Analysing German parliamentary debates: a special case of calls to order

Anonymous ACL submission

Abstract

Our research investigates incivility in parliamentary discourse, focusing on calls to order (CtO; plural: CtOs) in the German parliament. A notable gap exists in the analysis of CtOs in parliamentary discourse. Consequently, we propose a rule-based method to detect and annotate CtOs in parliamentary speeches and introduce a dataset of German parliamentary speeches spanning 72 years that include CtOs. Furthermore, this paper represents the first attempt to develop a classification system for the triggers of CtOs and to analyze the factors contributing to incivility in parliamentary discourse. Both statistical and empirical evidence suggest that despite strict regulations, issuing CtOs is often subjective, significantly influenced by the session president and prevailing parliamentary trends. Thus, the presidents of the parliament tend to call particular individuals to order preferentially. An insult towards individuals is the most frequent cause of CtO. Generally, male individuals and opposition party members receive more calls to order than their female and coalition party counterparts. Most CtO triggers were detected in speeches dedicated to governmental issues and presidency actions. Dataset is available at: https://anonymous.4open.science/r/cto_analysis-D126/.

1 Introduction

Our research investigates incivility in parliamentary discourse, focusing on calls to order (CtOs) in the German parliament (Bundestag). Call to order (CtO) is a valuable resource for examining the negativity and incivility in political debates, and offers a unique perspective on political polarization (Jenny et al., 2021). Moreover, analysis of CtOs as markers of disruptive language is a novel approach to studies of parliamentary corpora, going beyond traditional sentiment or stance analysis. Furthermore, methods of automatic analysis applied to parliamentary data support government transparency and accountability. However, a notable gap exists in the analysis of CtOs in parliamentary discourse. To the best of our knowledge, the sole effort in this area is that of (Jenny et al., 2021). In this study, we present a novel and comprehensive analysis of speeches delivered by German politicians spanning 72 years of parliamentary history, employing both automated and manual methodologies. CtOs have been largely overlooked in political research. Consequently, this paper represents the first attempt to develop a classification system for the triggers of CtOs and to analyze the factors contributing to incivility in parliamentary discourse. Moreover, we propose a rule-based method for the detection and annotation of CtOs within parliamentary speeches and introduce a novel dataset comprising annotated speeches that include a CtO.

In the present research, we will address the following research questions: **RQ1:** Which topics caused most CtOs? **RQ2:** What are the most frequent trigger classes for issuing a CtO? **RQ3:** How do factors such as political party affiliation, individual politicians, legislative periods, and topics relate to issuing CtOs?

1.1 Terminology used in this paper

A call to order, issued by the president of the session, serves as a disciplinary measure in response to breaches of parliamentary protocol, such as instances of personal insults among members or disruptions to the proceedings. Only the president may call members of the German parliament to order by stating their name (Schindler and Feldkamp, 2005, p. 447). Figure 1 demonstrates an example of CtO and how it is triggered during a parliamentary session. In the present paper, speeches of the president are referred to as presidency actions. An interjection is an interruption during a speech or introduction of another person[1].

[1] https://de.wiktionary.org/wiki/Zwischenruf

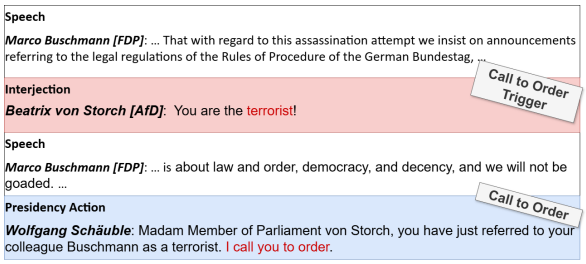

Figure 1: Example of a trigger (red) and an issued call to order (blue). Translated to English from German debates.

A legislative period (LP) is a period in which a parliament can act as a lawmaker and generally lasts four years in Germany. Our data spans a period from September 7, 1949, to September 7, 2021, which covers 19 legislative periods (LPs).

2 Related works

Recent research on parliamentary discourse has focused on the use of automated or semi-automated analytical methods. Within the framework of gender-based research, Ash et al. (2024) examined the differences between reactions to speeches given by male and female parliamentary members (PM) in the German parliament, focusing on interruptions and employing topic modelling techniques. Similarly, Mandravickaitė and Krilavičius (2017) investigated gender differences in language use in the professional environment based on parliamentary speeches in the Lithuanian Parliament using stylometric analysis. In the context of the United States, Miller and Sutherland (2023) analyzed interruptions in congressional hearings to explore interruption behavior influenced by gender and topic.

