GePaDeU – A Multi-layer Corpus of German Parliamentary Debates with Rich Semantic and Pragmatic Annotations

Ines Rehbein♠, Julian Schlenker♣, Lars Ostertag♣, Simone Paolo Ponzetto♣
University of Münster♠, University of Mannheim♣
rehbein@uni-muenster.de, {julian.schlenker, ponzetto}@uni-mannheim.de

Abstract

This paper presents GePaDeU, a new manually annotated corpus of **Ge**rman **Pa**rliamentary **De**bates with **U**nified layers of semantic and pragmatic information. The data includes parliamentary speeches from the German Bundestag, ranging over a time period from 2017–2021, with 267 speeches given by 197 members of parliament. The final release of our corpus unifies multiple annotation layers, including entity-level annotations, the annotation of speech events and their corresponding speakers, functional speech acts, clause-level aspect, and moral framing. We provide an overview of the various annotation layers and illustrate how the semantic and pragmatic annotations can be combined for corpus-linguistic studies and discourse analyses, and to answer research questions in the field of political science. The new resource will be made freely available for the research community.

Keywords: Parliamentary Debates, Annotation, Political Text Analysis

1. Introduction

Using text as data for studying research questions in the political and social sciences has emerged as a fruitful field of research. Work in the computational social sciences has made extensive use of unsupervised methods to produce evidential support from text, employing variations of lexical features extracted from text or learning latent representations in a mostly unsupervised manner. Many studies have focussed on inferring an actor's ideological position on a left–right scale or on investigating polarisation on social media (Laver et al., 2003; Slapin and Proksch, 2008; Glavaš et al., 2017; McPherson et al., 2001). These approaches scale well to large data, however, they are often limited by their lack of interpretability.

Other work has made extensive use of dictionaries to study political actors' and voters' attitudes and behaviour (Albaugh et al., 2013; Gründl, 2022; Lipsitz, 2018; Simonsen and Widmann, 2025). Many concepts, however, cannot be easily captured on the lexical level but require more context-sensitive methods that consider the semantic environment of a word. A case in point are pragmatic concepts like speech acts or the identification of framing in political text. For such constructs, a supervised approach that trains text classifiers on manually coded data to learn context-sensitive representations seems more promising. Our work belongs to the latter category, trying to capture complex semantic and pragmatic constructs via structured annotations by trained coders, to serve as training data for supervised approaches.

We present GePaDeU, a manually annotated corpus of **Ge**rman **Pa**rliamentary **De**bates with **U**nified layers of semantic and pragmatic information. While (parts of the) individual annotation layers have already been made available,[1] our new unified multilayer resource now allows users to search across layers and extract combinations of semantic and pragmatic labels to study the language of politics. We expect that our data will be useful for corpus- and discourse-linguistic studies of political communication as well as for research in the computational social sciences.

In the paper, we give an overview of the new, unified release of our corpus and its annotation layers and present use cases that illustrate how the different layers can be searched and combined and how users can export the results. We will release these use cases as Jupyter notebooks, for illustration and as a starting point for users who want to work with the data.

2. Data and Annotation Layers

We now describe our data and motivate the different annotation layers. In addition, we report descriptive statistics and baseline results for predicting the annotations for each layer in our corpus.

2.1. Data

The data includes debates from the German Bundestag, retrieved from Deutscher Bundestag – Open Data.[2] The data set includes 267 speeches, mostly from the 19th legislative term (2017-2021),

[1] See Reinig et al. (2024); Rehbein et al. (2024, 2025b).

[2] The data, datasheet and annotation guidelines (partly in German) are available from our github repository: https://github.com/umanlp/gepadeu.git.

Proceedings of the Fifteenth Language Resources and Evaluation Conference (LREC 2026), pages 1794–1810
11–16 May 2026.

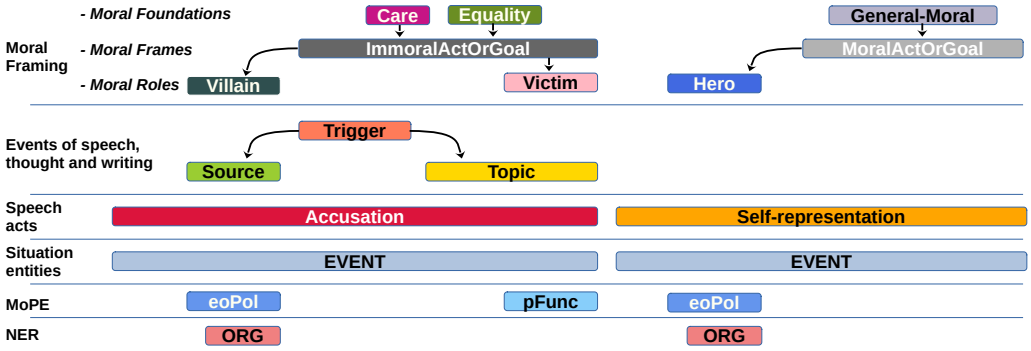

Figure 1: Example adapted from a quote in our data, illustrating the different annotation layers in GEPADEU.

given by 195 different speakers from 6 parties (CDU/CSU: 76, SPD: 58, AfD: 39, FDP: 34, The Left: 29, Greens: 27, non-attached: 4). The total size of the data is 214,272 tokens. Each speech is accompanied with metainformation, such as the speaker's name and party affiliation and the date the speech was held (see Table 10 in the Appendix). Due to time constraints, the number of annotated speeches differs across annotation types (NER: 40 speeches, all other annotation layers: 248-267 speeches). For more detailed information on data extraction and sampling, please refer to the datasheet.

2.2. Annotation Layers

Our corpus includes entity annotations on the token level (NER, Mentions of the People and the Elite), text sequence annotations (Speech Acts, Situation Entities) and more complex, structured annotations with relations between the coded concepts (Events of speech, thought and writing, Moral Framing). Below, we give a quick overview of the different annotation layers illustrated in Figure 1, outlining the main motivation, the most important properties of the annotations, inter-annotator agreement and baselines for automatic prediction.

2.2.1. Named Entities (NER)

Motivation Named Entities play an important role for the extraction of information from large text. In the context of political debates, we are interested in identifying not only persons and organisations, but also geo-political entities, events, and national, religious and political identities.

Annotation details Our annotations follow a scheme adapted from the OntoNotes NER inventory (Weischedel et al., 2013) that has been applied to and tested for German by Ruppenhofer et al. (2020) for news headlines and bibliographic interviews. The annotation inventory includes 30 different labels, grouped for reference into 5 super-categories, as well as a remainder category MISC for miscellaneous named entities (see Table 7 in the Appendix for an overview). Each file has been annotated by two coders with a background in linguistics.

Corpus statistics The NER annotations represent the smallest annotation subset in our corpus, covering 40 speeches with 1,639 annotated entities. As larger NER-annotated datasets already exist for German, we decided to only create a test set that can be used for domain adaptation and evaluation.

