Semantic analysis of German parliament speeches

Emanuel Fuchs
emanuel-fuchs@t-online.de

Arthur Jaques
arthur.jaques@live.com

Abstract

This paper performs a semantic analysis of transcribed speeches in the German parliament. Latent Dirichlet Allocation and sentiment detection are used to provide information about temporal shifts of topic shares and the different emphases and sentiments of parties on the extracted topics.

1 Introduction

This paper describes a semantic analysis of transcribed speeches in the German parliament. We first extract speech topics and analyze their temporal evolution. We then show the credibility of the extraction by visualizing the mean party topics and showing that the picture drawn is coherent to an intuitive view of the parties' differences. Since speech topics do not explain the positions of the parties regarding them, we extract sentiments and analyze their correlation with topics and parties.

2 Methods

We download the texts of the Bundestag speeches from 24 October 2017 to 7 September 2021 using the OpenParliament [2] API. We discard speeches not attributed to any of the major parliament factions (*SPD*, *FDP*, *CDU/CSU*, *Bündnis 90/Die Grünen*, *AfD*, *Die Linke*). We remove formalities (such as salutations) using regular expressions. The resulting dataset contains 8,470 speeches [1]. 207 different dates are present, and all parties held speeches in at least 91% of them [2].

For topic extraction, we use Latent Dirichlet Allocation [3], trained on the whole dataset. The texts are stemmed [4] and vectorized using the bag-of-words approach with word counts. We analyze the words having the highest weights for each topic and name the topics accordingly. We optimize the parameters for vectorization and topic extraction to obtain interpretable results (a subjective criterion). This results in keeping only words that appear in at least 20 speeches and at most 30% of them for vectorization, and extracting 9 topics. To visualize the average party topic distributions, we reduce the data to 2 dimensions using Principal Component Analysis.

We extract sentiments using the package germansentiment, which applies the Bert architecture trained on German texts [1]. We use p-value testing with a binomial assumption for the distribution of negative sentiments to test for differences between parties in the proportion of negative speeches.

3 Results

[5] The topics extracted from the speeches dataset and the name that was assigned to them are summarized in Table 1.

[1] 2,101 *CDU/CSU*, 1,547 *SPD*, 1,285 *AfD*, 1,188 *Bündnis 90/Die Grünen*, 1,186 *FDP*, 1,163 *Die Linke*.

[2] 203 CDU, 195 *SPD*, 198 *AfD*, 193 *Bündnis 90/Die Grünen*, 190 *Die Linke*, 189 *FDP*.

[3] In Scikit-learn's implementation.

[4] Using nltk's snowball German stemmer.

[5] Code and analyses: `https://github.com/nolan1999/bundestag-speeches-analysis`.

Table 1: Extracted topics

#	Assigned name	Strongest predictors
1	International	europa; deutsch; russland; staat; gemeinsam; international; eu; menschenrecht; welt; turkei.
2	Military	soldat; einsatz; bundeswehr; mandat; soldatinn; mission; afghanistan; unterstutz; mali; militar.
3	EU/Economy	europa; euro; eu; unternehm; milliard; deutsch; prozent; union; geld; wirtschaft.
4	Social	kind; euro; famili; prozent; arbeit; sozial; 000; rent; miet; hoh.
5	Decisions/Law	gesetzentwurf; bundestag; glaub; fall; thema; hatt; word; entscheid; punkt; regel.
6	Democracy/Freedom	emokrati; leb; gesellschaft; freiheit; grundgesetz; staat; gewalt; deutsch; demokrat; wer.
7	German History	deutsch; wer; ost; geschicht; stiftung; abstimm; 19; opf; stimmt; bitt.
8	Ecology	klimaschutz; co; energi; prozent; erneuerbar; bau; ziel; verbrauch; energiew; wirtschaft.
9	Health/Pandemic	pandemi; schul; unternehm; arbeit; stark; digital; bass; bereich; massnahm; wirtschaft.

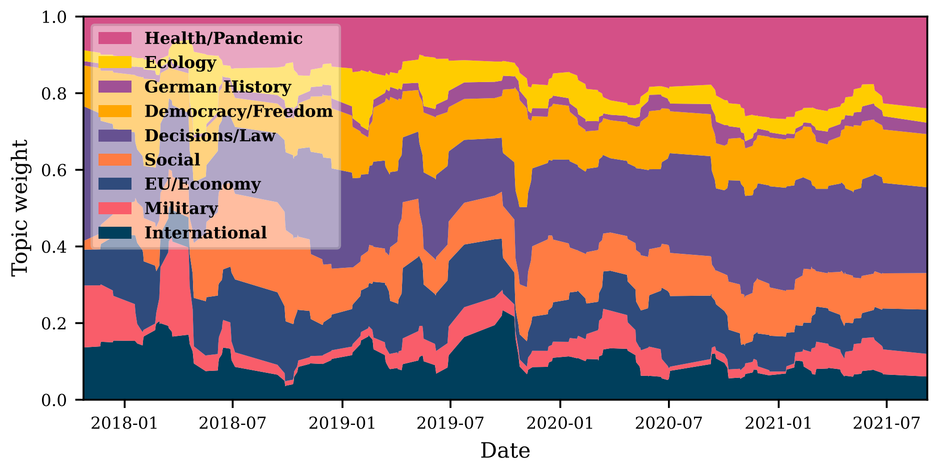

Figure 1: Average weight of topics discussed in the Bundestag (with Gaussian smoothing).