In the domain of sentiment analysis, Abercrombie and Batista-Navarro (2020) introduced ParlVote, a benchmark corpus designed for the evaluation of sentiment analysis methods in the political domain, utilizing transcripts from the UK House of Commons debates. Several experimental approaches were applied to assess sentiment analysis performance on this dataset. Additionally, Tarkka et al. (2024) compared the performance of generative (GPT) and fine-tuned BERT-based models in emotion detection tasks applied to transcripts of Finnish parliamentary plenary sessions.

Within the scope of discursive framing research, Reinig et al. (2024) analyzed speech acts in German parliamentary debates using a manually annotated dataset in combination with a fine-tuned BERT-based classifier. In a related effort, Rehbein et al. (2024) examined the use of factive expressions in political rhetoric and introduced GePaDe_SpkAtt, a corpus for speaker attribution based on the German parliamentary debates. This work also involved training a model for predicting speech events across a large corpus of parliamentary texts.

From a perspective of negativity analysis, Jenny et al. (2021) analysed negativity in Austrian parliamentary discourse by predicting instances of calls to order. Further, Haselmayer et al. (2022) explored whether the speaker's gender and debate context impact the level of negativity, utilizing sentiment analysis and word embedding techniques.

3 Data and Method

We utilized an annotated XML version of the GermaParl corpus (Blaette, 2017), which comprises a collection of transcribed protocols of debates in the German parliament. The raw data underwent processing, including conversion to a format optimized for analysis, splitting speech contributions into sentences and explicit parsing sentences containing CtOs. Calling to order in the German parliament is regulated, consequently, specific words indicating a CtO are used. Therefore, we employed a rule-based approach to identify CtOs within parliamentary speeches[2]. As Table 1 demonstrates, 42% of all speech contributions in GermaParl are presidency actions, and 0,1% of presidency actions contain a CtO.

	count
total speech contributions	958,098
presidency actions	399,807
speech contributions containing a call to order	558

Table 1: Number of speech contributions and calls to order in GermaParl corpus

Subsequently, we extracted references to individuals mentioned in these calls using a Named Entity Recognition (NER) model (Akbik et al., 2018), trained to recognize 4 types of entities in German texts, including names of individuals. CtOs that lacked identifiable individuals or referenced multiple individuals were manually annotated. Finally, we applied a rule-based method to resolve ambiguities among identified individuals and match them

[2]Data processing workflow and rules for matching sentences containing a CtO are provided in Appendix C. Dataset can be found at our GitHub page: https://anonymous.4open.science/r/cto_analysis-D126/.

with a comprehensive database of all members of the German parliament since 1949[3].

As the last step, we analysed and manually annotated speech contributions containing a CtO, categorizing them according to the underlying cause that triggered the CtO[4]. We developed a classification scheme according to the manual analysis of CtOs, as no such classification was conducted in previous research. Additionally, we detected the topic discussed in each speech using a classification model, trained to detect 21 topics in speeches of the German parliament (Klamm et al., 2022)[5].

4 Results

Analysis revealed that the insult towards an individual (ITO) is the most prevalent cause prompting a CtO, with a median occurrence of 17 per LP, followed by miscellaneous (MISC) (median of 6 per LP), general insult (GI) (median of 3 per LP), and non-verbal (NV) (median of 1 per LP). Additionally, a total of 48 instances of verbal actions that led to a CtO were identified but were either not transcribed or could not be located within our dataset (NDV) (Appendix A, Table 4). As illustrated in Figure 2-A, the distribution of causes across LPs is non-uniform, with high standard deviations observed for all causes. The most frequently occurring cause, ITO, is present in all LPs except for LPs 16, 17, and 18.

The χ^2 test for independence was conducted using the Monte Carlo method to assess relationships between variables, as most of the data did not meet the assumptions required for the χ^2 test. To assess the association strength between variables, we additionally applied Cramér's V measure using the χ^2 statistics from the Monte Carlo simulation[6]. As Table 3 demonstrates, statistically significant relationships were found between a CtO cause and LP, date, year, and the session's sequence number in the LP. However, the associations between these variables were negligible.

