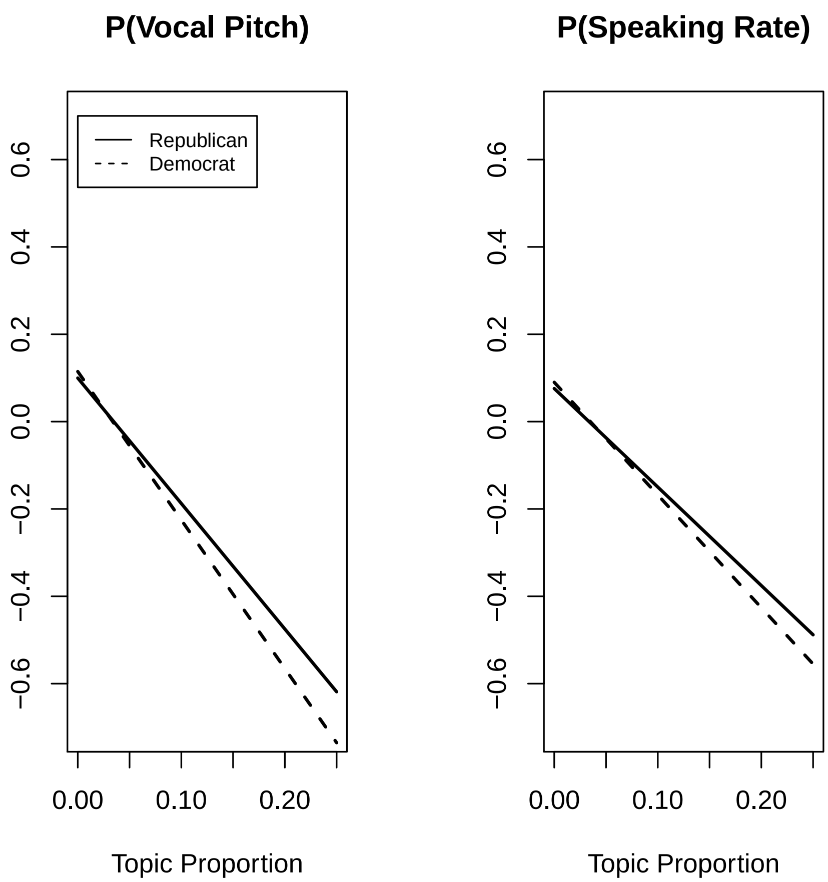
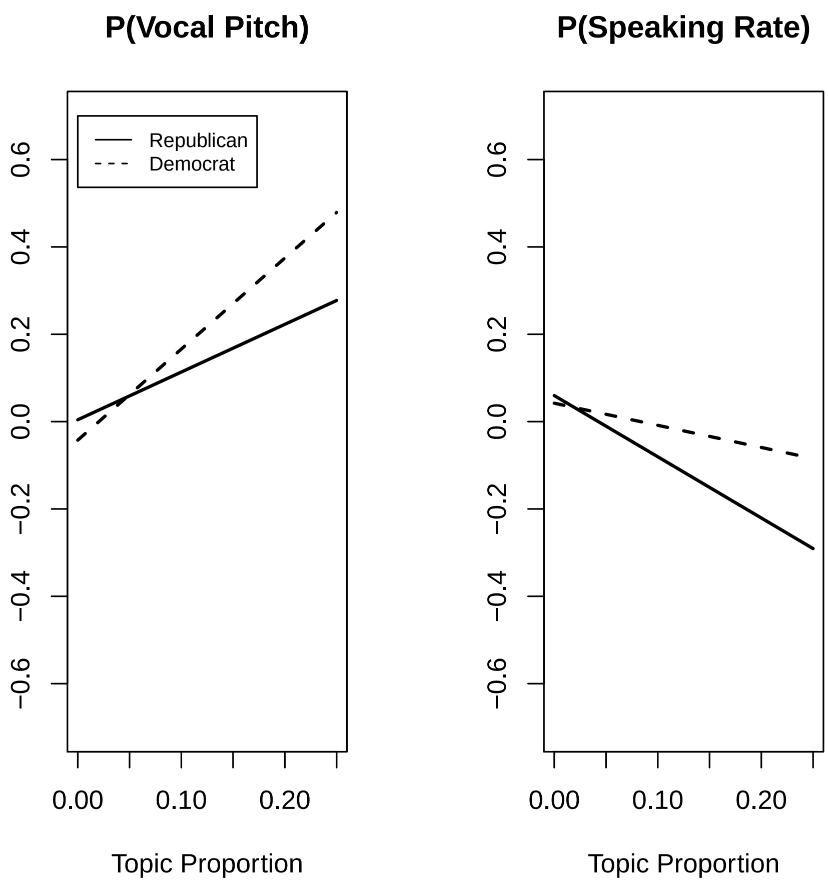


