Chapter 3

Marcelo Jenny, Martin Haselmayer and Daniel Kapla

Measuring incivility in parliamentary debates: Validating a sentiment analysis procedure with Calls to Order in the Austrian Parliament

(Pre-print version)

Cite as: Jenny, Marcelo, Martin Haselmayer, and Daniel Kapla (2021). Measuring incivility in parliamentary debates: Validating a sentiment analysis procedure with Calls to Order in the Austrian Parliament. In: Annemarie Walter (Ed.), *Crossing boundaries: Political Incivility in the Parliamentary, Electoral and Media Arena*. Abingdon: Routledge, pp.56-66.

Abstract

Parliamentary debates sometimes see uncivil behavior by MPs and parliamentary rules of procedure provide instruments such as Calls to Order to sanction uncivil behavior. Incivility is an extreme form of negativity encountered in parliamentary debates. The varied forms of negativity found in parliamentary debates has made them a popular test field for sentiment analysis, the systematic measurement of valence in statements. This chapter describes a sentiment analysis procedure which combines context-sensitive word representations, crowdcoding of negativity in training sentences and a neural network classifier to establish the level of negativity in debate statements. To validate the procedure we try to predict Calls to Order in the Austrian parliament. We find that we can predict Calls to Order from a statement's degree of negativity reasonably well. The procedure therefore offers great potential for a valid and reliable measurement of incivility in parliamentary debates.

Introduction

The study of negative rhetorical interaction in parliament forms part of what Soroka (2014) termed the study of "negativity in politics". Its most prominent subfield is negative campaigning (e.g. Ansolabehere and Iyengar, 1995; Lau and Brown Rovner, 2009; Nai and Walter, 2015; Haselmayer, 2019), but also includes forms of incivility (Stryker et al 2016) in the media (Jamieson, 1992; Mutz and Reeves, 2005), on the internet (Anderson et al 2014) or in parliamentary debates (Brooks and Geer, 2007; Herbst, 2010; Dodd and Schraufnagel, 2013). A recent addition is the use of rhetorical outrage by political actors (Sobieraj and Berry, 2011; Berry and Sobieraj, 2013).

Studying incivility in parliamentary debates provides a distinct perspective on the phenomenon of political polarization. The dominant way of approaching the topic of political polarization has been through a (left-right) spatial representation in which an increase in polarization is understood as political actors, such as parties, moving apart on a dimension of conflict (Dalton, 2008), or when there is a multitude of actors, such as legislators from two parties, sorting into antagonistic camps on political issues with the middle ground less and less populated by legislators from both sides (e.g. Sinclair, 2006; Binder et al, 2009; Abramowitz, 2010). Political polarization, as used in this chapter, is based on how political actors behave in deeds and words towards each other in the parliamentary or in other arenas (e.g. Uslaner, 1993; Iyengar et al, 2012; Jamieson et al, 2018). Rhetorical incivility, the use of 'strongly negative language' that disrespects social or legal norms of civility is one means through which political actors indicate that they disagree with or disapprove of each other. Iyengar et al. (2018) called that "affective polarization".

While patterns of rhetorical interaction between MPs and parties in different legislative terms and under varying government compositions are of interest to students of parliamentary politics and party competition, the chapter's focus is on making a methodological contribution. We use sentiment analysis (Liu, 2015), the systematic study of valence contained in textual data, to establish a measure for a statement's level of incivility. We use manually annotated negative sentences taken from political texts, such as parliamentary debates, to train a supervised classifier and validate the method's capabilities of identifying rhetorical incivility in Austrian parliamentary debates by predicting actual Calls to Orders.

Calls to Order in the Austrian National Council

Parliamentary debates are highly regulated rhetorical exchanges between MPs or between ministers and MPs. the National Council, which is the dominant lower chamber of the Austrian parliament, to whom government is responsible. Its Standing Orders provide an elaborate set of rules regulating the debates, which are further specified by precedent decisions by the president of parliament (Cerny and Fischer, 1982; Atzwanger and Zögernitz, 1999; Gibba, 2013; Jenny and Müller forthcoming).

