

How Audio, Image, and Video Data is Being Used to Study Legislative Behavior  
By  
Bryce J. Dietrich and Jielu Yao

Since perhaps Laver, Benoit, and Garry (2003), text data has been used to understand legislative politics. While studies like Proksch and Slapin (2012), Quinn et al. (2010), and Pearson and Dancey (2011) have added considerably to our understanding of legislative politics, a new line of research has begun to emerge in which audio, image, and video data are being used to understand a wide range of political behavior. This article highlights some of this literature and explains how future scholars could potentially apply these techniques to their own research.

To date, audio data has been used primarily to understand vote choice (e.g., Tigue et al., 2012) and emotional expression in elite settings (e.g., Dietrich, Enos, and Sen, 2019). For example, using 796 *YouTube* videos of the top two candidates from each House race in the 2012 midterm election, Klofstad (2016) found those with higher pitched voices were more likely to lose, although these effects became somewhat diminished once other electoral variables were added as controls. In more recent work, Klofstad and Anderson (2019) found little correlation between changes in vocal pitch and legislative effectiveness, a finding the authors find puzzling and suggest may be due to their measure of legislative effectiveness. Regardless, this work and other by this research team (e.g., Klofstad, Anderson, and Peters, 2012) clearly demonstrate how audio data can be used to better understand the U.S. Congress.

With that said, this work asks a very specific question about how non-verbal expressions, like changes in vocal pitch, can influence legislative behavior. More specifically, the experimental treatments in these studies tend to use manipulated audio in which the vocal pitch of the speaker is increased or decreased using software. If one were interested in knowing whether voters generally prefer higher or lower pitched voices, then this approach is completely reasonable. However, members of Congress (MCs) are exposed repeatedly to their colleagues' voices, so while they may initially react negatively to a speaker with a higher pitched voice, over time they likely become used to their colleagues' tone. After this point, deviations from this learned baseline is likely what is most important since such changes may signal whether their colleague is expressing a more or less intense expression.

In a recent article published in the *American Political Science Review*, this is precisely the argument advanced by Dietrich, Hayes, and O'Brien (2019). Instead of thinking about the basal reaction MCs may have to vocal pitch, these researchers explore how changes in vocal pitch may be indicative of underlying attitude preferences. Although certainly not the first to suggest female MCs are more likely to speak about women and the issues they care about on the House floor (e.g., Osborn and Mendez, 2010), this study uses the audio from over 74,000 floor speeches to demonstrate that such expressions are not simply "cheap talk," but instead are grounded in a deep attachment to women as a group. Ultimately, the authors show that when women speak about women with more intensity – as indicated by increased vocal pitch – their male colleagues are not only more likely to mention women, but they are also more likely to

vote with them. While the authors are quick to point out the difficulty of establishing a causal relationship using observational data, this study will undoubtedly serve as an important foundation for future scholars interested in using audio data in their own research, legislative or otherwise.

Despite the relatively widespread use of audio data in legislative studies, less work has been done which uses images or video in a similar way. One noticeable exception is Xi et al. (2019) who — similar to Yao et al. (2019) — applied deep learning techniques to Facebook photographs of 319 Members of Congress. Their classifier — which used transfer learning and the ResNet-34 architecture — achieved an accuracy of 82.35% when multiple images were used. Ultimately, they found conservative politicians are more likely to use image objects and people to advocate military and status quo political institutions while liberals are more likely to use images to show their concerns about the global community and environment. Xi et al. (2019) also found conservatives are more likely to express happiness in their photos, a result which is the opposite of what was previously reported by Wojcik et al. (2015). Given previous work in political communication which underlines the importance of facial expressions (e.g., Stewart, Walter, and Schubert, 2009; Grabe and Bucy, 2009), it is only a matter of time until scholars begin to use these results to better understand legislative behavior.