Figure 3: Predicted Probabilities (Social Welfare)



In Figure 3, both Democrats and Republicans increase their vocal pitch as the proportion of their speech dedicated to social welfare issues increases, but the increase is greater for Democrats. Here, the predicted probabilities for speaking rate are in the opposite direction. Instead, of increasing their speaking rate, both Democrats and Republicans decrease their speaking rate when talking about social welfare. It is unclear why this disconnect exists, but in all other instances, the effect of vocal pitch and speaking rate is in the same direction. It is also worth mentioning that the difference between Democratic and Republican vocal pitch is more pronounced in Figure 3, suggesting that social welfare issues may have produced a stronger emotional response during this time period. Given the contentious debates surrounding the Affordable Care Act, finding such a relationship is not too surprising.

Is Emotional Activation Indicative of Party Polarization?

From these results, it seems as though party members tend to become more emotionally activated when discussing party issues. Indeed, whether it is Republicans talking about security or Democrats talking about social welfare, when members of Congress talk about their party's issues they tend to raise their vocal pitch. In almost every instance these changes in vocal pitch mirror similar changes in speaking rate, suggesting members of Congress tend to become more emotionally activated when talking about their party's issues. Table 4 asks a slightly different question. Here, the question is not whether an individual member of Congress becomes more emotionally activated when they are engaged in partisan rhetoric, but whether party polarization is more likely when the overall tone of the debate becomes more heated.

In these models, the unit of analysis is an individual roll-call vote. The dependent variable is whether the vote was a party vote, with a 1 indicating the majority of one party voted against the majority of the other and a 0 otherwise. Given the dichotomous nature of the dependent variable, each model is a standard logistic regression. In the column labeled "Vocal Pitch" emotional activation is operationalized using the average vocal pitch (scaled) for the day on which the roll-call vote occurred. In the column labeled "Speaking Rate" emotional activation is operationalized using the average number of words per second (scaled) for the day on which the roll-call vote occurred. The number of positive and negative emotion words is calculated in a similar way. Instead of reflecting an individual's use of emotional language, each variable captures the average for the day of the vote. The roll-call data was obtained from *Voteview*. Covariates associated with each bill, such as the number of cosponsors, were obtained

Table 4: Logistic Regressions Predicting the Likelihood of a Party Vote

	Vocal Pitch	Speaking Rate
	(1)	(2)
(Intercept)	−1.20***	−1.27***
	(0.31)	(0.30)
Activation	0.59***	0.39**
	(0.14)	(0.19)
Cosponsors	−0.001	−0.002**
	(0.001)	(0.001)
|DW-Nominate|	0.84***	0.95***
	(0.24)	(0.23)
Majority Party	0.75***	0.80***
	(0.13)	(0.13)
Committee Chair	−0.002	−0.02
	(0.09)	(0.09)
Positive Emotion Words	0.82***	0.88***
	(0.11)	(0.10)
Negative Emotion Words	0.73***	0.69***
	(0.11)	(0.11)
N	4,037	4,190
Log Likelihood	−2,375.94	−2,472.41
AIC	4,813.88	5,006.81

*p < .1; **p < .05; ***p < .01

from the *Congressional Bills Project*. These include whether the bill's sponsor was a member of the majority party or served as a committee chair. The absolute value of the sponsor's DW-Nominate score was included to capture whether the sponsor was ideologically distinct from the rest of the chamber. Although not shown, fixed effects are also included for each bill's major topic area, as defined by the *Policy Agendas Project*.