At the start of a new parliamentary term the National Council elects a president and two deputy presidents. The president has traditionally been a long-serving MP from the largest parliamentary party and the deputy presidents from the second and third largest party (Jenny and Müller, 1995). These three presiding officers direct the parliamentary debates, uphold compliance with the Standing Orders and sanction MPs who fail to do so by violating decency or the dignity of the National Council through 'insulting statements' or behavior that 'does not comply with orders of the President' with a Call to Order (§ 102, Standing order of the National Council). Calls to Order that sanction uncivil rhetorical behavior provide an excellent resource to establish the validity of a sentiment analysis procedure.

Method and Data

We collected Calls to Order in the National Council issued over the course of five parliamentary terms (1996-2013). During that period two types of majority coalition governments were in office. First there was a SPÖ-ÖVP coalition (1996-1999), then came two terms of an ÖVP-FPÖ coalition (2000-2006), followed by a return to SPÖ-ÖVP coalition governments in the last two (2006-2013) terms studied.

Table 1 provides a count of Calls to Order issued based on a keyword query for Calls to Order ("Ordnungsrufe") on the Parliament's website. In these five parliamentary terms the National Council held 756 plenary sessions in which the presiding officers issued 410 Calls to Order. On average, every second plenary session saw a MP sanctioned by a Call to Order.

Two terms stand out with extraordinary high counts of Calls to Order, the 21st and the 24th term. In the 21st term an ÖVP-FPÖ coalition government – the so called 'black-blue' coalition – came into office for the first time. The highly controversial government formation initiated a period of left-right polarization in parliament (Müller and Fallend, 2004). The 24th term, with a SPÖ-ÖVP government coalition in office, saw five parliamentary parties, a record number of MPs expelled from or exiting their party group and a new parliamentary party established by renegade MPs, and temperamental outbursts by MPs sanctioned with Calls to Orders.

Table 3.1: Calls to Order in the National Council (lower house)

Legislative term	Calls to Order	Plenary sessions	Calls to Order per plenary session
20. (1996-1999)	68	182	0.4
21. (1999-2002)	85	117	0.7
22. (2002-2006)	69	163	0.4
23. (2006-2008)	25	75	0.3
24. (2008-2013)	163	219	0.7
20.-24th term	410	756	0.5

Source: calculated from a keyword search ("Calls to Order") at parliament's website (www.parlament.gv.at)

A Call to Order does not always follow immediately after a MP committed a transgression. Sometimes a MP or a parliamentary party lodges a protest to the president after the session and demands the sanctioning of an uncivil interruption by another MP. The president of the National Council then issues a decision on the matter in a subsequent plenary session. We extracted parliamentary debate contributions from the National Council's debate transcripts, but without statements of presiding officers that referred to a past debate. For that reason our dataset of Calls to Order covers only 83 per cent of the total number of Calls to Order issued.

Table 2 shows the number of speeches and Calls to Order that were issued directly in a debate per term and the share of speeches sanctioned with a Call to Order. Overall, one per cent of the speeches was sanctioned, with higher rates of sanctioning in the first term of the black-blue coalition government and the last term of the period studied.

Table 3.2.: Debate speeches and Calls to Order in the National Council

Legislative term	Speeches given in debates	Calls to Order in debates	Calls to Order in per cent of speeches
20. (1996-1999)	7,272	51	0.70
21. (1999-2002)	5,085	67	1.32
22. (2002-2006)	7,159	58	0.81
23. (2006-2008)	2,953	23	0.78
24. (2008-2013)	10,143	141	1.39
20.-24th term	32,612	340	1.04

Source: own dataset of Calls to Order extracted from parliamentary debates

We set up a case-control study that starts from the assumption that Calls to Order are issued to sanction strongly negative language, in line with the rules of order outlined above[i]. We use the texts of the speeches associated with a Call to Order and an identical number of randomly selected unsanctioned speeches to test whether sentiment analysis (Haselmayer and Jenny, 2017; Rudkowsky et al, 2018) can identify these uncivil – that is highly negative – parliamentary speeches.