Figure 2-B demonstrates that, overall, men are called to order more frequently than women. However, in LPs 11, 16, and 19, a greater proportion of female PMs were called to order compared to their male counterparts. The median number of men receiving a call to order per legislative period is 19, compared to 5.5 for women. This corresponds to 3.72% of all male parliamentarians, while the proportion for women is close to zero. However, high standard deviations were observed, indicating substantial variability in the data (Appendix A, Table 5). Overall, 5.25% of men and 2.62% of women were called to order through the history of the German parliament. Statistically significant relationships were found between the gender of a person called to order (PCO) and the cause of the CtO, LP, and the session's president. Additionally, a moderate association was observed between the session's president and the PCO's gender, as well as between the PCO's gender and the LP. In contrast, a weak association was found between PCO's gender and the cause of the CtO.

As presented in Figure 2-C, opposition party members receive more CtOs than coalition party members, with a median of 10 per legislative period compared to 6 for coalition members. However, high standard deviations indicate significant variability in the data (Appendix A, Table 6). A moderate association was found between the session president and the PCO's party affiliation. Additionally, a strong association was observed between the gender of the session president and the PCO's party, as well as between the president of the session and the PCO. In contrast, the president's party showed only a weak association with the PCO's party.

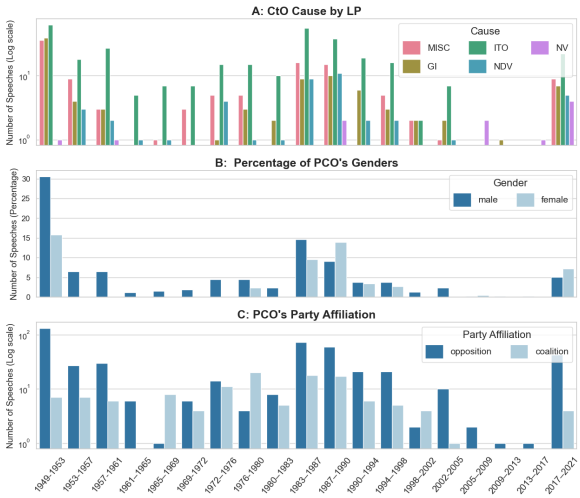

Figure 2: Distribution of causes, genders and party affiliations of PCOs over legislative periods (LPs).

A statistically significant relationship was found between the discussed topic and the presence of a CtO in speech, though a negligible association

[3] https://www.bundestag.de/services/opendata

[4] Annotation schema is provided in Appendix B.

[5] Classification model was applied to the whole speech, excluding presidency actions, as presidency actions merely include moderation of the session.

[6] For the analysis of the relationship between party affiliation, gender, the name and PCO's party affiliation, only disambiguated individuals were considered.

was observed. The highest number of CtO causes was observed in speeches related to governmental issues (188), followed by presidency actions (89), civil affairs (56), and international affairs (48). No CtOs were recorded in discussions on foreign affairs and culture (Appendix A, Figure 4). Figure 3 illustrates the distribution of the top 10 topics containing CtOs over the 72 years. Government remains the most discussed topic across all LPs, and the number of speeches on most topics has increased over time. However, there was a sharp decline in immigration-related speeches between LPs 1 and 3, with a continued decrease in subsequent LPs. No statistically significant relations were found between the gender of the president, and the PCO's gender and party affiliation, as well as the cause that triggered CtO. Additionally, no statistically significant relationship was found between the presence of CtO trigger in a speech and the speech's position (sequence number) in the agenda.

variable1	variable 2	p-value	Cramér's V
name of the president	name of the PCO	0.0	**0.795**
	gender of the PCO	0.0	***0.462***
	party of the PCO	0.0	***0.464***
	cause of the CtO	0.0	***0.4***
	PCO's party affiliation	0.0	**0.524**
gender of the president	gender of the PCO	*1.0*	
	party of the PCO	0.0	***0.326***
	cause of the CtO	*0.105*	
	PCO's party affiliation	*0.109*	
party of the president	party of the PCO	0.0	*0.28*
name of the PCO	cause of the CtO	0.0	**0.713**
gender of the PCO	cause of the CtO	0.043	*0.13*
	legislative period	0.0	***0.4***
party of the PCO	cause of the CtO	0.0	*0.267*
CtO trigger	date	0.0	*0.109*
	LP	0.0	0.035
	session's sequence number in LP	0.0	0.028
	speech's sequence number in the agenda	*0.052*	
	discussed topic	0.0	0.02
	year	0.0	0.038

Table 2: The χ^2 test with a Monte Carlo method and Cramér's V. P-value above the threshold marked with italicized text. A small association is marked with italicized text, a medium association with italicized bold text, and a large association with bold text.