Table 4 shows vocal pitch is a strong predictor of party votes. Not only is the coefficient positive and statistically significant ($p < 0.01$), but when predicted probabilities are calculated holding all other variables at their mean and modal values the substantive effect is sizable. More specifically, when the average vocal pitch for the day of the roll-call vote is allowed to vary from its minimum (-2.38) to maximum (2.43) the probability of a party vote increases from 0.21 to 0.82, suggesting heightened vocal pitch is likely indicative of a polarized environment. A similar result is found for the average speaking rate for the day of the vote. Here, when the average speaking rate is allowed to vary from its minimum (-0.73) to maximum (1.07), the probability of a party vote increases from 0.48 to 0.63. Unsurprisingly, the coefficient associated with the average speaking rate is also statistically significant ($p \leq .05$). To put these results into perspective, when the average number of negative emotion words is allowed to vary from its minimum (-0.64) to maximum (0.82), the probability of a party vote only increases from

Table 5: Logistic Regressions Predicting the Likelihood the Vote Passed

	Vocal Pitch	Speaking Rate
	(1)	(2)
(Intercept)	4.23***	4.27***
	(0.51)	(0.51)
Activation	−0.37**	0.0001
	(0.15)	(0.19)
Cosponsors	0.002**	0.002**
	(0.001)	(0.001)
|DW-Nominate|	−1.09***	−1.15***
	(0.32)	(0.31)
Majority Party	−2.93***	−2.95***
	(0.42)	(0.42)
Committee Chair	0.13	0.10
	(0.09)	(0.09)
Positive Emotion Words	−0.78***	−0.79***
	(0.13)	(0.13)
Negative Emotion Words	−0.44***	−0.40***
	(0.13)	(0.13)
N	4,037	4,192
Log Likelihood	−1,917.47	−1,989.47
AIC	3,896.94	4,040.94

*p < .1; **p < .05; ***p < .01

0.51 to 0.55. A similar result is found for the average number of positive words which decreases the probability of a party vote from 0.61 to 0.39 when it is allowed to vary from its minimum (-0.59) to maximum (0.97). Collectively, this suggests not only is vocal pitch is a statistically significant predictor of party votes, but its effect is sizable relative to other variables.

While there are a number of reasons why party polarization makes policy-making more difficult, one of the main concerns is legislative gridlock. Beginning with Mayhew (1991), scholars have been concerned that party politics may prevent Congress from passing any legislation, let alone legislation that addresses the important issues of the day. These studies initially focused on whether legislative productivity was higher in a united versus divided government (Edwards III, Barrett, and Peake, 1997; Howell et al., 2000; Kelly, 1993). In order to pass legislation, the constitution requires agreement among the three branches of government. When the same party controls the House of Representatives, Senate, and Presidency, it is easier to overcome this hurdle since agreement between the three branches is more likely. Unfortunately, empirical support for the divided government hypothesis has been mixed, which is why scholars have increasingly considered whether party polarization makes legislation more difficult to passed (Binder, 1999; Binder, 2004; Jones, 2001). When political parties are polarized, not only are they more internally cohesive, but they are also diametrically opposed to the opposition. In this environment, strong incentives exist for partisans to distinguish their records and positions from the opposition, decreasing the incentive to cut legislative deals. If compromises are necessary to get things done, then any indication of increased polarization, non-verbal or otherwise, should also be an indication of legislative gridlock.