Baselines We train a transformer-based token classification model (Devlin et al., 2019) on the Twitter news headline data (Ruppenhofer et al., 2020) and apply the classifier to the parliamentary debates. Using the sequeval package in strict mode for evaluation,[3] we obtain an F1 score (micro-F1 averaged over all entities) of 90% (weighted avg. F1: 89%).

2.2.2. Mentions of the People and the Elite

Motivation Our next annotation layer encodes group mentions as instances of *the Elite* or as representatives of *the people*. This is motivated by our interest in studying *thin* and *thick* populism. *Thin* populism is defined as a "communication style of political actors that refers to the people" (Jagers and Walgrave, 2007, p.233) while the concept of *thick* populism combines people-centrist references with anti-elitism and the exclusion of certain minority groups, also considered as *othering* (Staszak, 2009; Fielder and Catalano, 2017). An important step towards detecting populist rhetoric is the identification of references to *the People* and to *the Elite*, which motivated us to annotate those instances in our corpus. The annotations are also relevant

[3] Available from https://github.com/chakki-works/seqeval.

for studying *group-based appeals* in political text. Group-based appeals are positive or negative references towards social groups, which political actors often make strategic use of, i.e., when communicating with voters in order to gain new voters' support or to mobilise their core supporter base (Thau, 2019; Huber, 2022; Thau, 2024; Huber and Haselmayer, 2024). Our schema includes references to social groups, for example to People by Ethnicity (pEth), People by Nation (pNat), People by Age (pAge), or People by Social Variables (pSoc) (e.g., gender, class). This will enable political scientists to search for group mentions and study their strategic use by politicians in German parliamentary debates.

Annotation details Our annotation scheme is adapted from Wirth et al. (2019) and extends their classification by adding new classes and converting it into a hierarchical schema (see the overview document in our github repository), with a coarse-grained binary distinction at the top and 27 fine-grained labels in lower levels of the schema. The task was set up as a multilabel task, as entities could belong to more than one class. For example, references to "the German farmers" belongs to the classes "people by nationality" (pNat) and "people by functionality" (pFun). The data has been annotated by two trained student annotators with background in political science and sociology. Inter-annotator agreement was measured as the Micro F1 for *exact match* between the annotations for both coders (0.69%) and, to provide a more realistic assessment of the quality of the annotations, as the *overlap* between the labelled spans identified by each coder (0.80%). A substantial part of the annotation disagreements was due to mismatches in span boundaries where one annotator decided to include an adverb, PP or relative clause in the annotation span while the other annotator did not.

Corpus statistics The annotations are available for all 267 speeches in GePaDeU. Table 8 shows the data statistics and distribution of labels in the corpus. Not surprisingly, the most frequent references in parliamentary debates are to instances of political actors or parties (epPol, eoPol). Looking at references to *the people*, we see that the majority class is People by Function (pFun), referring to social groups defined by their vocation (farmers, craftsmen, the unemployed) or societal function (e.g., parents, tax payers).

Baselines Results for an transformer-based token classification model trained on the annotations are in the range of to 78% F1 (micro) for the coarse-grained labels and around 74% for the fine-grained entity classes. When applying the classifier to gold spans and only predicting the labels, F1 increases to over 88%, giving evidence that the model mostly struggles with the span identification part, not the labelling of the entity classes.

2.2.3. Situation Entities (SE)

Motivation Situation Entity (SE) types (Smith, 2003) are clause-level aspectual annotations, covering eventualities (States, Events, Reports), general statives (Generics, Generalising sentences) and abstract entities (Facts, Propositions). In addition, the classification scheme considers two types of speech acts (Questions, Imperatives). SE types have been shown to support the analysis of argumentation structure (Becker et al., 2016). In addition, we argue that this type of information offers new possibilities for the analysis of political communication, for example, for analysing *epistemological bias* (Recasens et al., 2013), referring to the (un)certainty of a proposition, i.e., whether something has been presented as a universally accepted truth or simply as a belief. Another potential application is based on the identification of generic and generalising statements, which have often been associated with stereotypes in the literature (Geurts, 1985; Leslie, 2014; Radden, 2009; Novoa et al., 2023; Bosse, 2024; Ralston, 2024).

Annotation details The SE type annotations include 9 separate classes (States, Events, Report, Generics, Generalising sentences, Facts, Propositions, Questions, Imperatives), where the abstract entities (Facts, Propositions) themselves can evoke SE types. As a result, we use a multilabel setting as our annotation setup. The annotations have been conducted by two trained annotators, both MA students of linguistics. IAA is 0.54 Cohen's κ. For comparison, Becker et al. (2016) report a Cohen's κ of 0.4 on argumentative microtexts and a slightly higher κ of 0.5 on texts from different genres that have been used for annotator training. Looking at the disagreements, we see that the annotators often disagree on the newly introduced label Event-Perfect-State, which has been added by Mavridou (2016) to annotate German clauses in perfect tense that are ambiguous between the SE types State and Event. The percentage agreement between our annotators is 73% (for argumentative microtexts: 52-61%, see Becker et al. (2016, p.26).

Corpus statistics 19,676 SE instances have been labelled by the two annotators. The most frequent SE types are State (12,321), followed by Generic sentences (2,360) and Event (2,145). This is different from what is predicted by the theory for argumentative text genres (Smith, 2005) and what has been empirically observed by Becker et al.

Cooperation			Conflict		
regulatory	*informative*	*consolidating*	*declarative*	*confrontational*	*argumentative-critical*
Macro	Report	Request	Self-representation	Rhetorical-question	Demand
Expressive	Question	Support	Promise	Accusation	Rejection
					Bad-outcome
Subjective-statement			Subjective-statement		

Table 1: Our linguo-pragmatic taxonomy for speech acts in parliamentary discourse, based on and extending Kondratenko et al. (2020).

(2016) for the annotation of argumentative microtexts, where the majority of SE types are general statives. It indicates that the parliamentary debates cannot be described as "100% argumentative" but also serve other functions, such as providing information or for self-representation through highlighting one's own achievements and by attacking the political opponent (also see Kondratenko et al. (2020) on the different functions of communication in parliamentary settings). Future studies combining SE types with functional speech act annotations (§2.2.4) could provide new insights into genre distinctions and communication in institutions.

Baselines We obtain baseline results for a fine-tuned German BERT model for sequence classification, trained on the SE type annotations of around 84% micro F1. For comparison, previous work on German SE type classification, using an RNN with a Gated Recurrent Unit (GRU), report results in the range of 65-72% F1 (macro), with accuracies around 72% on a diverse set of genres (Becker et al., 2017). Results, however, are not directly comparable, due to different evaluation strategies (see Appendix, §A.5 for details).

2.2.4. Speech Acts

Motivation Being able to determine the rhetorical function of speech acts in political communication has great potential for studies in the area of political text analysis, as it will allow us to investigate how different political players interact with each other in order to achieve their goals or legitimise their actions. For example, corpus users can examine which demands are made by the different political parties over time, what promises are made and who supports whom on what topic over time.