5 Conclusion

In this study, we conducted a manual analysis of CtOs in the German parliament and developed a classification consisting of 5 underlying reasons (causes) for issuing a CtO. Our analysis indicates that ITO is the most frequent CtO trigger. NV comprises the smallest parts of CtO triggers. Moreover, statistical testing suggests that certain presidents are more likely to be prompted by specific triggers. Additionally, particular parliamentary members tend to employ specific classes of insults. At the same time, no gender-specific classes of insult were observed.

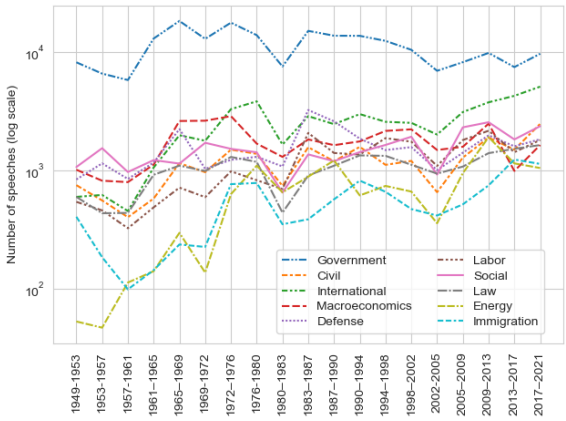

Figure 3: Distribution of the top 10 topics that caused CtOs over legislative periods.

Following, we applied a classification model that distinguished speech contribution to 21 topics, including presidency actions, as an additional category. Most CtO triggers were detected in speeches dedicated to governmental issues and presidency actions. The χ^2 test revealed a statistically significant association between the topic and the CtO trigger; however, a Cramér's V showed only a weak association between these variables, which indicates that this association is not of practical interest and might occur due to the large data sample size.

Notably, session presidents tend to call particular individuals to order preferentially. Moreover, presidents are more likely to call representatives of certain parties and genders to order. In addition, CtOs are associated with the party's affiliation. Generally, male individuals and opposition party members receive more calls to order than their female and coalition party counterparts. This supports the hypothesis that opposition members are more prone to breaching parliamentary order. Historically, there are fewer women than men in the German parliament, which can contribute to the pattern. Furthermore, the likelihood of being called to order varies by gender, depending on LP. However, no statistically significant relationship was detected between the gender of the session president and that of the PCO.

In conclusion, both statistical and empirical evidence suggest that despite strict regulations, issuing CtOs is often subjective and significantly influenced by the session president and prevailing parliamentary trends.

6 Limitations

This study is subject to several limitations. Firstly, we employed a semi-automated method to annotate the corpus. Sentences containing CtO instances were identified using a rule-based approach, which is a legitimate choice in this context, given that disciplinary measures in the German parliament are strictly regulated and, therefore, exhibit specific patterns. Nevertheless, a manual review revealed that these patterns occasionally resulted in false positives, as illustrated in the following example:

- **DE:** *Ich kann nur wegen der Zwischenrufe zur Ordnung rufen, die ich selber höre.*
- **EN:** *I can only call to order the interjections that I hear myself.*

Furthermore, the rule-based approach may not detect CtOs issued using non-conventional phrasing if such occurs in the dataset. We opted against the approach proposed by Jenny et al. (2021), as it showed a correct prediction rate of only 75.3%, and we believe that this would not capture CtOs triggered by speeches lacking explicit negative connotations, as demonstrated in the following example:

- **DE:** *Die Oder-Neiße-Grenze ist die Grenze des Friedens.*
- **EN:** *The Oder-Neisse border is the border of peace.*

Secondly, a semi-automated approach was utilized to extract and disambiguate called-to-order individuals, which also may lead to false annotations.

Additionally, for different reasons, we were not able to disambiguate all individuals mentioned in CtOs, nor all speakers in the corpus. Therefore, some statistical tests were conducted only with the disambiguated data.