Table 5 shows vocal pitch is a strong predictor of whether a vote passed. Not only is the coefficient positive and statistically significant ($p < 0.01$), but when predicted probabilities are calculated holding all other variables at their mean and modal values the substantive effect again is sizable. More specifically, when the average vocal pitch for the day of the roll-call vote is allowed to vary from its minimum (-2.38) to maximum (2.43) the probability of the vote passing decreases from 0.77 to 0.36, suggesting heightened vocal pitch is not only indicative of a polarized environment, but also an environment in which legislative gridlock is more likely. When it comes to whether the vote passed, speaking rate is not a statistically significant predictor. Here, when the average speaking rate is allowed to vary from its minimum (-0.73) to maximum (1.07), the probability of the vote passes decreases from 0.60 to 0.58, suggesting that speaking rate is less indicative of legislative gridlock. Again, to put these results into perspective, when the average number of negative emotion words is allowed to vary from its minimum (-0.64) to maximum (0.82), the probability of the vote passing only decreases from 0.64 to 0.49. The average number of positive words seem to have a more pronounced effect. Here, the probability of the vote passing increases from 0.31 to 0.90 when the average number of positive words are allowed to vary from the minimum (-0.59) to maximum (0.97). Collectively, this suggests not only is vocal pitch is a statistically significant predictor of party votes, but its effect is sizable relative to other variables.

Discussion

Do members of Congress remain "on message" when speaking about partisan issues? Non-verbal expressions are one way party members can achieve this end. Using the audio from 74,158 floor speeches, this study has demonstrated the following:

1. Party members are more likely to raise their vocal pitch when speaking about their party's issues.

2. On days in which members of Congress on average raise their vocal pitch, party polarization is more likely. More specifically, when vocal pitch increases from its minimum to maximum value

the probability of a party vote increases from 0.21 to 0.82, holding all other variables constant at their mean and modal values.

3. On days in which members of Congress on average raise their vocal pitch, legislative gridlock is more likely. More specifically, when vocal pitch increases from its minimum to maximum value the probability of vote passing decreases from 0.77 to 0.36, holding all other variables constant at their mean and modal values.

In each instance, these effects are highly significant even when text-based measures designed to capture emotional expression are included as controls.

This study does not make any claims to whether emotion causes polarization or the other way around. From previous literature, we know that vocal pitch is more likely to increase when one is emotionally activated and that speakers are typically unaware that such changes take place. Since vocal pitch is correlated with speaking rate, it is unlikely members of Congress are intentionally increasing their vocal pitch *and* speaking rate to achieve some political objective. Although it is impossible to conclusively say whether the non-verbal displays studied in this paper are strategically motivated, each variable is scaled to standard deviations above or below the speaker's baseline, meaning the speaker would not only have to intentionally change their vocal pitch and speaking rate, but they would have to do so knowing their own baseline. Members of Congress are certainly strategic, but it is unlikely they are making such calculations, even though this is, at best, an educated guess. Similarly, party leaders may have an incentive to ask members to advance a certain "tone," but it is unclear whether this translates to asking party members to change their non-verbal behavior. It is more likely that members of Congress generally care about the issues they are advancing and their level of emotional commitment is reflected in their vocal pitch.

Even though this study cannot definitively isolate the underlying causal mechanism, it still makes four sizable contributions to the study of floor speeches. First, this study is the first to demonstrate the importance of understanding non-verbal behavior on Capitol Hill. Morris (2001) argues members of Congress are increasingly using floor speeches to create their own "legislative sound bite[s]" (115). This study suggests we should begin to think of these "sound bites" as more than words on a page. Whether one is talking about hand gestures or vocal inflections, non-verbal behavior is an important part of what transpires on the House floor. Members of Congress express commitment in a variety of ways. Sometimes they say, "I oppose this bill" and other times they say "I OPPOSE THIS BILL!" This study offers vocal cues as a way to differentiate between these two types of expressions.

Second, in some instances non-verbal behavior may lead to different inferences. For example, even though Republicans tended to talk at a higher vocal pitch when discussing security issues, Democrats tended to dedicate more of their speech time to talking about security. If one were to simply look at the number of security speeches, then one would wrongly conclude that Democrats paid more attention security issues than Republicans. If one were to look at vocal pitch, not only would one conclude differently, but one's conclusion would be consistent with well-established theoretical expectations. It is more difficult for speakers to manipulate their non-verbal behavior, meaning cues such as vocal pitch are going to be more honest indicators of where speakers stand on an issue. In this instance, Democrats may attempt to seize a Republican issue by delivering more speeches, but their vocal pitch may suggest their commitment to the issue is waning. Either way, vocal pitch carries unique weight, suggesting it is not synonymous with text-based measures.