Annotation details As a starting point, we take the linguo-pragmatic schema developed by Kondratenko et al. (2020) used in their qualitative study of parliamentary discourse from the Ukrainian parliament (2004–2019). The focus of their work was on investigating communication strategies employed by the different politicians. In their work, they distinguish Cooperation from Conflict communication, with each communication type further divided into three subclasses. Cooperation includes regulatory, informative and consolidating speech acts while Conflict communication consists of declarative, confrontational and argumentative-critical utterances (see Table 1).

We extend their schema by adding another layer with 13 fine-grained speech act types that represent different communicative functions in political discourse. The final type, SUBJECTIVE-STATEMENT, consists of evaluative or opinionated statements that could not unambiguously be assigned to one of the other classes of Cooperation or Conflict communication. We use this class for subjective statements that assess a situation, an event or an action, including judgments and evaluations of people and personal impressions and memories.

Annotating functional speech acts is a challenging task which is reflected in the inter-annotator agreement scores. On the first training batches we observed an IAA of .56 Cohen's κ which increased to .67 κ after training.

Corpus statistics Overall, our corpus includes 12,947 annotated speech acts for 250 speeches. The most frequent ones are SUBJECTIVE STATEMENT (5,992) and REPORT (2,476), followed by MACRO speech acts that inlude meta-communication (987). Also frequent are DEMAND (829) and REQUEST (780). Among the low-frequency speech acts are SUPPORT (85) and REJECTION (79). The distribution of speech act types in GEPADEU is shown in Table 5 in the Appendix.

Baselines Speech act identification in parliamentary debates can be decomposed into two tasks. First, the speeches have to be segmented into speech acts and second, each speech act needs to be labelled. Baselines for segmentation, based on a supervised BERT token classifier, are in the

Cue/Role name	Description	Example
CUE	the cue that evokes the STW event	Merkel spoke$_{Cue}$ to the people.
SOURCE	Source of the STW event	Merkel$_{Source}$ spoke to the people.
MEDIUM	Medium of the STW event	The constitution$_{Medium}$ states ...
MESSAGE	Message / content of the STW event	She said that she would resign$_{Message}$.
TOPIC	Topic of the STW event	Merkel addressed the theme of taxation$_{Topic}$.
EVIDENCE	Evidence for the message	The survey$_{Evidence}$ shows that ...
ADDRESSEE	Addressee of the STW event	Merkel spoke to the people$_{Addressee}$.
PARTICLE	Separated verb prefix or	Merkel schlug$_{Cue}$ vor$_{Particle}$ (proposed) ...
(PTC)	obligatory particle	Merkel stellt sich$_{Particle}$ vor (imagines herself) ...
MULTIWORD	multiword cue	I now give$_{Cue}$ you the floor$_{Multiword}$

Table 2: Overview over our schema for annotating events of **S**peech, **T**hought and **W**riting (STW).

range of 91% F1 (macro). The most difficult part is the identification of tokens that are not part of any speech act (mostly vocatives). For speech act labelling, using the fine-grained taxonomy described above, a transformer-based sequence classification model achieves a micro F1 of around 82%.

2.2.5. Events of Speech, Thought and Writing

Motivation Identifying the speakers and mesages in political text is an essential prerequisite for analysing political communication. The complexity of the task, however, is often underestimated by assuming that the words produced by a speaker reflect only their own point of view. In political debates, political actors often cite and refer to what has been said by others in order to criticise it. It is therefore important to identify speech events, their sources, addressees and the messages conveyed on a particular topic.

Annotation details An important basis for our work is Brunner et al. (2020) who annotate events of speech, thought and writing in narrative German text (historical fictional and non-fictional texts from 1840-1919). Our annotation scheme shares several ideas with Brunner's work but has a somewhat different label inventory, inspired by FrameNet (Baker et al., 1998). We first identify cues that trigger events of speech, thought and writing. In the next step, the annotators mark the arguments of the trigger, including the SOURCE, ADDRESSEE, MESSAGE, MEDIUM, TOPIC and EVIDENCE for the speech event. Table 2 shows examples for the different categories in our schema. The annotations are available for 265 speeches and have been done by three trained annotators with background in linguistics. For details on IAA, please refer to Rehbein et al. (2024).

Corpus statistics The data includes over 7,700 annotated events of speech, thought and writing. Speech event triggers include multiword units like "eine Rede halten" (give a speech). The number of annotated speech events and their roles in the GEPADEU corpus is shown in Table 4.

Baselines As the annotations are similar in structure to frame-semantic annotations, we can train any out-of-the-box Semantic Role Labelling system on our data. Baselines for the system of Conia and Navigli (2020) for the prediction of trigger words are very high with over 97% F1. This is a word sense disambiguation task that, for a set of given speech trigger words, disambiguates whether the word in context evokes a speech event or not. Results for roles are lower, with an average of 81%. As is common for supervised learning tasks, we get high scores for the more frequent roles (SOURCE 84%, MESSAGE 83%) while results for the less frequent roles are lower with an F1 of 66-78%.

2.2.6. Moral Framing

Motivation The final annotation layer in the GEPADEU corpus captures moral rhetoric in the debates, annotating spans of text that include moral framing, together with the narrative roles that participate in each frame. On top of each frame, we encode the Moral Foundation(s) (Haidt et al., 2009; Graham et al., 2013) used to frame the moral message conveyed by the speaker. This type of annotation is relevant for studies of moral framing in political text where, up to now, most work has relied on the use of dictionaries to identify moral values in the discourse (see Frimer et al. (2019); Araque et al. (2020); Carvalho et al. (2020); Araque et al. (2021); Hopp et al. (2021); Wu et al. (2023), *inter alia*). The expression of moral values, however, is highly context-dependent and often implicit, and cannot be understood by looking at word forms without considering their context. This has been demonstrated by Rehbein et al. (2025a) who have shown that dictionary-based predictions of moral foundations do not correlate well with human annotations that have been disambiguated in context. We therefore believe that our annotations provide a valuable new resource for studying moral framing

in German political text.

Annotation details Our schema is based on Moral Foundations Theory (Haidt et al., 2009; Graham et al., 2013) and inspired by ideas from the Narrative Policy Framework (NPF) (Shanahan et al., 2017). We identify (im)moral values, acts, goals and events as text spans in the discourse and assign the corresponding frame type (either MoralValue/ImmoralValue or MoralActOrGoal/ImmoralActOrGoal). In the next step, we look for frame participants that play a role in the moral event and assign them the labels HERO, VICTIM, VILLAIN and BENEFICIARY. Finally, each frame is labelled with its corresponding moral foundation(s) (MF), using the following MF label inventory {CARE, EQUALITY, PROPORTIONALITY, LOYALTY, AUTHORITY, PURITY, LIBERTY, GENERAL-MORAL}.

