PJIM PARSONS JOURNAL FOR INFORMATION MAPPING

THE PARSONS INSTITUTE FOR INFORMATION MAPPING
68 Fifth Avenue New York, NY 10011
212 229 6825 piim.newschool.edu

Congress Speaks

DINO CITRARO, BA
& KIM REES, BA

KEYWORDS *bifurcated mosaic plot, data visualization, Gov 2.0 speeches, information visualization, politics, strip treemap, tag cloud, tornado chart, U.S. Congress, voting records, word cloud*

ABSTRACT *Congress Speaks* is an entertaining, online exploration of the more than 14.5 million words spoken by the United States 110th Congress. Individual legislators can be compared according to what, and how often, they spoke, their congressional tenure, and their voting record. In addition, states can be visually compared by viewing: their party representation, party vocalization breakdown, and a state-by-state word count.

While exploring which topics are the most emphasized by individual legislators, interesting and unexpected words appear, such as: "wolves," "slavery," and "wilderness." These kinds of associations are also revealed while exploring the words used for individual states.

In addition to being educational the site is entertaining. Visitors can pit legislators against each other to see who is more verbose. Each legislator's mouth is animated and portrayed on a caricature of their political party's mascot: a donkey for Democrats, an elephant for Republicans, and in the case of Independents, an ostrich. Congress Speaks even includes crowd reactions, unexpected animations, confetti, and balloons.

INTRODUCTION *Congress Speaks* was developed as a public service. The data used in this piece, while publicly available, was difficult to access and rarely presented in an engaging manner. As an information visualization firm we knew that people would benefit from the information if it were presented in a more intuitive and entertaining manner. We developed Congress Speaks to be an example of how data visualization can be used to engage the public and inspire greater civic involvement.

OVERVIEW AND INITIAL INTERACTION

On the surface, the visualization is an exploration of the words spoken during the United States 110th Congress. It is through the question, "Who spoke the most?" that we created the pathway by which visitors explore not only the volume of words spoken, but also the associated information that comes as a result of these words. Associative information includes the total words that all congress members spoke from a particular state and how that number compares to other states. In addition, supporting information is introduced to provide further context for the words, including the most frequently used words for each legislator, how those words compare to the entire congress, how the legislator voted on key issues, and their tenure.

To provide a mechanism for exploration of the words spoken by each legislator, we created an interface that reveals a public speaking, debate-like, metaphor. When first accessing the tool visitors see the map of the United States with word balloons emerging from each state enticing and prompting the visitor to choose a state. The interface also displays two microphones, one on each side of the screen; these provide clues to the visitor that they should explore the states and then select individuals to stand at the respective microphones. While each Senator and Representative is listed within the manifest of each state, the entire state is also selectable and can be chosen for comparison (FIGURE 1).