Perhaps the most direct application of image- or video-as-data is a forthcoming paper in *Political Analysis* titled “Using Motion Detection to Understand Social Polarization in the U.S. House of Representatives.” In this paper, Dietrich (2019) uses over 6,000 hours of C-SPAN videos to measure the extent to which Democrats and Republicans mingle after floor votes. Not only does he find these types of bipartisan social interactions have been steadily declining since 1997, but he also shows that they may have important downstream consequences. More specifically, party votes are found to be more likely after videos in which Democrats and Republicans do not literally cross the aisle. Similar to Dietrich, Hayes, and O’Brien (2019), the author emphasizes the difficulty of establishing a firm causal relationship using observational data, but he finds more evidence that the lack of bipartisan social interactions is “predictive of party voting and not the other way around” (12).

The use of audio, image, and video data will become increasingly important to the study of legislatures both inside and outside of the United States, especially as the methodology needed to understand the dynamics in these non-traditional data sources continues to develop. Casas and Webb Williams (2019), Knox and Lucas (2019), Neumann (2019), Rheault, Ludovic, and Sophie Borwein (2019), and Torres (2019) will be important future works in this regard. These studies will likely motivate a new wave of research in which non-traditional data sources are used in combination with machine learning to better understand legislative politics and hopefully this article will help scholars better navigate what may seem like unfamiliar waters.

## References

Casas, Andreu, and Nora Webb Williams. 2019. "Images that matter: Online protests and the mobilizing role of pictures." *Political Research Quarterly* 72(2): 360-375.

Dietrich, Bryce J., Ryan D. Enos, and Maya Sen. 2019. "Emotional arousal predicts voting on the US supreme court." *Political Analysis* 27(2): 237-243.

Grabe, Maria Elizabeth, and Erik Page Bucy. *Image bite politics: News and the visual framing of elections*. Oxford University Press, 2009.

Laver, Michael, Kenneth Benoit, and John Garry. 2003. "Extracting policy positions from political texts using words as data." *American Political Science Review* 97(2): 311-331.

Klofstad, Casey A. 2016. "Candidate voice pitch influences election outcomes." *Political Psychology* 37(5): 725-738.

Klofstad, Casey A., and Rindy C. Anderson. 2018. "Voice pitch predicts electability, but does not signal leadership ability." *Evolution and Human Behavior* 39(3): 349-354.

Klofstad, Casey A., Rindy C. Anderson, and Susan Peters. 2012. "Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women." *Proceedings of the Royal Society B: Biological Sciences* 279(1738): 2698-2704.

Knox, Dean and Chris Lucas. 2019. "A dynamic model of speech for the social sciences." Working Paper.

Neumann, Markus. 2019. "Hooked with Phonetics: The Strategic Use of Style-Shifting in Political Rhetoric." Working Paper.

Osborn, Tracy, and Jeanette Morehouse Mendez. 2010. "Speaking as women: Women and floor speeches in the Senate." *Journal of Women, Politics & Policy* 31(1): 1-21.

Proksch, Sven-Oliver, and Jonathan B. Slapin. 2012. "Institutional foundations of legislative speech." *American Journal of Political Science* 56(3): 520-537.

Quinn, Kevin M., et al. 2010. "How to analyze political attention with minimal assumptions and costs." *American Journal of Political Science* 54(1): 209-228.

Rheault, Ludovic, and Sophie Borwein. 2019. "Multimodal Techniques for the Study of Affect in Political Videos." Working Paper.

Stewart, Patrick A., Bridget M. Waller, and James N. Schubert. 2009 "Presidential speechmaking style: Emotional response to micro-expressions of facial affect." *Motivation and Emotion* 33(2): 125-135.

Tigue, Cara C., et al. 2012. "Voice pitch influences voting behavior." *Evolution and Human Behavior* 33(3): 210-216.

Torres, Michelle. 2019. "Give me the full picture: Using computer vision to understand visual frames and political communication." Working Paper.

Wojcik, Sean P., et al. 2015. "Conservatives report, but liberals display, greater happiness." *Science* 347(6227): 1243-1246.

Xi, Nan, et al. 2019. "Understanding the Political Ideology of Legislators from Social Media Images." arXiv preprint arXiv:1907.09594.

Yao, Jielu et al. 2019. "Do Women Candidates "Run as Women" Online? Applying and Validating Automated Textual and Visual Analyses to Facebook Campaign Advertising." Working Paper.