The frame identifcation step and annotation of narrative roles has been performed by the two trained students of linguistics who also participated in the SE annotation. Each frame instance has been labelled by four coders (the two student annotators, a PhD student and a postdoc, all of them with a degree in linguistics) who assigned a max. of two MF labels to each frame.[4] IAA for the annotation of MFs across the four coders is .54 Fleiss' κ. To put this into context, the creators of the Moral Foundation Twitter Treebank report Kappa scores in the range of .16 to .44 (Fleiss' κ) for a slightly different label set, with a κ of .27 across all coded foundations (Hoover et al., 2020).

Corpus statistics The moral frame annotations are available for 248 speeches in GEPADEU, with 5,465 annotated frame spans (MORALACTORGOAL: 2,1446, IMMORALACTORGOAL: 1,916, MORALVALUE: 343, IMMORALVALUE: 19, POLITICALACTORGOAL: 1,041) and 2,705 annotated narrative roles (HERO: 426, VILLAIN: 1,126, VICTIM: 497, BENEFICIARY: 656), see Table 9 in the Appendix.

Baselines First baselines for the identification of frame spans seem high with a weighted avg. F1 of around 85%. This, however, is misleading as it mostly reflects the high accuracy for non-frame tokens (label: O) with around 93% while the results for the B and I classes are much lower, with F1 scores of roughly 57% (B) and 67% (I). This highlights the difficulty of the task and shows that token-level classification might not be the optimal setup for moral frame identification where moral values are often expressed with implicit language, making it hard for the model to find reliable lexical cues for span detection.

[4]Most (>98%) frames are assigned one MF label only.

Preliminary results for frame type classification and moral foundation classification on gold spans are higher with results around 82% for frame types (i.e., detecting the correct frame type for a given span in a multiclass setting with 6 labels[5]) and around 78% for MF classification (multiclass, multilabel task with 9 labels,[6] (Rehbein et al., 2025b). The prediction of moral roles is work in progress.

2.3. Automatically Labelled Data

In addition to the manually created high-quality annotations, we will release a large dataset of parliamentary debates, augmented with automatically predicted annotations for the different semantic and pragmatic layers, based on the models described above. For that, we use the raw transcripts from the GermaParl Corpus of Plenary Protocols (Blaette, 2017), spanning a time range from 1949 to 2023 and augment the data with newer transcripts that we downloaded from the Open Data website of the German Bundestag.[7] Our silver standard thus covers more than 75 years of parliamentary debates, up to September 24, 2025.[8]

3. Use Cases

We now demonstrate how the rich semantic and pragmatic annotations in GEPADEU can be used for political text analysis. For this purpose, we provide Jupyter notebooks that allow users to import the corpus, extract individual annotation layers, combine annotations, sort the extracted data according to various metadata variables, and export the output as a CSV file.

To illustrate some use cases, we provide examples below.[9] Please note that the manually annotated subcorpus is not a representative sample of the whole population of Bundestag debates. Therefore, the results shown below are meant to illustrate what type of analyses can be conducted, based on our annotations, but no claim is made concerning their representativity.[10]

[5]The four frame types described above, an additional type POLITICALACTORGOAL and the NONE label.

[6]The 7 MFs and the labels GENERAL-MORAL, NONE.

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

[8]The end-to-end prediction of moral frames is work in progress and will be added to the final release.

[9]We provide Jupyter notebooks for each example use case that we will release with the data: https://github.com/umanlp/gepadeu.git.

[10]The final release of our data will represent the whole population, i.e., all Bundestag debates from 1949 until Sep. 2025.

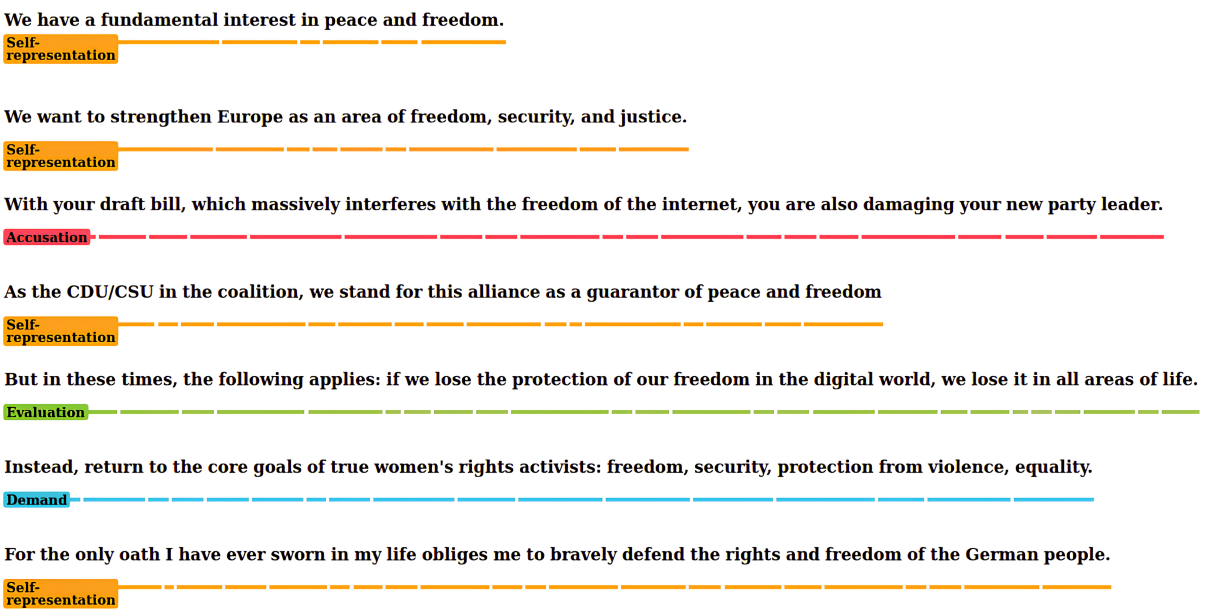

Figure 2: Examples for keyword search for "freedom" with speech act labels (English translation).

3.1. Example 1: Keyword-based Search

In our first example, we search for specific keywords in the debates. For each keyword, we also record its functional speech act to provide more information on its usage. The extracted data can be exported as a csv file, together with the metadata for each instance.

Figure 2 shows example uses of the high value term *freedom*, together with their speech act labels. Although the manually annotated dataset is too small for representative analyses, we find that the rare speech act type SELF-REPRESENTATION, which accounts for about 1% of speech act annotations in our corpus, is overrepresented in the results with over 12% of cases, making it the third most frequent speech act type for the term *freedom*. This is not surprising, since high value concepts such as *freedom* or *justice* – at least in today's Western societies – are generally acknowledged as something desirable and are therefore well suited for positive self-presentation. The keyword-based search, in combination with the functional speech act tags, allows users to identify the pragmatic functions in which specific search terms are used and to compare the usage across parties and political actors.

3.2. Example 2: Moral Foundations

Our next example deals with moral framing in political discourse. We first look at the distribution of moral foundations across political parties and then combine the speech act annotations with moral frames to study moral accusations in political rhetoric.