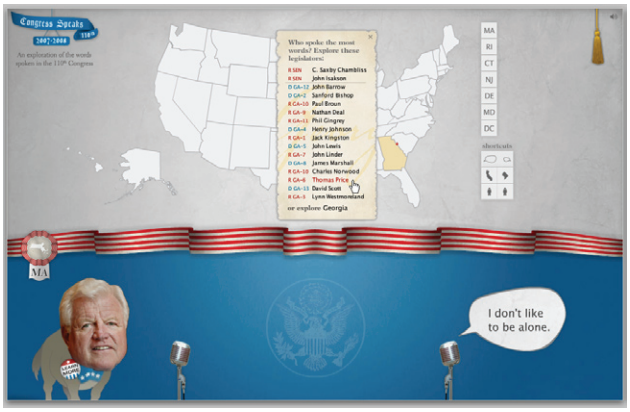

FIGURE 1: *Choosing legislators for comparison*

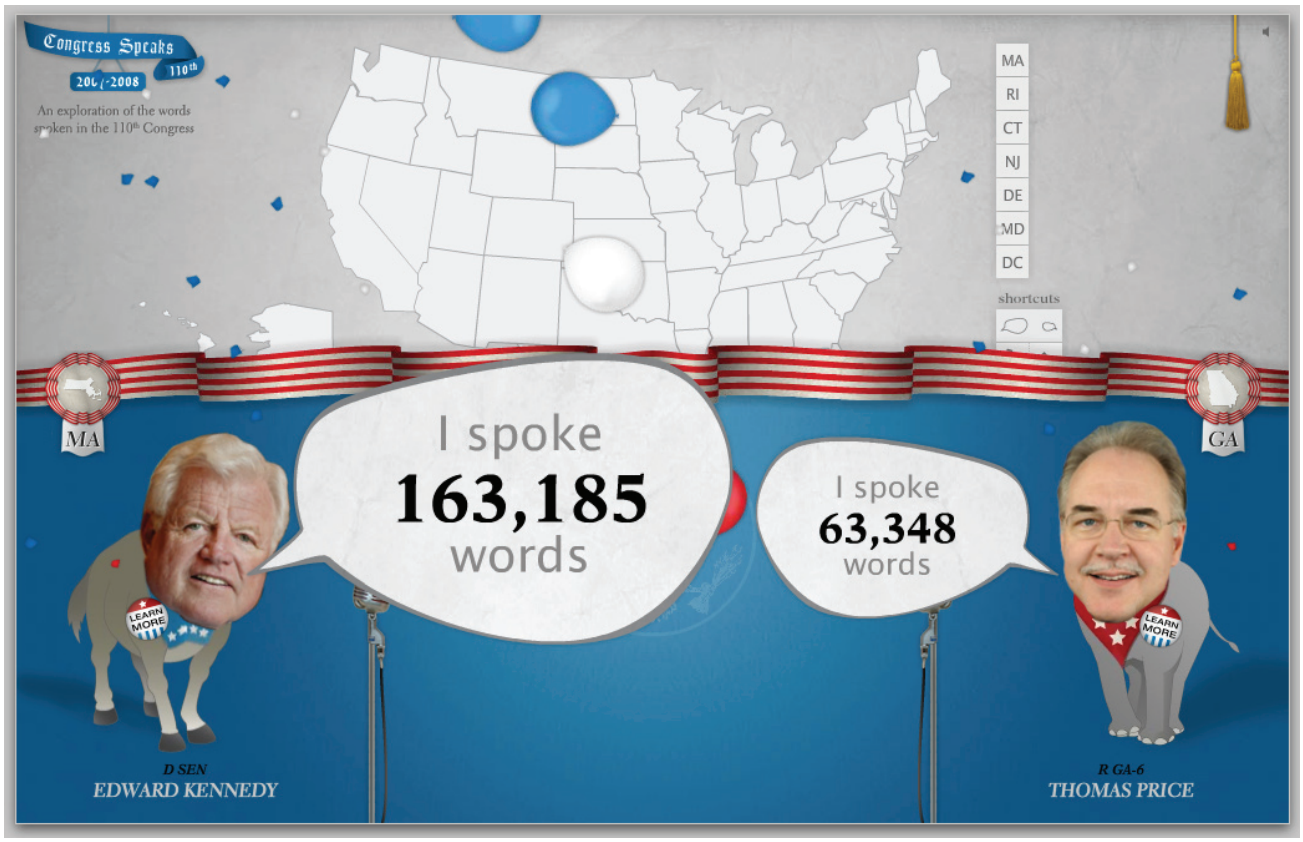

FIGURE 2: *Two legislators selected for comparison*

After two legislators have been selected, each appears at the microphone and begins talking. The resulting word balloons show the number of words spoken both as a numerical value, and by the overall size of the balloons. Legislators are color-coded when listed in each state manifest, and assume a character body of their political party's mascot when standing at the microphone. Additionally, each wears a bandana around their neck that denotes the color of their political party (FIGURE 2).