Third, a large percentage of the House floor is unexplored. One reason for this has to do with the structured nature of floor debates. Outside of one-minute speeches, members of Congress are constrained to talking about certain topics when speaking on the House floor, making the former a more interesting area of research. Since non-verbal cues are more difficult to control, these cues are also unstructured, making them particularly useful for understanding position taking on the House floor. Members of Congress can speak about whatever issue they want during one-minute speeches. Even though what issues they talk about may be constrained, in these more structured speeches *how* they talk about those issues is entirely up to the speaker. Not only does this mean non-verbal behavior opens up a large portion of the House floor, but it also means that much of what we have learned about one-minute speeches can be applied to cues such as vocal pitch, advancing previous research.

Finally, there is a broader discussion about whether emotion is good or bad for democracy. For example, Matthews (1959) and others (e.g., Darr, 2005; Dodd and Schraufnagel, 2012; Jamieson and Falk, 2001; Schraufnagel, 2005; Uslaner, 1993) argue incivility actually prevents members of Congress from being productive in the legislature. Indeed, "[p]ersonal attacks, unnecessary unpleasantness, pursuing a line of thought or action that might embarrass a colleague needlessly, are all thought to be self-defeating–'after all, your enemies on one issue may be your friends on the next'" (Matthews, 1959, p. 1070). Even though raising the tone of one's voice is not the same thing as being uncivil, this study suggests emotions are expressed both verbally and non-verbally, meaning something as simple as vocal pitch can add considerably to our understanding of all emotional expressions, including those that are uncivil.

Floor speeches are more than words on a page. Even though previous scholars have demonstrated the importance of *what* is said on the House floor, this study is the first to suggest *how* those floor speeches are delivered is of equal, if not greater, import. In doing so, not only does this advance our understanding of floor behavior, but it lays an important foundation for future research. Whether one is talking about floor debates, oral arguments, or presidential addresses, “sound bites” are spoken, not written. Future work is needed to fully understand the importance of elite-non-verbal behavior, but this study takes an important first step.

References

Apple, William, Lynn A. Streeter, and Robert M. Krauss (1979). “Effects of Pitch and Speech Rate on Personal Attributions”. In: *Journal of Personality and Social Psychology* 37.5, pp. 715–727.

Bachorowski, Jo-Anne and Michael J. Owren (2003). “Sounds of Emotion: The Production and Perception of Affect-Related Vocal Acoustics”. In: *Annals of the New York Academy of Sciences* 1000, pp. 244–265.

Banse, Rainer and Klaus R. Scherer (1996). “Acoustic Profiles in Vocal Emotion Expression”. In: *Journal of Personality and Social Psychology* 70.3, pp. 614–636.

Binder, Sarah A. (1999). “The Dynamics of Legislative Gridlock, 1947–96”. In: *American Political Science Review* 93.3, pp. 519–533.

— (2004). *Stalemate: Causes and Consequences of Legislative Gridlock*. Brookings Institution Press.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003). “Latent Dirichlet Allocation”. In: *Journal of Machine Learning Research* 3, pp. 993–1022.

Breitenstein, Caterina, Diana van Lancker, and Irene Daum (2001). “The Contribution of Speech Rate and Pitch Variation to the Perception of Vocal Emotions in a German and an American Sample”. In: *Cognition & Emotion* 15.1, pp. 57–79.

Calvo, Rafael A. and Sidney D’Mello (2010). “Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications”. In: *IEEE Transactions on Affective Computing* 1.1, pp. 18–37.

Cowie, R. et al. (2001). “Emotion Recognition in Human-Computer Interaction”. In: *IEEE Signal Processing Magazine* 18.1, pp. 32–80.

Darr, Christopher R. (2005). “Civility as Rhetorical Enactment: The John Ashcroft “Debates” and Burke’s Theory of Form”. In: *Souther Journal of Communication* 70.4, pp. 316–328.