Moral framing in politics has been the subject of numerous studies in recent years (see, e.g., Fulgoni et al. (2016); Johnson and Goldwasser (2018); Hoover et al. (2020); Araque et al. (2021); Roy and Goldwasser (2021); Beiró et al. (2023), *inter alia*. Many of these studies are based on Moral Foundations Theory (MFT) (Haidt et al., 2009; Graham et al., 2013), a descriptive, pluralist theory of morality rooted in social psychology. MFT assumes the existence of a number of moral intuitions or "gut feelings" that drive moral reasoning and turn it into rationalisation. Knowledge about the "moral intuitions" held by certain populations has been used to investigate a wide range of research questions, including cultural differences in moral values (Wu et al., 2023) or the driving factors behind human attitudes and behaviour, such as COVID-19 vaccine hesitancy (Weinzierl and Harabagiu, 2022). Others examine the use of moral rhetoric in political debates and try to determine when moral rhetoric successfully contributes to the mobilisation of copartisan voters (Lipsitz, 2018; Simonsen and Widmann, 2025; Jung, 2020).

While most of the studies mentioned above either use context-insensitive dictionaries to extract moral values from text (Fulgoni et al., 2016; Jung, 2020; Araque et al., 2021; Beiró et al., 2023; Wu et al., 2023) or predict moral values at the document level (Hoover et al., 2020; Trager et al., 2022; Liscio et al., 2022; Zangari et al., 2025), our conceptualisation of moral frames on the span level that includes narrative roles and moral foundations (see §2.2.6) allows us extract more fine-grained representations of moral framing that are context-sensitive and more informative than document-level labels. In addition, the annotations can be aggre-

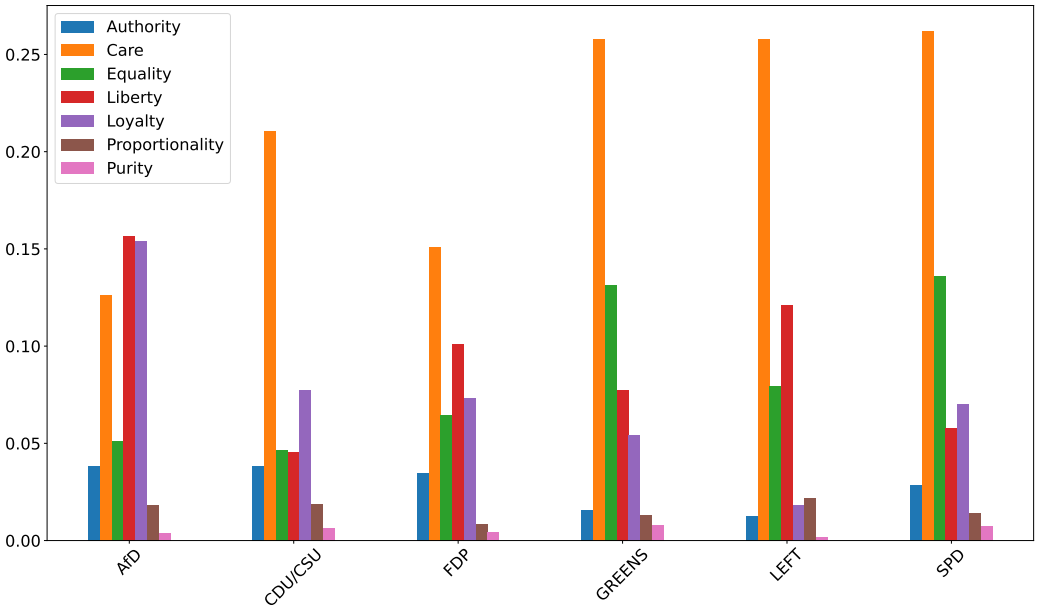

Figure 3: Proportion of each Moral Foundation in the morally framed messages for each party in GEPADEU (normalised by the total number of moral frames for each party).

gated and compared across parties and over time and can also be combined with annotations from other annotation layers.

Moral Foundations per party Figure 3 shows the distribution of moral foundations in GEPADEU for the six parties in the German Bundestag.[11] When looking at the German political landscape (Figure 5 in the Appendix), based on the Chapel Hill Expert Survey (2019) (Jolly et al., 2022), we notice some interesting differences between the parties. The figure plots the left-right economic position of the different parties as well as their ideological orientation for the libertarian-authoritarian dimension. We can see that the three parties positioned mostly to the left on the left-right economic scale, the Left, Greens and SPD, all show a higher usage of CARE frames in their speeches, a Moral Foundation encoding "virtues of kindness, gentleness, and nurturance".[12] Members of the AfD, a right-wing extremist party, frame most of their messages based on the Moral Foundation LOYALTY, which is related to virtues of patriotism and self-sacrifice for the group. In addition, they display the highest proportion of LIBERTY frames, followed by the Left.[13]

Studying moral rhetoric Simonsen and Widmann (2025) investigate moral rhetoric in parliamentary speeches on immigration from eight European countries, using a dictionary-based approach. They provide qualitative insights, based on a small sample of 160 speeches (20 speeches per country), manually coded for pragmatic categories like "moral self-praise", "moral call for action", or "moral attacks on political opponent" (see Table 3 in Simonsen and Widmann (2025, p.13)). Our annotation layers will allow users to conduct similar analyses on larger data by combining the functional speech act classes with moral foundations. We illustrate this, based on the class "moral attacks on political opponent" from Simonsen and Widmann (2025) which we operationalise by combining the speech act tag ACCUSATION with the annotations of Moral Foundations.

In the first step, we extract speeches on immigration, using the regular expressions provided in Simonsen and Widmann (2025) (supplementary materials). We list the English and German terms below:

English: immigr*, migr*, asyl*, refuge*, foreigner*, "guest worker"*

German: immigr*, migrat*, migrant*, migrier*, einwander*, zuwander*, zugewander*, eingewander*, asyl*, flüchtling*, geflücht*, ausländ*, gastarbeit*

A speech is considered relevant if at least three matching tokens are found (either for the same

[11] Please note that the CDU/CSU is the union of two conservative parties, the Christian Democratic Union and Christian Social Union. The two parties are forming a joint parliamentary group in the Bundestag, however, the CSU is only running for election in a single German province, Bavaria.

[12] See the descriptions on the Moral Foundations website: https://moralfoundations.org.

[13] For a brief description of the different Moral Foundations, please refer to the annotation overview.

	Authority	Care	Equality	Liberty	Loyalty	Proport.	Purity	General	total
	%	%	%	%	%	%	%	%	#
AfD	2.0	13.7	11.8	2.0	29.4	0.0	0.0	41.2	51
CDU/CSU	2.1	2.1	2.1	2.1	6.2	0.0	0.0	85.4	48
FDP	0.0	0.0	0.0	0.0	14.3	0.0	0.0	85.7	7
GREENS	0.0	15.4	0.0	7.7	15.4	0.0	0.0	61.5	13
LEFT	0.0	23.5	14.7	14.7	2.9	0.0	0.0	44.1	34
SPD	0.0	3.8	26.9	19.2	7.7	0.0	0.0	42.3	26

Table 3: Total number of morally framed accusations per party in debates on immigration and percentage of frames for each Moral Foundation.