Once the question "Who was doing all that talking?" is answered, a debate moderator appears and prompts us to further investigate the information. We took this opportunity to subtly show protocol and hierarchy within the Senate and House of Representatives by using the first Vice President, John Adams, as the moderator for comparisons between Senators, and the first Speaker of the House, Fredrick Augustus Muhlenberg, as the moderator for comparisons between Representatives.

COMPARING LEGISLATORS

Our information mapping efforts were primarily concentrated on the details of the congressional members. In this view, we have mapped information under five categories, each focusing on various aspects of the individual legislator (FIGURE 3).

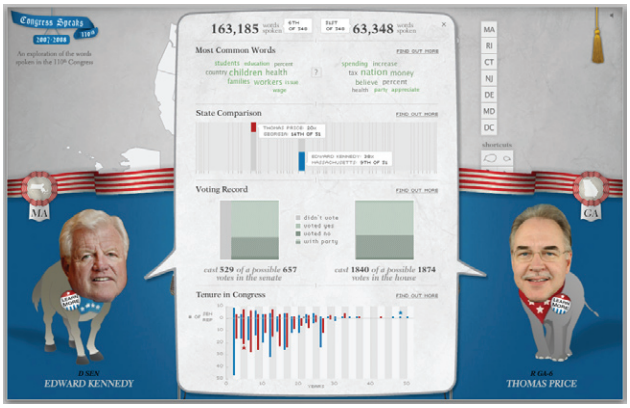

FIGURE 3: *Legislator comparison details*

WORD VOLUME

We initiated the design of the detail screen by reiterating the number of words spoken by each individual. To contextualize, we show how that number ranks among all 548 members of the 110th Congress.

FIGURE 4: *Legislator comparison details*

Here we map the ten most common words for each legislator. While we use traditional word clouds such as Wordle[1] and Flickr[2] as our inspiration, we provide additional information within our visualization. We use size to denote how often the word was used and color to symbolize uniqueness. Words that are greener in color indicate that these words are spoken more by this individual than by other members of congress. In FIGURE 5, for example, Rep. Price said the word "appreciate" more often than other legislators.

Additionally, a rollover provides details about each word. FIGURE 5 shows that Senator Kennedy used the word "wage" 607 times; his usage was over ten percent of the overall congressional usage.

In a side-by-side comparison, this view of word-usage can provide insight. For instance, FIGURE 5 shows that Kennedy was concerned about social issues such as education, health, and labor, while Price was focused primarily on national finance issues such as spending and taxes.

WORD COMMONALITY

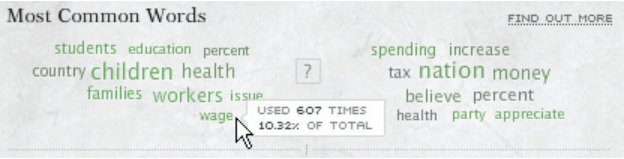

FIGURE 5: *Charting the most common words*

This category expands on the legislator's total word count and provides contextual information about how their specific state compares to states nationwide. Each state is displayed as a vertical bar whose thickness is proportionate to its word count compared to the word count of the entire U.S. (State abbreviations are provided on rollover to minimize visual complexity.)

Legislators can be compared using various data points. Each legislator is displayed as a box within his or her state. The size of the bar indicates how much that individual spoke within their state; the percentage is also shown as a number in the tooltip. FIGURE 6 shows that Rep. Price provided twenty percent of the words for Georgia. Additionally, the color of the box denotes their political party: red for republican, blue for democrat, and white for independent.

The rank of verbosity for the state is also displayed in each legislator's tooltip. Only the states of the two compared legislators are listed in this way. In FIGURE 6, we see that Massachusetts is the ninth most loquacious in terms of its legislators.

We used a strip treemap[3] to indicate this, as it offered a way to compare this data within a constrained space. The strips represent the states and the nested rectangles are the legislators within each state. By limiting the nested boxes to appear only for the comparison of two individuals visual complexity is reduced and the user is allowed to focus on what is meaningful within the current context.

