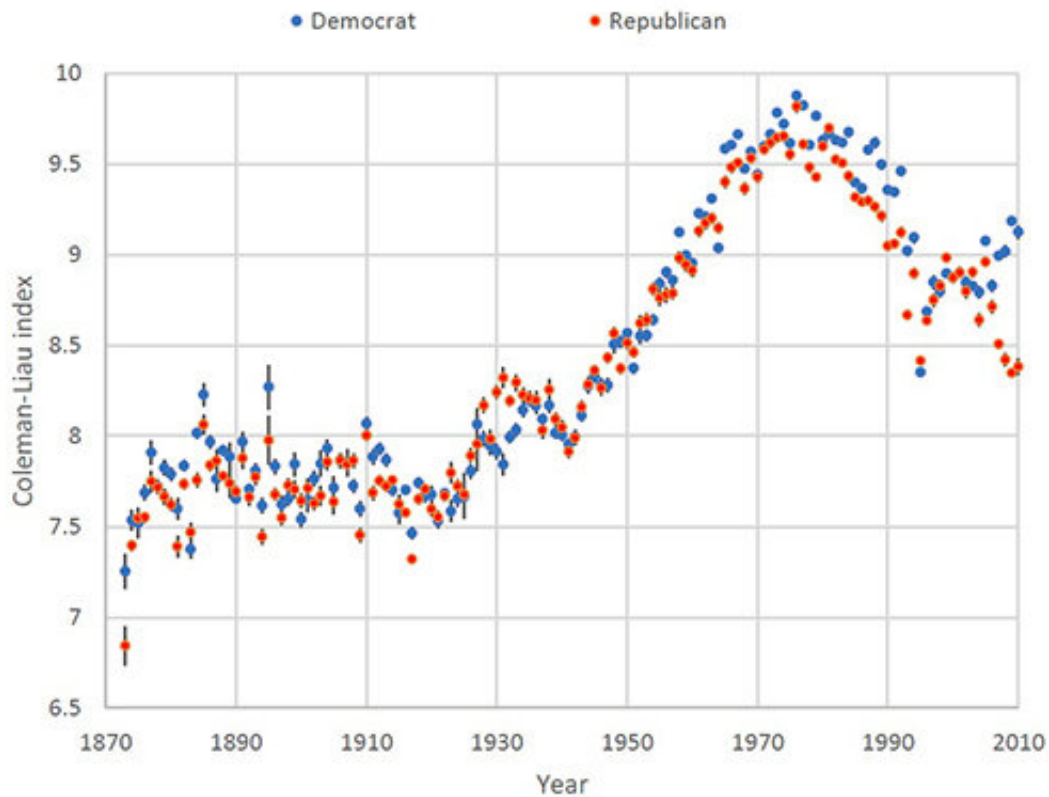


Analysis shows that political speeches now use simpler language, express more sentiments

August 18 2020



The Coleman-Liau index shows the reading level of Democratic and Republican congressional speeches in from 1873 to 2010. Credit: Kansas State University

Research by Kansas State University shows how politicians from both

major parties have changed their political speech from previous centuries.

A computer science research team at K-State analyzed nearly 2 million congressional speeches made by Republican and Democrat legislators from 1873 to 2010. Their computer analysis shows that [political speeches](#) are in fact very different in their style from political speeches made in Congress several decades ago.

In the [research paper](#) "A data science approach to 138 years of congressional speeches" published recently in the journal *Heliyon*, K-State computer science students Ethan Tucker and Colton Capps and computer science associate professor Lior Shamir used automatic text analysis algorithms to analyze congressional speeches in different years.

"The research results show that more recent speeches use a smaller vocabulary, simpler language, express more positive or negative sentiments, and have more noticeable differences between Democratic and Republican speakers," Shamir said.

The algorithms measured different aspects of the speeches such as the vocabulary, the [reading level](#), the positive or negative sentiments expressed in the speeches, and more. The sentiments are measured by using artificial intelligence reading of the text and associating words and phrases with positive or negative sentiments given their context.

"Based on that analysis, the algorithm determines whether a piece of text is positive, very positive, negative, very negative or neutral," Shamir said.

The algorithms also measured the frequency in which different topics were discussed. These quantitative [speech](#) elements were computed from thousands of congressional speeches made in each year, and the average

of each year allowed to measure the changes in the language and topics discussed in Congress during a period of 138 years, Shamir said.

The research showed that the frequency of words related to women's identity—such as she, her, hers, woman, women, etc.—has been increasing consistently since the early 1980s, while the frequency of words that identify men have been decreasing. The frequency of words related to women's identity in the 21st century is five times higher compared to the 1950s, but still lower than the frequency of words related to men's identity. Since the 1990s, terms related to women's identity are more frequent in speeches made by Democratic legislators compared to speeches made by Republican legislators.

"For most of the 20th century, however, there were no substantial differences between women's identity in Democratic and Republican speeches, and expressions of women's identity were about 10 times less frequent than expressions of men's identity by legislators from both parties," Shamir said.

The research also showed that the reading level of the speeches changed significantly over the years. The analysis measured the Coleman-Liau readability index, which estimates the reading level of a certain text and associates it with the appropriate school grade. The analysis showed that the reading level of congressional speeches made by both Republican and Democratic legislators increased consistently from the eighth-grade reading level in the 19th century, to the 10th-grade level in the 1970s. But since 1976 the reading level of political speeches has been declining consistently, and as of the 21st century, it is below the ninth-grade reading level. The same trend was also observed with the vocabulary used by congressional members in speeches, which had been increasing consistently until the early 1970s, and then started to decline—and it is still declining, Shamir said.

The researchers' analysis of the speeches also showed that more recent congressional speeches express more positive and negative sentiments than the speeches made in Congress during the 19th century and early 20th century. The sentiments in political speeches became gradually more positive and peaked in the 1960s, but declined sharply during the 1970s. Since the 1970s the sentiments expressed in congressional speeches have been becoming more positive.

According to the study, the decline in reading level and vocabulary of the speeches can be related to the increasing presence of media—including live radio and TV coverage—in Congress beginning in the 1970s. Members of Congress started to gradually adjust their speech styles, addressing the public through the media rather than addressing their fellow legislators.

Another aspect reflected through the analysis was the partisan split, Shamir said. Starting in the mid-1990s, Republican and Democratic speeches became increasingly different from each other and also correlated with the political affiliation of the president. For instance, during the George W. Bush administration, speeches of Democratic legislators expressed more negative sentiments compared to their Republican counterparts. That difference flipped immediately after 2008, with the beginning of the Obama administration, during which Republican speeches became more negative.

"With natural language processing we can extract new knowledge from old data," Shamir said. "There is no practical way to quantify and profile such a large number of speeches without using computers."

More information: Ethan C. Tucker et al, A data science approach to 138 years of congressional speeches, *Heliyon* (2020). [DOI: 10.1016/j.heliyon.2020.e04417](https://doi.org/10.1016/j.heliyon.2020.e04417)

Provided by Kansas State University

Citation: Analysis shows that political speeches now use simpler language, express more sentiments (2020, August 18) retrieved 25 April 2024 from <https://phys.org/news/2020-08-analysis-political-speeches-simpler-language.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.