Moreover, we used a classification model (Klamm et al., 2022) to find discussed topics in the speeches. This model was specifically trained to distinguish topics in speeches in the German parliament. However, the F1-score for some categories, such as Social Welfare and Public Lands, was under 0.5, which can cause false classification of speeches containing this topic. We decided to apply the model to the whole speech text and not to single sentences or paragraphs, as generally, one speech is dedicated to a specific topic. There could be some variations from the topic during the speech due to interjections, but the general topic stays the same. For future work, we also consider applying other techniques, such as the seeded Latent Dirichlet allocation as in Watanabe and Zhou (2022).

In addition, this paper focuses on the first analysis of calls to order; to our knowledge, there is no equivalent research in this area. Therefore, we focused on a general analysis of factors influencing incivility in parliamentary debates without a deeper investigation of single factors.

Notably, in our statistical analysis, we accounted for the uneven distribution of some variables (such as gender, occurrence of a CtO in speech or topics) in our dataset and adjusted the statistical analysis accordingly. All the reported findings are statistically significant.

Moreover, to the best of our knowledge, no pre-existing classification frameworks for calls to order currently exist. This paper marks the first attempt to systematically analyze and categorize such calls to order. While calls to order can be classified in various ways, such as by focusing on specific insult types, this study emphasizes overarching features of insulting behaviour.

Finally, due to the absence of a benchmark dataset for this task, a quantitative evaluation of the rule-based methods was not feasible. However, because of the limited size of the analyzed dataset, all rule-based annotations were verified manually.

References

Gavin Abercrombie and Riza Batista-Navarro. 2020. ParlVote: A corpus for sentiment analysis of political debates. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 5073–5078, Marseille, France. European Language Resources Association.

Alan Akbik, Duncan Blythe, and Roland Vollgraf. 2018. Contextual string embeddings for sequence labeling. In *COLING 2018, 27th International Conference on Computational Linguistics*, pages 1638–1649.

Elliott Ash, Johann Krümmel, and Jonathan B. Slapin. 2024. Gender and reactions to speeches in german parliamentary debates. page ajps.12867.

Andreas Blaette. 2017. Germaparl. corpus of plenary protocols of the german bundestag. https://github.com/PolMine/GermaParlTEI. Accessed: 2024-12-09.

Martin Haselmayer, Sarah C Dingler, and Marcelo Jenny. 2022. How women shape negativity in parlia-

mentary speeches—a sentiment analysis of debates in the austrian parliament. 75(4):867–886.

Marcelo Jenny, Martin Haselmayer, and Daniel Kapla. 2021. *Measuring incivility in parliamentary debates: validating a sentiment analysis procedure with calls to order in the Austrian Parliament*, pages 56–66.

Christopher Klamm, Ines Rehbein, and Simone Paolo Ponzetto. 2022. FrameASt: A framework for second-level agenda setting in parliamentary debates through the lense of comparative agenda topics. In *Proceedings of the Workshop ParlaCLARIN III within the 13th Language Resources and Evaluation Conference*, pages 92–100, Marseille, France. European Language Resources Association.

Justina Mandravickaitė and Tomas Krilavičius. 2017. Stylometric analysis of parliamentary speeches: Gender dimension. In *Proceedings of the 6th Workshop on Balto-Slavic Natural Language Processing*, pages 102–107, Valencia, Spain. Association for Computational Linguistics.

Michael G. Miller and Joseph L. Sutherland. 2023. The effect of gender on interruptions at congressional hearings. *American Political Science Review*, 117(1):103–121.

Ines Rehbein, Josef Ruppenhofer, Annelen Brunner, and Simone Paolo Ponzetto. 2024. Out of the mouths of MPs: Speaker attribution in parliamentary debates. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 12553–12563, Torino, Italia. ELRA and ICCL.

Ines Reinig, Ines Rehbein, and Simone Paolo Ponzetto. 2024. How to do politics with words: Investigating speech acts in parliamentary debates. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 8287–8300, Torino, Italia. ELRA and ICCL.

Peter Schindler and Michael F. Feldkamp. 2005. *Datenhandbuch zur Geschichte des Deutschen Bundestages 1994 bis 2003: eine Veröffentlichung der Wissenschaftlichen Dienste des Deutschen Bundestages*. Nomos Verlagsgesellschaft.