LEGISLATOR VOTING RECORDS

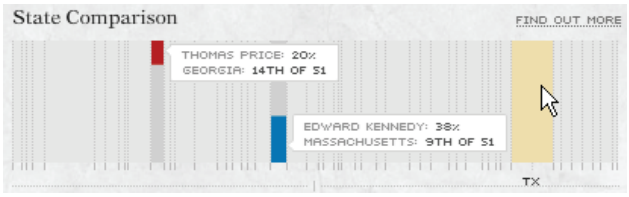

FIGURE 6: *State by state comparison*

Here we used a mosaic plot to chart how many votes the legislator cast, as well as how many were in support of bills and how many votes were in opposition. We then mapped how many of those votes were in line with their political party.

By applying a line pattern to the "with party" categorization, we were able to overlay that variable on top of the yes/no column. Since the party line voting variable is simply a subset of a bifurcated vote variable this method worked well to concisely articulate this data point.

To contextualize this information further, we noted in text how many voting opportunities were available, and for which chamber of congress they applied. Percentages for each value in the chart are displayed in a rollover tooltip. FIGURE 7 shows that Senator Kennedy voted "yes" 50.08% of the time on his 529 votes.

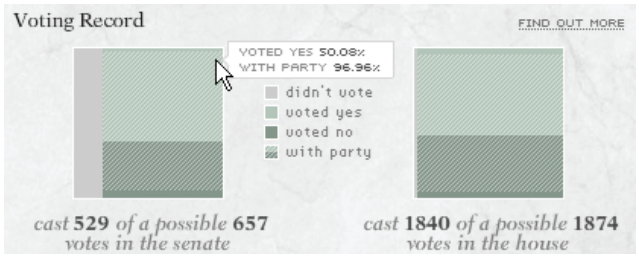

FIGURE 7: *Voting records graphs*

LEGISLATOR TENURE IN CONGRESS

Finally, to show how the legislator fit into the overall temporal landscape of congress members, we created a chart to communicate tenure. All 548 members of congress are represented in the chart.

This tornado bar chart is divided to show senators on the top and representatives on the bottom. Identical tenures are grouped together by political party and color-coded. The multiple series for each year is helped visually by using alternate column shading for each two-year term along the x-axis.

The x-axis represents years in congress. The height of each bar signifies how many legislators have been in office for that particular time period. For example, the tooltip in FIGURE 8 shows that 24 democrats (noted by the height) have served 26 years (denoted by the position on the x-axis) in congress as representatives.

Each of the legislators being compared is indicated with a star for their particular bar on the chart.

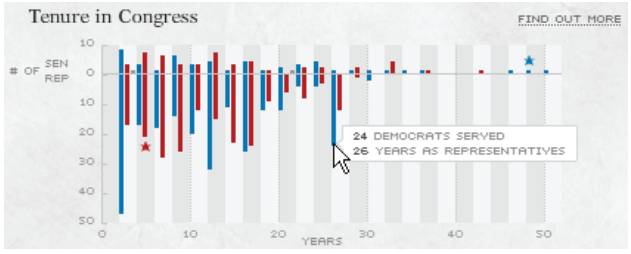

FIGURE 8: *Chart depicting tenure in congress*

COMPARING STATES

As noted earlier, visitors can choose to compare individual states in addition to comparing legislators. In this scenario we show party representation within the state instead of legislator tenure, and use a data graphic to show party affiliation denoted by color. The scale of the person icon represents the percentage of party legislators within the given state.

We also use a simple bar chart to denote the percentage of words spoken by each party. These too, are depicted by party color, and have a word balloon connector from the person icon to reinforce the relationship between the two.

This comparison sometimes reveals that party affiliation and vocal representation are disproportionate. For example, FIGURE 9 shows that 47% of the state on the left is Democrat, yet the vocal representation is a majority of 72%. The state on the right displays the opposite of this.

ACCENTING DATA

To provide further context for the legislators who comprised the 110th Congress we included a collection of demographic facts about the members, including: statistics about ethnic diversity, gender, religion, and age. Continuing the moderator metaphor established earlier, we used imagery of Dick Cheney, the presiding Vice President, and Nancy Pelosi, the presiding Speaker of the House as the presenters of this information.