Method

Our approach for identifying uncivil speeches starts with measuring the level of negativity of sentence-wise statements. It builds on a combination of negativity scores for words and sentences computing word embeddings scores for a sentence first and then running it through a neural network classifier to establish a negative sentiment score. The basics of the approach were presented in Rudkowsky et al (2018). We have improved it here by drawing on a bigger database of word embeddings and a different procedure for the sentiment classification.

Sentiment scores of sentences in a speech derive from words contained in it. Word embeddings capture their relationship with other words that do not have to appear in a speech or text. Words with similar meaning and valence tend to appear in similar contexts: they are `embedded' similarly. This contextual information was utilized in Rudkowsky et al (2018) through the *Word2Vec* program (Mikolov et al, 2013a; 2013b). Here, we use a set of German word embeddings provided by the multilingual NLP library Polyglot[ii] (Al-Rfou et al, 2013). Polyglot calculates individual word embedding features from a corpus of Wikipedia articles in the respective language. We computed the average of the word vectors to get values for a sentence vector as input for a Multilayer Perceptron (MLP) neural network that classified the strength of negative sentiment for the complete sentence.

FastText[iii] (Bojanowski et al, 2016, Grave et al, 2018, Joulin et al, 2016; Mikolov et al, 2018) is an alternative library with a larger set of items (words and subwords).[iv] Using subwords, word components, the fastText library, with help from the *Gensim* library[v], can establish meaningful word vectors even for words that are not found in the corpus. As another improvement to the procedure previously used (Rudkowsky et al, 2018) the procedure now respects word order. A sentence consists of a set of words in a distinct order, which are represented as a sequence of embeddedness word vectors. That preserves information on word order and captures dependences between words. To deal with a sequential data input, a Recurrent Neural Network processes the sequence of vectors and creates a single vector summary that is passed on to a three-layered Multilayer Perceptron (MLP) performing the actual classification of a sentence as neutral, negative or very negative.[vi][vii]

Text preprocessing is minimal: all stopwords and punctuations are kept. To compare results the new procedure was trained with the dataset used in Rudkowsky et al (2018). The training dataset consists of 20,580 sentences with a continuous negativity score per sentence ranging from 0 (“neutral”) to 4 (“very negative”) (Haselmayer and Jenny, 2017, 2018). These continuous scores were matched to one of three output classes (“neutral”, “negative”, “very negative”)[viii]. The new model was trained 60 times with a dropout of 40 per cent over the entire network. The average accuracy of was 63 per cent, outperforming the procedure presented in Rudkowsky et al (2018) by 5 per cent. There we reported an overall accuracy of 58 per cent for a word embeddings approach compared to 55 per cent accuracy for a ‘classic’ bag-of-words approach.

To show that high negativity scores correspond with uncivil parliamentary behavior, we build a balanced case-control sample dataset with 680 speeches selected from parliamentary debates. Half of the speeches were sanctioned by a Call to Order, the other half constitute a random sample of unsanctioned speeches. All speeches have sentiment scores assigned by the procedure described above. Following similar research, we calculate speech sentiment scores by aggregating the mean value of all sentences in a speech (Haselmayer and Jenny, 2017; Rudkowsky et al, 2018). These scores constitute the input variable and issuance of Call to Order the output variable for a logistic regression.

The goal is to establish the degree of correspondence between a high degree of negativity (or incivility) expressed in a parliamentary speech and the issuance of a Call to Order. We posit that the better the negative sentiment scoring procedure works in practice in scoring the degree of negativity contained in a speech the better should be the predicted sorting in sanctioned and unsanctioned speeches.