Davitz, Joel R. (1964). “Auditory Correlates of Vocal Expression of Emotional Feeling”. In: *The Communication of Emotional Meaning*. Ed. by Joel R. Davitz. New York, NY: McGraw-Hill, pp. 101–112.

Debruyne, Frans et al. (2002). “Speaking Fundamental Frequency in Monozygotic and Dizygotic Twins”. In: *Journal of Voice* 16.4, pp. 466–471.

Dietrich, Bryce J. et al. (2016). “Racial Rhetoric in Black and White: Symbolic and Substantive References in U.S. House Speeches”.

DiMaggio, Paul, Manish Nag, and David Blei (2013). "Exploiting Affinities Between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding". In: *Poetics* 41.6, pp. 570–606.

Dodd, Lawrence C. and Scot Schraufnagel (2012). "Congress and the Polarity Paradox: Party Polarization, Member Incivility and Enactment of Landmark Legislation, 1891-1994". In: *Congress and the Presidency* 39.2, pp. 109–132.

Edwards III, George C., Andrew Barrett, and Jeffrey Peake (1997). "The legislative impact of divided government". In: *American Journal of Political Science* 41.2, pp. 545–563.

Farrell, Justin (2016). "Corporate Funding and Ideological Polarization about Climate Change". In: *Proceedings of the National Academy of Sciences* 113.1, pp. 92–97.

Fitch, W. Tecumseh and Jay Giedd (1999). "Morphology and Development of the Human Vocal Tract: A Study Using Magnetic Resonance Imaging". In: *Journal of the Acoustical Society of America* 106.3, pp. 1511–1522.

Friedmann, Erika et al. (1982). "The Effects of Normal and Rapid Speech on Blood Pressure". In: *Psychosomatic Medicine* 44.6, pp. 545–553.

Genovese, Federica (2015). "Politics Ex Cathedra: Religious Authority and the Pope in Modern International Relations". In: *Research & Politics* 2.4, p. 2053168015612808.

Gerrity, Jessica C., Tracy Osborn, and Jeanette Morehouse Mendez (2007). "Women and Representation: A Different View of the District?" In: *Politics & Gender* 3.2, pp. 179–200.

Hall, Richard L. (1998). *Participation in Congress*. New Haven, CT: Yale University Press.

Hayes, Matthew and Bryce J. Dietrich (2016). "Race, Deliberation, and Voter Identification Laws".

Howell, William et al. (2000). "Divided Government and the Legislative Productivity of Congress, 1945–94". In: *Legislative Studies Quarterly* 25.2, pp. 285–312.

Jamieson, Kathleen Hall and Eriak Falk (2001). *Civility in the House of Representatives: The 106th Congress*. Tech. rep. The Annenberg Public Policy Center.

Jones, David R. (2001). "Party Polarization and Legislative Gridlock". In: *Political Research Quarterly* 54.1, pp. 125–141.

Keele, Luke John (2008). *Semiparametric Regression for the Social Sciences*. John Wiley & Sons.

Kehrein, Roland (2002). "The Prosody of Authentic Emotions". In: *Speech Prosody 2002, International Conference*.

Kelly, Sean Q. (1993). “Divided We Govern? A Reassessment”. In: *Polity* 25.3, pp. 475–484.

Laukka, Petri, Patrik Juslin, and Roberto Bresin (2005). “A Dimensional approach to vocal expression of emotion”. In: *Cognition & Emotion* 19.5, pp. 633–653.

Maltzman, Forrest and Lee Sigelman (1996). “The Politics of Talk: Unconstrained Floor Time in the U.S. House of Representatives”. In: *The Journal of Politics* 58.3, pp. 819–830.

Matthews, Donald R. (1959). “The Folkways of the United States Senate: Conformity to Group Norms and Legislative Effectiveness”. In: *The American Political Science Review* 53.4, pp. 1064–1089.

Mayhew, David R. (1991). *Divided We Govern*. New Haven, CT: Yale University Press.

Morris, Jonathan S. (2001). “Reexamining the Politics of Talk: Partisan Rhetoric in the 104th House”. In: *Legislative Studies Quarterly* 26.1, pp. 101–121.