Otto Tarkka, Jaakko Koljonen, Markus Korhonen, Juuso Laine, Kristian Martiskainen, Kimmo Elo, and Veronika Laippala. 2024. Automated emotion annotation of Finnish parliamentary speeches using GPT-4. In *Proceedings of the IV Workshop on Creating, Analysing, and Increasing Accessibility of Parliamentary Corpora (ParlaCLARIN) @ LREC-COLING 2024*, pages 70–76, Torino, Italia. ELRA and ICCL.

Kohei Watanabe and Yuan Zhou. 2022. Theory-driven analysis of large corpora: Semisupervised topic classification of the un speeches. *Social Science Computer Review*, 40(2):346–366.

A Data statistics

	count
total number of issued CtOs	596
number of CtOs with disambiguated individuals called to order	513
number of CtOs with not disambiguated individuals called to order	96
number of PMs with the presidency role who issued calls to order	50

Table 3: Number of speech contributions and calls to order in GermaParl corpus

cause	total frequency	median per LP	standard deviation
ITO	344	17	18.80
MISC	120	6	10.54
GI	106	3	11.20
NDV	48	2	3.20
NV	13	1	1.46

Table 4: Number of speeches containing specific cause triggered CtO.

PCO gender	number of PMs called to order	% of PMs called to order	number of PMs in parliamnet	median number of PMs called to order per LP	standard deviation	median % of PMs called to order per LP	standard deviation (%)
male	493	5.25	9390	19	31.99	3.72	7.07
female	59	2.62	2249	5.5	6.57	0	4.99

Table 5: Number of PCOs distinguished by their gender.

PCO's party affiliation	total frequency	median per LP	standard deviation
opposition	460	10	33.14
coalition	123	6	5.72

Table 6: Number of CtO distinguished by PCO's party affiliation.

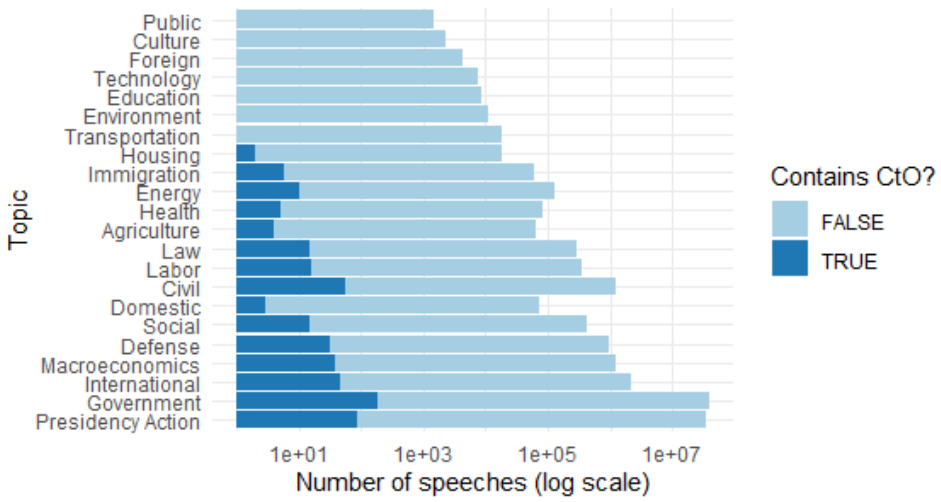

Figure 4: Distribution of topics and CtO inclusion in the topic.

B Annotation schema

Based on the manual review of the dataset, we propose the following classification of actions that caused a call to order (Table 7).

class name	abbreviation	description	example DE	example EN
insult towards individual	**ITO**	insult towards an individual	Schreiner [SPD]: Wild gewordener Gartenzwerg!	Schreiner [SPD]: Garden gnome gone wild!
general insult	**GI**	insult towards a group of people, party, event, actions, etc.	Abg. Renner: Die Union der Faschisten von gestern ist fertig!	PM Renner: The Union of Fascists of yesterday is finished!
non-verbal	**NV**	non-verbal actions that caused a call to order	Abgeordnete der Fraktion Die Linke halten Transparente und Fahnen hoch.	Members of the parliamentary group Die Linke hold up banners and flags.
not documented verbal	**NDV**	verbal actions that caused a call to order but were not transcribed.	Der Abg. Dr. Richter [Niedersachsen] wendet sich dem amtierenden Präsidenten zu und spricht unter andauernder großer Unruhe des Hauses auf ihn ein, ohne daß seine Worte vom Haus und am Stenographentisch verstanden werden können.	PM Dr Richter [Lower Saxony] turns to the President-in-Office and speaks to him, to the continued great agitation of the House, without his words being understood by the House and the stenographers' table.
miscellaneous	**MISC**	all other verbal actions excluding direct insults that caused a call to order	Gerd Andres [SPD]: Wie lange darf der eigentlich noch reden, Herr Präsident? Ist das unbegrenzt?	Gerd Andres [SPD]: How long is he actually allowed to talk, Mr President? Is that unlimited?