Evaluating the procedure as proposed rests on a set of assumptions:

a) a Call to Order sanctions uncivil language,

b) uncivil language is very negative language,

c) issuance of a Call to Order is rule-based, but modified over time by new precedent decisions to the president's rulebook, and

d) variation in presiding officers' sanctioning of uncivil language is small.

We have already presented evidence in support of the first two assumptions. The remaining assumptions may be more controversial. Some variation in how the three presidents of the National Council conduct debates and some leeway in dealing with transgressions has to be expected. The first president of the National Council enjoys more of it as the highest parliamentary authority than the other two presidents. The first president also directs more often the most controversial agenda items of a plenary session day.

For a number of reasons a president's leeway in sanctioning uncivil language and behavior in plenary sessions is bounded. Experienced MPs and parliamentary staff members know what typically warrants a Call to Order. Lists of previously sanctioned terms and behavior have been published in answers to parliamentary questions. Parliamentary sessions are open to the public, broadcast on the internet and session transcripts available online. Party competition leads MPs to lodge a protest to the president whenever a transgression by an MP from another party is not immediately sanctioned.

That is one side of the matter. However, because Members of Parliament know where the boundary between the acceptable and the unacceptable has been drawn in the past, there is also continuous innovation in how to rhetorically bash and insult opponents and, perhaps, get away with it. That includes using synonyms of items on the list of forbidden terms or the coining new offensive phrases. Variation in sanctioning behavior by presiding officers, innovative ways of expressing incivility by MPs, and parliamentary rulebook updating via precedent decisions, means that we cannot expect a perfect prediction of Calls to Order issuance from the degree of rhetorical incivility in a parliamentary speech measured. In spite of these challenges, the next section shows the sentiment analysis performs reasonably well in predicting sanctioned uncivil statements.

Results

The aim is to predict ex post speeches from debates in the National Council sanctioned with a Call to Order. The degree of rhetorical incivility is the strength of negative sentiment in a speech based on the scoring method outlined above. We did a binary regression of Calls to Order issued on the negativity of a statement for our sample of 680 speeches, 50 per cent of which were sanctioned, the other 50 per cent were not. Table 3 shows the confusion matrix from the logistic regression model of predicted Calls to Order compared with their actual issuance.

Table 3.3: Confusion matrix of logistic regression model of Calls to Order on Statement negativity

Model predictions	True values Call to Order issued	No Call to Order issued	Row totals
Call to Order predicted	268	96	364
Not predicted	72	244	316
Column totals	340	340	680

The logistic regression model performs slightly better in correctly classifying statements that did not result in a Call to Order than the set of sanctioned statements. The positive predictive value is 73.6 per cent (for statements sanctioned with a Call to Order), the negative predictive value is 77.2 per cent. Highly negative statements – based on the sentiment scoring – that were not sanctioned constitute interesting cases of variation in the sanctioning behavior of presiding officers.

The logistic regression model has a correct prediction rate of 75.3 per cent and improves on the 69.7 per cent of the earlier version presented in Rudkowsky et al (2018). That is considerably better than random guessing. For a dataset with balanced event classes random guessing produces on average a correct prediction rate of 50 per cent. Comparing the results to state-of-the-art sentiment analysis applications is difficult, due to the human decision-making component involved in the issuance of Calls to Order. We see predicting a Call to Order as a harder task than a dichotomous classification of positive and negative texts. Such tasks obtained accuracy scores of about 85% in recent English language applications (e.g. Jianqiang et al, 2018). Levels of accuracy for more complex classification tasks, such as hate speech detection in German language texts, are similar to our study (Ross et al, 2016; Bai et al, 2018).

Conclusion

The study of incivility in parliament provides an interesting perspective on party competition and negativity in politics. Valid and reliable measurement of incivility, however, remains a challenge. We have described a procedure to get a graded measure of incivility and applied it to parliamentary speeches, particularly ones that were sanctioned by a Call to Order. Sanctioned speeches should be located in the upper range of a negativity scale and thus constitute valuable data to do validity tests for sentiment analyses procedures.