Osborn, Tracy and Jeanette Morehouse Mendez (2010). “Speaking as Women: Women and Floor Speeches in the Senate”. In: *Journal of Women, Politics & Policy* 31.1, pp. 1–21.

Owens, Ryan J. and Justin P. Wedeking (2011). “Justices and Legal Clarity: Analyzing the Complexity of U.S. Supreme Court Opinions”. In: *Law & Society Review* 45.4, pp. 1027–1061.

— (2012). “Predicting Drift on Politically Insulated Institutions: A Study of Ideological Drift on the United States Supreme Court”. In: *The Journal of Politics* 74.2, pp. 487–500.

Owren, Michael J. and Jo-Anne Bachorowski (2007). “Measuring Emotion-Related Vocal Acoustics”. In: *Handbook of Emotion Elicitation and Assessment*. Ed. by James A. Coan and John J. B. Allen. New York, NY: Oxford University Press, pp. 239–265.

Pearson, Kathryn and Logan Dancey (2011a). “Elevating Women?s Voices in Congress Speech Participation in the House of Representatives”. In: *Political Research Quarterly* 64.4, pp. 910–923.

— (2011b). “Speaking for the Underrepresented in the House of Representatives: Voicing women’s Interests in a Partisan Era”. In: *Politics & Gender* 7.4, pp. 493–519.

Petrocik, John R. (1996). “Issue Ownership in Presidential Elections, with a 1980 Case Study”. In: *American Journal of Political Science* 40.3, pp. 825–850.

Poole, Keith T. and Howard Rosenthal (2001). “D-Nominate after 10 Years: A Comparative Update to Congress”. In: *Legislative Studies Quarterly*, pp. 5–29.

Popkin, Samuel L. (1994). *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago, IL: University of Chicago Press.

Przybyla, Beata D., Yoshiyuki Horii, and Michael H. Crawford (1992). “Vocal Fundamental Frequency in a Twin Sample: Looking for a Genetic Effect”. In: *Journal of Voice* 6.3, pp. 261–266.

Puts, David Andrew, Steven J.C. Gaulin, and Katherine Verdolini (2006). “Dominance and the Evolution of Sexual Dimorphism in Human Voice Pitch”. In: *Evolution and Human Behavior* 27, pp. 283–296.

Roberts, Margaret E., Brandon M. Stewart, and Edoardo M. Airoldi (2016). “A Model of Text for Experimentation in the Social Sciences”. In: *Journal of the American Statistical Association*, pp. 1–49.

Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley (2014). “stm: R package for structural topic models”. In: *R Package* 1, pp. 1–12.

Roberts, Margaret E. et al. (2013). “The Structural Topic Model and Applied Social Science”. In: *Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation*.

Roberts, Margaret E. et al. (2014). “Structural Topic Models for Open-Ended Survey Responses”. In: *American Journal of Political Science* 58.4, pp. 1064–1082.

Scherer, Klaus R. and James S. Oshinsky (1977). “Cue Utilization in Emotion Attribution from Auditory Stimuli”. In: *Motivation and Emotion* 1.4, pp. 331–346.

Schraufnagel, Scot (2005). “Testing the Implications of Incivility in the United States Congress, 1977-2000: The Case of Judicial Confirmation Delay”. In: *The Journal of Legislative Studies* 11.2, pp. 216–234.

Siegman, Aron Wolfe, Robert A. Anderson, and Tal Berger (1990). “The Angry Voice: Its Effects on the Experience of Anger and Cardiovascular Reactivity”. In: *Psychosomatic Medicine* 52.6, pp. 631–643.

Tausczik, Yla R. and James W. Pennebaker (2010). “The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods”. In: *Journal of Language and Social Psychology* 29.1, pp. 24–54.

Titze, Ingo R. (2000). *Principles of Voice Production*. Iowa City, IA: National Center for Voice and Speech.

Uslaner, Eric M. (1993). *The Decline of Comity in Congress*. Ann Arbor, MI: University of Michigan Press.

Zeng, Zhihong et al. (2009). “A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31.1, pp. 39–58.