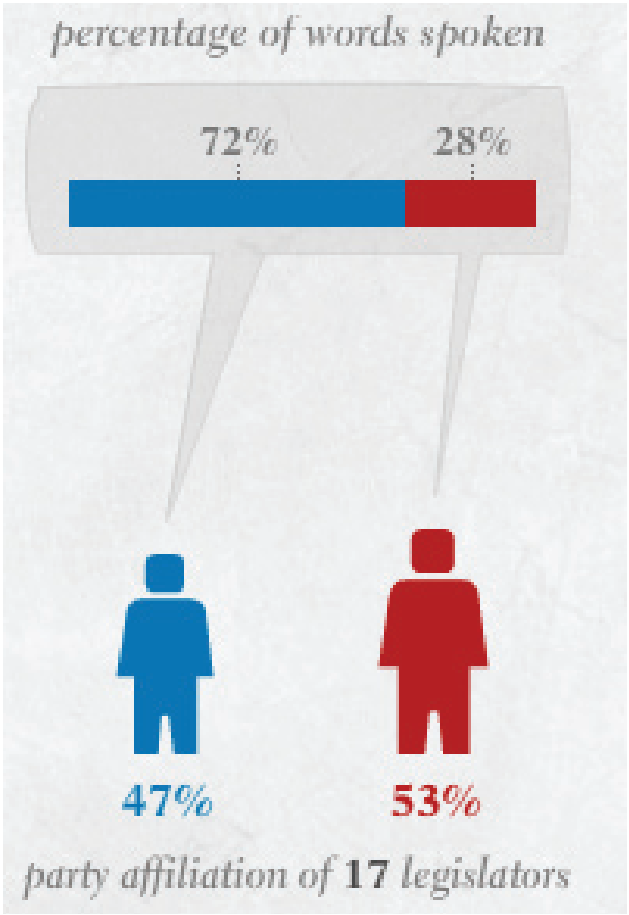

FIGURE 9: *Party representation by state*

CREDITS

All the data used in *Congress Speaks* is public domain information. We used the Capitol Words API,[4] *New York Times* developer network,[5] and information published in the *San Francisco Chronicle*.[6]

ACKNOWLEDGEMENTS

Additional members of Periscopic who contributed to this visualization are *Brett Johnson*, *Jacob O'Brien*, and *Zach Krausnick*.

BIOGRAPHY

Dino Citraro and *Kim Rees* are partners in Periscopic, an award-winning interactive design and development firm, specializing in user-centric design with a strong focus on information visualization. The company's work has appeared in several publications, including the *2009 Communication Arts Interactive Annual* and in the *Information Design Sourcebook*.

Citraro is a 15-year veteran of the multimedia industry. His work has spanned interactive motion pictures, multi-player online games, immersive data visualizations, and interactive hardware installations.

Rees has over fifteen years of experience in the interactive industry and is a prominent individual in the information visualization community. She has spoken at numerous events including the CERF Biennial Conference, Tableau Conference, Web Visions, and Portland Data Visualization among others.

NOTES

1 Jonathan Feinberg, "Wordle," http://www.wordle.net/

2 Flickr, "Popular Tags," http://www.flickr.com/photos/tags/

3 Ben Shneiderman, "Tree visualization with treemaps: a 2-D space-filling approach," *ACM Transactions on Graphics* 11, no. 1 (January 1992): 92–99.

4 Capitol Words, http://www.capitolwords.org/api/.

5 "Developer Network," *The New York Times*, http://developer.nytimes.com/.

6 The Associated Press, "Some Facts About Members of New Congress," *San Francisco Chronicle* (January 3, 2007), http://www.sfgate.com/cgi-bin/article.cgi?f=/n/a/2007/01/03/national/w170412S68.DTL&hw=muslim+congress&sn=001&sc=1000.