Topics evolution over time An overview of the evolution of topics in the speeches is shown in Figure 1. Notice the increased weights of both *Health/Pandemic* and *Decisions/Law* from the start of 2020 on, with the Covid 19 outbreak and consequent discussions about effective policy to tackle it. The historical evolution of the topic weights shows the focus of discussions in the parliament. An analysis of the *Social* topic (Figure 2) shows for example an increase in social speeches in the second half of 2018, which can be attributed to discussions about the statutory pension insurance and the so-called Participation Opportunities Act. The Strong Families Act passed at the beginning of 2019 may also have contributed to the rise of the topic [3][4].

Party topics differences In Figure 3 we show the results of 2-dimensional Principal Component Analysis on the average topic weights for each party. The results show the parties more or less sorted on the left-right axis along the second principal component. Furthermore, it shows the two parties that are considered more extremist (*Die Linke* and *AfD*) significantly distant from the other parties. We notice how the new *Ampel Koalition*, that wasn't governing at the time, was already closer together than the governing *Schwarz-rote Koalition* (topic similarity might indicate similar policy priorities).

Sentiment analysis The extraction of sentiments from the speeches yielded in total 7,293 neutral speeches, 1,179 negative speeches, and 13 positive speeches. Figure 4 shows the average sentiment by party and topic. The plot suggests more negative speeches of non-governing parties (with *FDP* as

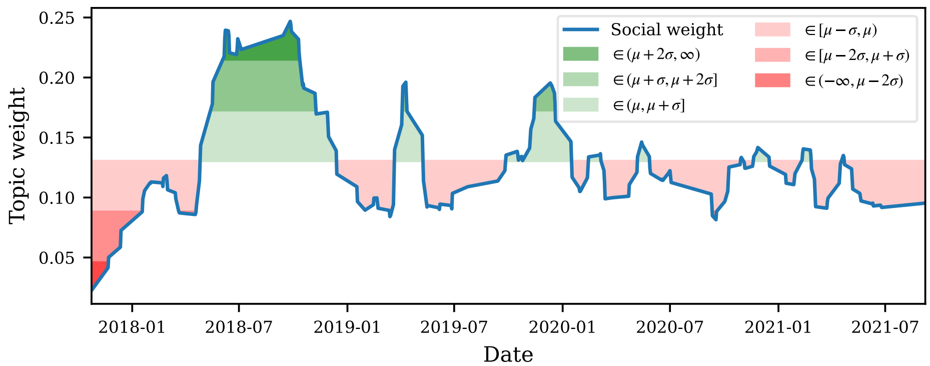

Figure 2: Average weight of *Social* in the Bundestag speeches (with Gaussian smoothing).

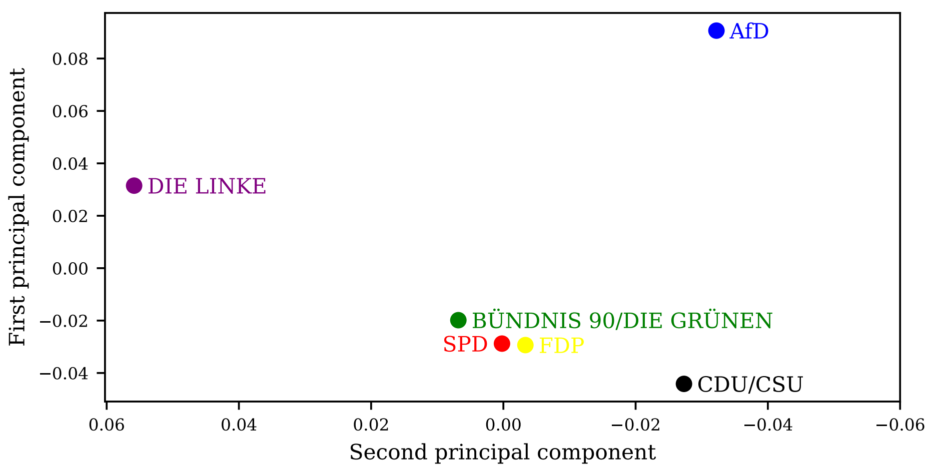

Figure 3: Principal Component Analysis on average topics per party.

an exception). We perform a p-value test for significant differences between the average proportion of negative speeches over all parties and single parties [6]. We assume a binomial distribution of the amount of negative speeches, and use beta-distributed priors. We obtain p-values for the asymmetrical differences between the binomial negative speech probability of the whole set of speeches and the speeches of specific parties. Using a threshold α of 0.05 with Bonferroni correction for multiple testing [7] we refute the null hypotheses that both *SPD* and *CDU/CSU* have a bigger or equal probability of a negative speech than the Bundestag's mean, and that *Bündnis 90/Die Grünen*, *AfD*, and *Die Linke* have smaller or equal probability of a negative speech than the Bundestag's mean. We did not obtain significant results for both asymmetrical tests in the case of *FDP*. These results are coherent with the hypothesis that non-governing parties tend to hold more negative speeches.