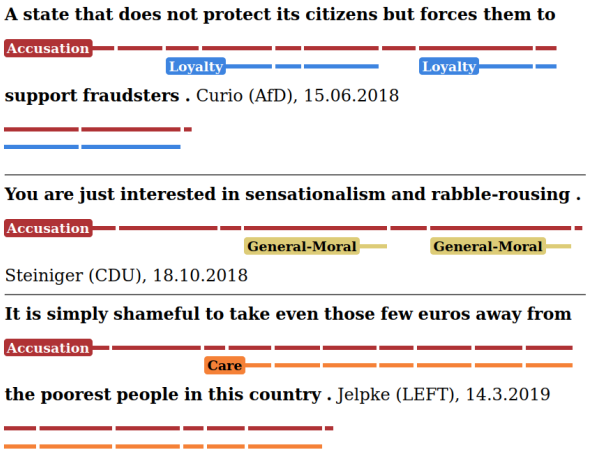

Figure 4: Examples for moral accusations.

or for different regular expressions). Following this approach, we extract 41 (out of 267) speeches from our data.

In the next step, we use the filtered speeches as input and extract all instances labelled as ACCUSATION, based on the speech act layer. Then we extract all moral foundations for the filtered speeches and look for moral spans that overlap with the speech act annotations. As a result, we get a list of accusations made in the context of immigration debates, together with their moral framing (Table 3). The last column in the table shows the raw number of moral accusations (accusations that do not include moral framing are excluded) while the other columns show the percentage of moral accusations for each Moral Foundation. We see a high number of accusing speech acts made by members of the extreme right, AfD. A large percentage of these accusations (29.4%) are framed using the Moral Foundation LOYALTY. A typical example is shown in Figure 4 (top). The conservative union of CDU/CSU also uses a high number of moral attacks, however, most of the accusations cannot clearly be assigned to one of the Moral Foundations but have been coded as GENERAL-MORAL, a label used to refer to general moral principles and values concerning dishonesty, trickery, or selfish behaviour (Figure 4, middle). The last example in Figure 4 (bottom) shows the use of moral framing, based on the CARE foundation, which is frequent in accusations made by members of the LEFT party.

Similar analyses can be done by combining the speaker attribution roles (such as MESSAGE, TOPIC, EVIDENCE) or Situation Entity types (e.g., GENERIC, EVENT) with functional speech acts (for example, ACCUSATION, DEMAND, SUPPORT) and Moral Foundations.

4. Conclusions

In the paper, we introduced GEPADEU, a new resource for analysing German political debates. Our new corpus is richly annotated with semantic and pragmatic information, allowing users to search for named entities and references to social groups, to extract events of speech, thought and writing or to filter the data, according to semantic clause types and functional speech acts. In addition, GEPADEU includes fine-grained annotations of moral framing. The manually annotated corpus includes 267 speeches with more than 214,000 tokens and will be augmented with an automatically annotated dataset that includes all debates from the German Bundestag from 1949 to September 2025.

We show-cased how the different annotation layers can be combined to study the language of politics and to investigate research questions from the political sciences. We hope that our new resource will support future research in the computational political sciences.[14]

5. Limitations

Our work comes with several limitations. Most importantly, while the combined resources of manually labelled and automatically predicted data will

[14] To improve data accessibility, we will make the manually annotated GEPADEU corpus available via ANNIS (Zeldes et al., 2009), a browser-based search and visualisation tool (see `linguistischeempirieannis.uni-muenster.de`).

be representative of the parliamentary discourse in the German Federal Republic, it is clear that the automatically assigned labels will be error-prone. We thus advise researchers who want to work with the data to conduct additional evaluations to confirm that the analyses based on the predicted labels are accurate enough for the planned analyses. It is also not clear how well the classifiers will perform on the older transcripts of the Bundestag, and whether we miss important concepts through language change and topic shifts.

Acknowledgments

The work presented in this paper is funded by the German Research Foundation (DFG) under the UNCOVER project (PO1900/7-1 and RE3536/3-1). We would like to thank our annotators, Lilly Brauner and Florian Ertz, for their dedicated work. We would also like to thank the anonymous reviewers for their constructive feedback.

6. Bibliographical References

Quinn Albaugh, Julie Sevenans, Stuart Soroka, and Peter John Loewen. 2013. The automated coding of policy agendas: A dictionary-based approach. In *The 6th Annual Comparative Agendas Conference*.

Oscar Araque, Lorenzo Gatti, and Kyriaki Kalimeri. 2020. Moralstrength: Exploiting a moral lexicon and embedding similarity for moral foundations prediction. *Knowledge-Based Systems*, 191:105184.

Oscar Araque, Lorenzo Gatti, and Kyriaki Kalimeri. 2021. The language of liberty: A preliminary study. In *Companion Proceedings of the Web Conference 2021*, WWW '21, page 623–626, New York, NY, USA. Association for Computing Machinery.

M. Atari, J. Haidt, J. Graham, S. Koleva, S. T. Stevens, and M. Dehghani. 2023. Morality beyond the weird: How the nomological network of morality varies across cultures. *Journal of Personality and Social Psychology*, 5(125):1157–1188.

Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. In *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Volume 1*, pages 86–90, Montreal, Quebec, Canada. Association for Computational Linguistics.

Maria Becker, Alexis Palmer, and Anette Frank. 2016. Argumentative texts and clause types. In *Proceedings of the Third Workshop on Argument Mining (ArgMining2016)*, pages 21–30, Berlin, Germany. Association for Computational Linguistics.

Maria Becker, Michael Staniek, Vivi Nastase, Alexis Palmer, and Anette Frank. 2017. Classifying semantic clause types: Modeling context and genre characteristics with recurrent neural networks and attention. In *Proceedings of the 6th Joint Conference on Lexical and Computational Semantics (*SEM 2017)*, pages 230–240, Vancouver, Canada. Association for Computational Linguistics.

Mariano Beiró, Jacopo D'Ignazi, Victoria Bustos, María Prado, and Kyriaki Kalimeri. 2023. Moral narratives around the vaccination debate on facebook. In *Proceedings of the ACM Web Conference 2023*, WWW'23), pages 4134–4141.

Anne Bosse. 2024. Stereotyping and generics. *Inquiry: An Interdisciplinary Journal of Philosophy*, 67(10):3876–3892.

Annelen Brunner, Stefan Engelberg, Fotis Jannidis, Ngoc Duyen Tanja Tu, and Lukas Weimer. 2020. Corpus REDEWIEDERGABE. In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC), May 11-16, 2020, Palais du Pharo, Marseille, France*, pages 803 – 812, Paris. European Language Resources Association.