Our sentiment analysis procedure combines the crowdcoding of a sample set of statements with a word embeddings approach and a neural network classifier to arrive at a graded measure of negativity. We produced negative sentiment scores for 680 parliamentary speeches in the Austrian National Council, the lower house of the national parliament. Half of them were sanctioned by Calls to Order, the other half were not. This balanced set of statements constitutes a case control design with negativity scores as input variable and Call to Order issuance as outcome variable in a logistic regression model. If measurement of negative sentiment strength works, a logistic regression should be able to separate sanctioned from unsanctioned statements. Overall, the presiding officers' actions were correctly predicted in three out of four cases. The validation exercise performed reasonably well, and the graded measure of negativity obtained by the procedure presented should be useful for analyzing ups and downs of incivility in parliamentary debates over time.

The empirical analysis was on debates in German, but the sentiment analysis procedure can be easily applied to other languages. A strong point of the word embeddings approach is the availability of multi-lingual libraries. Providing a valid and reliable measurement of incivility in parliamentary debates can improve our understanding of negativity in democratic politics in general and its contribution to affective polarization in contemporary societies in particular (Iyengar et al, 2018).

References

Abramowitz, A.I. (2010). *The disappearing center: engaged citizens, polarization, and American democracy*. New Haven: Yale University Press.

Anderson, A.A., Dominique B., Scheufele, D.A., Xenos, M.A., and Ladwig, P. (2014). 'The "nasty effect:" online incivility and risk perceptions of emerging technologies'. *Journal of Computer-Mediated Communication,* 19 (3), p373-387.

Ansolabehere, S. and Iyengar, S. (1995). *Going negative: how attack ads shrink and polarize the electorate*. New York: Free Press.

Brooks, D. J., & Geer, J. G. (2007). Beyond negativity: The effects of incivility on the electorate. *American journal of political science*, *51*(1), 1-16.

Atzwanger, K. and Zögernitz, W. (1999). *Nationalrat-Geschäftsordnung: samt Verfahrensordnung für parlamentarische Untersuchungsausschüsse und umfangreichen Anmerkungen.* 3. revised edition. Vienna: Manz.

Bai, X., Merend, F., Zaghi, C., Caselli, T. and Nissimet, M. (2018). RuG at GermEval: Detecting Offensive Speech in German Social Media. *14th Conference on Natural Language Processing KONVENS 2018*.

Berry, J.M. and Sobieraj, S.(2013). *The Outrage Industry: Political Opinion, Media and the New Incivility*. New York: Oxford University Press.

Binder, A. R., Dalrymple, K. E., Brossard, D. and Scheufele, D. A. (2009). 'The soul of a polarized democracy: Testing theoretical linkages between talk and attitude extremity during the 2004 presidential election'. *Communication Research* 36 (3), p 315-340.

Bojanowski, P., Grave, E., Joulin, A. and Mikolov, T. (2016). CoRR abs/1607.046.06. https://dblp.org/rec/bib/journals/corr/BojanowskiGJM16.

Czerny, W.F. and Fischer, H. (1982). *Kommentar zur Geschäftsordnung des Nationalrates und zum Unvereinbarkeitsgesetz*. 2. Revised edition. Wien: Verlag der Österreichischen Staatsdruckerei.

Dalton, R.J. (2008). 'The quantity and the quality of party systems: Party system polarization, its measurement, and its consequences'. *Comparative Political Studies*, 41 (7), p 899-920.

Dodd, L. C. and Schraufnagel, S. (2013). Taking incivility seriously. In Kelly, S.Q. and Frisch, S.A. (eds.). *Politics to the Extreme: American Political Institutions in the Twenty-First Century*. New York: Palgrave Macmillan, pp. 71-92.

Gervais, B.T. (2014). 'Following the news? Reception of uncivil partisan media and the use of incivility in political expression'. *Political Communication,* 31 (4), p 564-583.