4 Discussion and conclusion

Limitations entailed by our methodology suggest caution when interpreting the results. Some speeches are missing in the data; however, since the data loss is due to technical reasons, we might assume that the effect is random and the available data provide a good representation. Furthermore, the data cover a short period of 4 years, which leads to a high influence of individual fluctuations. Besides the usual limitations of the sentiment detection and topic discovery, in topic extraction the best parameters and topics names were selected subjectively. It must be noted that since the discussed

[6] We concentrate the analysis on negative sentiments, since very few positive sentiments are found.

[7] 2 asymmetrical tests per party and 6 parties yield 12 hypotheses, and a corrected $\alpha = 0.004$.

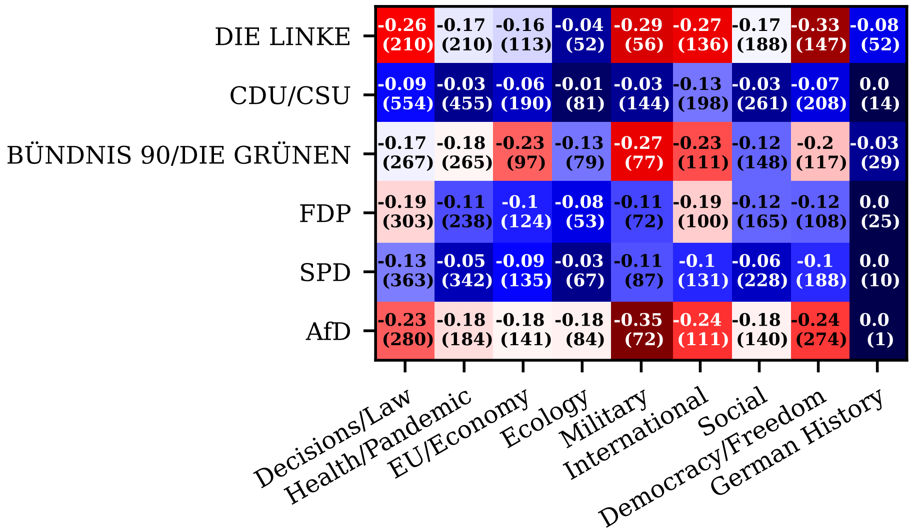

Figure 4: Average sentiment (from -1, fully negative, to +1, fully positive) by party and topic. The number in brackets indicates the number (significance) of speeches from which the mean is constructed. Each speech is assigned to a single topic (highest weight).

topics are mostly defined by the parliament's agenda, only small topic differences can be detected between parties. When implementing p-value testing, we use the party data in the aggregate data as well. This however makes the distributions that are tested for differences more similar and increases the resulting p-value, making it more difficult to show significant results. This effectively decreases the number of significant results strengthening our results.

Despite these limitations, the methodology used is promising. Meaningful results were obtained and topics such as the *Health/Pandemic* and *Social* behave as expected. In addition, the created figures provide information about the temporal behavior and the differences between parties. More data is needed to test for party-sentiment-topic correlations, analyze the topics' evolution over time, and further test the hypothesis that non-governing parties hold more negative speeches (for this, we need to test over different governing coalitions).

References

[1] Oliver Guhr, Anne-Kathrin Schumann, Frank Bahrmann, and Hans Joachim Böhme. Training a broad-coverage german sentiment classification model for dialog systems. In *Proceedings of The 12th Language Resources and Evaluation Conference*, pages 1620–1625, Marseille, France, May 2020. European Language Resources Association.

[2] Joscha Jäger, Alexa Steinbrück, Michael Morgenstern, Olivier Aubert, and Philo van Kemenade. Open parliament tv. https://de.openparliament.tv/, 2021. Accessed 31-January-2022.

[3] Lukas Stern. Das Jahr 2018 im Deutscchen Bundestag: Ereignisse und Beschlüsse. https://www.bundestag.de/dokumente/textarchiv/2018/kw51-jahresrueckblick-584212, 2018. Accessed 02-February-2022.

[4] Lukas Stern. Ereignisse, Debatten, Beeschlüsse – das Jahr 2019 im Bundestag. https://www.bundestag.de/dokumente/textarchiv/2019/kw52-jahresruecklick-673588, 2019. Accessed 02-February-2022.