Flavio Carvalho, Helder Yukio Okuno, Lais Baroni, and Gustavo Guedes. 2020. A brazilian portuguese moral foundations dictionary for fake news classification. In *The 39th International Conference of the Chilean Computer Science Society (SCCC)*, pages 1–5.

Simone Conia and Roberto Navigli. 2020. Bridging the gap in multilingual semantic role labeling: a language-agnostic approach. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 1396–1410, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Alexander Philip Dawid and Allan M. Skene. 1979. Maximum likelihood estimation of observer errorrates using the EM algorithm. *Applied Statistics*, 1(28):20–28.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human*

Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Grace E. Fielder and Theresa Catalano. 2017. Othering others: Right-wing populism in uk media discourse on "new" immigration. In Jan Chovanec and Katarzyna Molek-Kozakowska, editors, *Representing the Other in European Media Discourses*, pages 207–234. Amsterdam: John Benjamins.

Jeremy A. Frimer, Reihane Boghrati, Jonathan Haidt, Jesse Graham, and Morteza Dehgani. 2019. Moral foundations dictionary for linguistic analyses 2.0. Unpublished manuscript.

Dean Fulgoni, Jordan Carpenter, Lyle Ungar, and Daniel Preoţiuc-Pietro. 2016. An empirical exploration of moral foundations theory in partisan news sources. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 3730–3736, Portorož, Slovenia. European Language Resources Association (ELRA).

Bart Geurts. 1985. Generics. *Journal of Semantics*, 3(4):247–255.

Goran Glavaš, Federico Nanni, and Simone Paolo Ponzetto. 2017. Unsupervised cross-lingual scaling of political texts. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 688–693, Valencia, Spain. Association for Computational Linguistics.

Jesse Graham, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, and Peter H. Ditto. 2013. Chapter two - moral foundations theory: The pragmatic validity of moral pluralism. In Patricia Devine and Ashby Plant, editors, *Advances in Experimental Social Psychology*, volume 47, pages 55–130. Academic Press.

Johann Gründl. 2022. Populist ideas on social media: A dictionary-based measurement of populist communication. *New Media & Society*, 24(6):1481–1499.

Jonathan Haidt, Jesse Graham, and Conrad Joseph. 2009. Above and below left–right: Ideological narratives and moral foundations. *Psychological Inquiry*, 20(2-3):110–119.

Joe Hoover, Gwenyth Portillo-Wightman, Leigh Yeh, Shreya Havaldar, Aida Mostafazadeh Davani, Ying Lin, Brendan Kennedy, Mohammad Atari, Zahra Kamel, Madelyn Mendlen, Gabriela Moreno, Christina Park, Tingyee E. Chang, Jenna Chin, Christian Leong, Jun Yen Leung, Arineh Mirinjian, and Morteza Dehghani. 2020. Moral foundations twitter corpus: A collection of 35k tweets annotated for moral sentiment. *Social Psychological and Personality Science*, 11(8):1057–1071.

Frederic R Hopp, Jacob T Fisher, Devin Cornell, Richard Huskey, and René Weber. 2021. The extended moral foundations dictionary (emfd): Development and applications of a crowd-sourced approach to extracting moral intuitions from text. *Behavior Research Methods*, 53:232–246.

Lena M. Huber. 2022. Beyond Policy: The Use of Social Group Appeals in Party Communication. *Political Communication*, 39(3):293–310. Publisher: Routledge _eprint: https://doi.org/10.1080/10584609.2021.1998264.

Lena M. Huber and Matthias Haselmayer. 2024. Promising links: How parties combine policy issues with group appeals. *West European Politics*, pages 1–28.

R. Iyer, S. Koleva, J. Graham, P. Ditto, and J. Haidt. 2012. Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PLoS ONE*, 8(7).

Jan Jagers and Stefaan Walgrave. 2007. Populism as Political Communication Style: An Empirical Study of Political Parties' Discourse in Belgium. *European Journal of Political Research*, 46(3):319–345.

Kristen Johnson and Dan Goldwasser. 2018. Classification of moral foundations in microblog political discourse. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 720–730, Melbourne, Australia. Association for Computational Linguistics.

Seth Jolly, Ryan Bakker, Liesbet Hooghe, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen, and Milada Anna Vachudova. 2022. Chapel Hill Expert Survey Trend File, 1999-2019. *Electoral Studies*, 75.

Jae-Hee Jung. 2020. The mobilizing effect of parties' moral rhetoric. *American Journal of Political Science*, 64(2):341–355.

Natalia V. Kondratenko, Anastasiia A. Kiselova, and Liubov V. Zavalska. 2020. Strategies and tactics of communication in parliamentary discourse. *studies about languages*, 36:17–29.

Michael Laver, Kenneth Benoit, and John Garry. 2003. Extracting policy positions from political texts using words as data. *American Political Science Review*, 97(2).

Sarah-Jane Leslie. 2014. Carving up the social world with generics. *Oxford studies in experimental philosophy*, 1:208–232.

Keena Lipsitz. 2018. Playing with emotions: The effect of moral appeals in elite rhetoric. *Political Behavior*, 40:57–78.

Enrico Liscio, Alin E. Dondera, Andrei Geadău, Catholijn M. Jonker, and Pradeep K. Murukannaiah. 2022. Cross-domain classification of moral values. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 2727–2745, Seattle, United States. Association for Computational Linguistics.

Kleio-Isidora Mavridou. 2016. *Situation Entity Types: a Cross-linguistic Corpus Study and a Comparison of Automatic Classifiers*. Ph.D. thesis, Institute for Computational Linguistics at Saarland University, Germany.

Miller McPherson, Lynn Smith-Lovin, and James M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444.

Gustavo Novoa, Margaret Echelbarger, Andrew Gelman, and Susan A. Gelman. 2023. Generically partisan: Polarization in political communication. *Proceedings of the National Academy of Sciences*, 120(47):e2309361120.

Günter Radden. 2009. Generic reference in english: A metonymic and conceptual blending analysis. In Klaus-Uwe Panther, Linda L. Thornburg, and Antonio Barcelona, editors, *Metonymy and Metaphor in Grammar*, pages 199–228. John Benjamins Publishing Company.

Tom Ralston. 2024. Reconceptualising the psychological theory of generics. *Philosophical Studies*, 181:2973–2995.

Marta Recasens, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. 2013. Linguistic models for analyzing and detecting biased language. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1650–1659, Sofia, Bulgaria. Association for Computational Linguistics.

Ines Rehbein, Lilly Brauner, Florian Ertz, Ines Reinig, and Simone Ponzetto. 2025a. Moral reckoning: How reliable are dictionary-based methods for examining morality in text? In *Proceedings of the 5th International Conference on Natural Language Processing for Digital Humanities*, pages 232–250, Albuquerque, USA. Association for Computational Linguistics.