Gibba, M. (2013). *Von Lügen, Pinschern und Ehrabschneidern - Ordnungsrufe im Nationalrat: Eine Analyse der seit 1945 erteilten Ordnungsrufe*. Abschlussarbeit der Grundausbildung, Parlamentsdirektion: Wien.

Grave, E., Bojanowski, E., Gupta, P., Joulin, A. and Mikolov, T. (2018). Learning Word Vectors for 157 Languages. *Proceedings of the International Conference on Language Resources and Evaluation* (LREC 2018).

Haselmayer, M. (2019). 'Negative campaigning and its consequences: a review and a look ahead'. *French Politics,* 17(3), p 355-372.

Haselmayer, M., Hirsch, L., & Jenny, M. (2020). Love is blind. Partisan bias in the perception of positive and negative campaign messages. Love is blind. Partisanship and perception of negative campaign messages in a multiparty system. *Political Research Exchange* 2(1): 1-17. DOI: 10.1080/2474736X.2020.1806002

Haselmayer, Martin and Marcelo Jenny (2018). Friendly fire? Negative campaigning among coalition partners. Research and Politics 5(3): 1-9. DOI: 10.1177/2053168018796911

Haselmayer, M. and Jenny, M. (2017). ‘Sentiment analysis of political communication: Combining a dictionary approach with crowdcoding’. *Quality & Quantity*, 51(6), p2623-2646.

Herbst, S. (2010). *Rude Democracy. : Civility and Incivility in American Politics*. Philadelphia: Temple University Press.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N. and Westwood, S.J. (2018). ‘The origins and consequences of affective polarization in the United States’. *Annual Review of Political Science,* 22(1), p129-146.

Jamieson, K. H. (1992). *Dirty politics: Deception, distraction, and democracy*. Oxford University Press, USA.

Jamieson, K.H., Volinsky, A., Weitz I. and Kenski, K. (2018). ’The political uses and abuses of civility and incivility’. In Kenski, K. and Jamieson, K.H. (eds.). *The Oxford Handbook of Political Communication*. Oxford Handbooks Online. DOI: 10.1093/oxfordhb/9780199793471.013.79_update_001

Jenny, M. and Müller, W.C. (1995). ‘Presidents of Parliament: Neutral Chairmen or Assets of the Majority’ in Döring, H. (ed.), *Parliaments and Majority Rule in Western Europe*. Frankfurt/New York: St. Martin’s Press, pp. 326-364.

Jenny, M. and Müller, W.C. (forthcoming). ‘Austria: Individual talent, official functions, and party size’ in Bäck, Debus and Fernandes (eds.), *The Politics of Legislative Debates*. Oxford: Oxford University Press.

Jianqiang, Z., Xiaolin, G. and Xuejun, Z. (2018). Deep convolution neural networks for Twitter sentiment analysis. *IEEE Access*, *6*, 23253-23260.

Joulin, A., Grave, E., Bojanowski, P. and Mikolov, T. (2016). Bag of Tricks for Efficient Text Classification. *CoRR* abs/1607.01759. https://dblp.org/rec/bib/journals/corr/JoulinGBM16.

Lau, R.R. and Brown Rovner, I. (2009). 'Negative Campaigning'. *Annual Review of Political Science, 12* (1), p285-306.

Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University press.

Mason, L. (2015). ' "I disrespectfully agree": The differential effects of partisan sorting on social and issue polarization'. *American Journal of Political Science,* 59 (1), p128-145.

Mikolov, T., Grave, E., Bojanowski, Puhrsc, C. and Joulin, A. (2018). 'Advances in Pre-Training Distributed Word Representations' in Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018).

Mikolov, T., Sutskever, I., Chen, K., Corrado G. and Dea, J. (2013a). Distributed Representations of Words and Phrases and their Compositionality. CoRR abs/1310.4546. https://dblp.org/rec/bib/journals/corr/MikolovSCCD13.