Table 7: Classification schema

C Data Processing

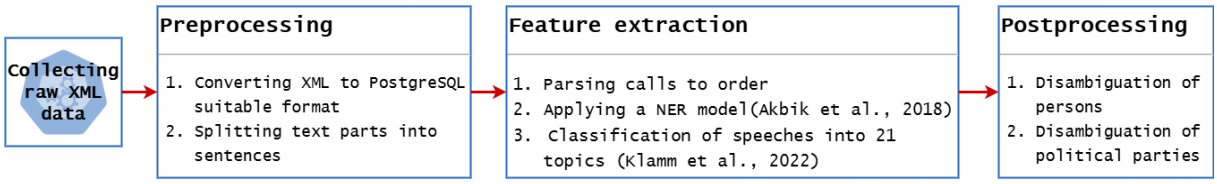

Figure 5: Data processing workflow.

1. Raw parliamentary speeches were sourced from the GermaParl corpus (Blaette, 2017).

2. After collecting the data, we preprocessed the raw data, including converting it to a format more suitable for our analysis. At the last step, speeches were split into sentences using the sentence-splitter library[7].

3. Calls to order, unlike interjections, are not explicitly indicated in the GermaParl corpus. Therefore, in the first step, we manually reviewed a part of the dataset containing only the speeches of the session's president. Based on this review, we developed a set of rules to identify calls to order, as illustrated in Table 8. Following, we applied these rules to analyze only the speeches given by the session's president to detect instances of calls to order, using regular expressions[8].

rule number	rule description
rule 1	match substring *'ordnungsruf'* if substrings *'erteile'* or *'erteilen'* are also in the string but substring *'ordnungsruf'* is not preceeded by substrings *'keinen'*, *'kein'* or *'erteilten'* and substring *'nicht'* is not in the string.
rule 2	match substring *'zur ordnung'* only if subsrtring *'rufe'* is in the string and *'zur ordnung'* is not preceded by *'gesetz'* or *'gesetzes'*.

Table 8: Rules to match calls to order

Following that, we proceeded to look for the persons mentioned in the call of order. For this aim, we applied a Named Entity Recognition (NER) model (Akbik et al., 2018), trained to recognise four types of entities in German texts, including names of individuals.

[7] https://github.com/mediacloud/sentence-splitter
[8] https://docs.python.org/3/library/re.html

4. For the topic analysis, we applied a model (Klamm et al., 2022) specifically trained on German parliament data to classify text into 21 categories. The model was applied to the whole speech text, excluding presidency actions, as presidency actions merely include moderation of the session.

5. At the next step, we applied a rule-based approach to disambiguate names of individuals found in our dataset. For this, we utilized a database containing the names of all members of the German Parliament throughout its history[9]. This analysis helped classify the individuals into three distinct groups.

   - **Group 1:** unique surnames or surname-name combinations
   - **Group 2:** multiple occurrences of the same surname or surname-name combination
   - **Group 3:** no occurrence of surnames or surname-name combinations in the database

   Individuals from Group 2 required additional disambiguation. To achieve this, we aligned them with the legislative periods during which the calls to order occurred. If a call to order date fell within the time frame of a person's tenure in the parliament, that individual was considered a match. If still multiple matches were found in the database, these names were disambiguated manually. Individuals from group 3 were not disambiguated.

   Additionally, we proceeded with the disambiguation of political parties using pattern matching and a comprehensive list of all German political parties and their possible abbreviations throughout the history of the parliament. This process enabled us to standardize party mentions into a unified format. For example, *Christlich Demokratische Union Deutschlands* and *CDU* would be recognized as the same entity. Finally, all the annotations were verified manually.

[9] https://www.bundestag.de/services/opendata