Mikolov, T., Chen K.,Corrado, K. and Dean, J. (2013b). Efficient Estimation of Word Representations in Vector Space. CoRR abs/1301.3781. https://dblp.org/rec/bib/journals/corr/abs-1301-3781

Mutz, D.C. and Reeves, B. (2005). 'The new videomalaise: Effects of televised incivility on political trust'. *American Political Science Review,* 99 (1), p1-15.

Müller, W.C. and Fallend, F. (2004). 'Changing Patterns of Party Competition in Austria: From Multipolar to Bipolar System'. *West European Politics,* 27 (5), p801-835.

Nai, A. and Walter, A. (2015). *New perspectives on negative campaigning: Why attack politics matters*. Colchester: ECPR Press.

Rami, A., Perozzi, B., and Skiena, S. (2013). 'Polyglot: Distributed word representations for multilingual NLP' in Association for Computational Linguistics (ed.), *Proceedings of the Seventeenth Conference on Computational Natural Language Learning*. Sofia, pp. 183-19. http://www.aclweb.org/anthology/W13-3520.

Ross, B., Rist, R., Carbonell, G., Cabrera, B., Kurowsky, N. and Wojatzki, M. (2017). Measuring the Reliability of Hate Speech Annotations: The Case of the European Refugee Crisis. *arXiv preprint* arXiv:1701.08118.

Rudkowsky, E., Haselmayer, M., Wastian, M., Jenny, M., Emrich, S. and Sedlmair, M. (2018). 'More than Bags of Words: Sentiment Analysis with Word Embeddings'. *Communication Methods and Measures* 12, (2-3), p140-157.

Sinclair, B. (2006). *Party Wars: Polarization and the Politics of National Policy Making*. Norman: University of Oklahoma Press.

Sobieraj, S. and Berry, J.M. (2011). 'From incivility to outrage: Political discourse in blogs, talk radio, and cable news'. Political Communication, 28 (1), p19-41.

Soroka, S.N. (2014). *Negativity in Democratic Politics: Causes and Consequences*. Cambridge: Cambridge University Press.

Stryker, R., Conway, B.A. and Danielson, J.Z. (2016). 'What is political incivility?' *Communication Monographs,* 83 (4), p535-556.

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

---

[i] The president's answer to a parliamentary question on the issuing of Calls to Order in the 24th legislative period contains a precise list of reasons along with exact keywords. Based on 108 Calls to Order, the vast majority (92%) were due to rhetorical negativity, such as personal insults or denigration. https://www.parlament.gv.at/PAKT/VHG/XXIV/ABPR/ABPR_00065/fnameorig_239591.html

[ii] https://polyglot.readthedocs.io

[iii] https://fasttext.cc/. The vocabulary of Polyglot's German language library consists of the 100.000 most frequent words in the corpus. The length of individual feature vectors of word embeddings is 64 dimensions, the length of the feature vector in Fasttext is 300 dimensions.

[iv] It includes approximately two million items for the German language. Subwords can be used to obtain information for unknown compound words. As compounds are common in the German language, and particularly relevant in the context of political speeches, it improves word coverage substantially.

[v] https://radimrehurek.com/gensim/

[vi] The GRU units' dimension is 128 and uses a Rectified Linear Unit Activation function. The MLP transforms the output vector in three layers (with dimensions of 128, 128 and 64 using a Rectified Linear Unit Activation function). The final output layer (with a Softmax activation function) produces a single score per sentence from a three-valued negative sentiment scale.

[vii] We tested a range of structurally more complex model options, including ‚larger‘ models with more layers and/or vectors with higher dimensions but classification performance did not improve further.
[viii] Class assignment of sentences in the dataset was performed by splitting the score interval into equal subintervals: sentences with a score in the interval [0.4 to 3] were considered 'neutral', in the range [4/3 - 8/3] 'negative' and larger values were classified as 'very negative'.