Ines Rehbein, Ines Reinig, and Simone Paolo Ponzetto. 2025b. Moral framing in politics (MFiP): A new resource and models for moral framing. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 34643–34663, Suzhou, China. Association for Computational Linguistics.

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.

Shamik Roy and Dan Goldwasser. 2021. Analysis of nuanced stances and sentiment towards entities of US politicians through the lens of moral foundation theory. In *Proceedings of the Ninth International Workshop on Natural Language Processing for Social Media*, pages 1–13, Online. Association for Computational Linguistics.

Josef Ruppenhofer, Ines Rehbein, and Carolina Flinz. 2020. Fine-grained named entity annotations for German biographic interviews. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 4605–4614, Marseille, France. European Language Resources Association.

Elizabeth A. Shanahan, Michael D. Jones, Mark K. Mcbeth, and Claudio M. Radaelli. 2017. The Narrative Policy Framework. In C.M. Weible and P.A. Sabatier, editors, *The Theories of the Policy Process*, pages 173–213. Boulder, CO: Westview Press.

Kristina Bakkær Simonsen and Tobias Widmann. 2025. When do political parties moralize?: A cross-national study of the use of moral language in political communication on immigration. *British Journal of Political Science*, 55:e33.

Jonathan Slapin and Sven-Oliver Proksch. 2008. A scaling model for estimating time-series party positions from texts. *American Journal of Political Science*, 52(3):705–772.

Carlota S. Smith. 2003. *Modes of Discourse: The Local Structure of Texts*. Cambridge Studies in Linguistics. Cambridge University Press.

Carlota S. Smith. 2005. *Aspectual entities and tense in discourse*, pages 223–237. Springer.

Jean-François Staszak. 2009. Other/otherness. In Rob Kitchin and Nigel Thrift, editors, *International Encyclopedia of Human Geography*, pages 43–47. Elsevier, Oxford.

Michael Thau. 2019. How political parties use group-based appeals: Evidence from britain 1964–2015. *Political Studies*, 67(1):63–82.

Michael Thau. 2024. The group appeal strategy: Beyond the policy perspective on party electoral success. *Political Studies*, 72(4):1653–1674.

Jackson Trager, Alireza S. Ziabari, Aida Mostafazadeh Davani, Preni Golazizian, Farzan Karimi-Malekabadi, Ali Omrani, Zhihe Li, Brendan Kennedy, Nils Karl Reimer, Melissa Reyes, Kelsey Cheng, Mellow Wei, Christina Merrifield, Arta Khosravi, Evans Alvarez, and Morteza Dehghani. 2022. The Moral Foundations Reddit Corpus. ArXiv preprint.

Sarah Wagner, L. Constantin Wurthmann, and J. Philipp Thomeczek. 2023. Bridging left and right? How Sahra Wagenknecht could change the German party landscape. *Politische Vierteljahresschrift*, 64.

Maxwell A. Weinzierl and Sanda M. Harabagiu. 2022. From hesitancy framings to vaccine hesitancy profiles: A journey of stance, ontological commitments and moral foundations. *Proceedings of the International AAAI Conference on Web and Social Media*, 16(1):1087–1097.

Werner Wirth, Martin Wettstein, Dominique Wirz, Nicole Ernst, Florin Büchel, Anne Schulz, and Frank Esser et al. 2019. *Codebook: NCCR democracy Module II: The Appeal of populist Ideas and Messages*. Unpublished paper.

Winston Wu, Lu Wang, and Rada Mihalcea. 2023. Cross-cultural analysis of human values, morals, and biases in folk tales. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 5113–5125, Singapore. Association for Computational Linguistics.

Lorenzo Zangari, Candida M. Greco, Davide Picca, and Andrea Tagarelli. 2025. ME2-BERT: Are events and emotions what you need for moral foundation prediction? In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 9516–9532, Abu Dhabi, UAE. Association for Computational Linguistics.

Amir Zeldes, Anke Lüdeling, Julia Ritz, and Christian Chiarcos. 2009. Annis: a search tool for multi-layer annotated corpora. In *Proceedings of Corpus Linguistics*.

7. Language Resource References

Andreas Blaette. 2017. *GermaParl. Corpus of Plenary Protocols of the German Bundestag*. PolMine. 2023.

Weischedel, Ralph and Palmer, Martha and Marcus, Mitchell and Hovy, Eduard and Pradhan, Sameer and Ramshaw, Lance and Xue, Nianwen and Taylor, Ann and Kaufmen, Jeff and Franchini, Michelle and El-Bachouti, Mohammed and Belvin, Robert and Houston, Ann. 2013. *OntoNotes Release 5.0*. Linguistic Data Consortium. 5.0, ISLRN ISLRN 151-738-649-048-2.

Supplementary Material

A. Appendix

A.1. Named Entities

Table 7 below illustrates the NER inventory that we use to annotate a subset of the GePaDeU corpus. The set of labels is adapted from Ruppenhofer et al. (2020) and includes 5 coarse-grained domains (Agentive, Spatial, Human-created, Times and events, Numeric expressions) with 31 fine-grained labels, adapted from the OntoNotes scheme (Weischedel et al., 2013). Not all labels occur in our data, for example, we did not observe any instances of non-human creatures in the parliamentary debates.

A.2. MoPE

Table 8 shows the hierarchical annotation scheme used to code **M**entions **o**f the **P**eople and the **E**lite (MoPE). The table also displays the distribution of labels in the GePaDeU, amounting to 7,351 annotated mentions (including GPEs).

A.3. Speaker Attribution

Table 2 (main paper) gives an overview of the annotation scheme for coding events of Speech, Thought and Writing (STW) in parliamentary debates and Table 4 below shows the distribution of speech triggers and roles in the corpus.

Cue/Role	Freq.	Avg. length
Cue	7,706	1.1
Source	4,663	1.7
Message	4,578	9.7
Topic	1,188	5.4
Addressee	717	3.2
Particle	561	1.0
Medium	321	3.2
Evidence	151	4.3

Table 4: Statistics for the speaker attribution labels in GePaDeU (Cue also includes multiword cues; 773 of the 7,706 cues are Multiword cues).

A.4. Speech acts

Table 1 (main paper) illustrates the linguo-pragmatic taxonomy for speech acts in parliamentary debates, adapted from Kondratenko et al. (2020). Our final schema comprises a set of 14 functional speech act labels to encode the different types of speech acts in cooperation and conflict communication. The distribution of speech acts in the GePaDeU is shown in Table 5.

Speech act label	No. instances
Subjective statement	5,992
Report	2,476
Macro	987
Demand	829
Request	780
Accusation	709
Expressive	538
Question-All	414
Bad-outcome	202
Self-representation	163
Promise	162
Support	85
Rejection	79
I-S-Humour	35
Total	**12,947**

Table 5: Number of annotated speech act instances per label in the GePaDeU. Please note that instances can have more than one speech act label. The number of speech act labels assigned to the 12,947 spech acts is 13,451.

A.5. Situation Entities

Table 6 shows the different Situation Entity Types and their distribution in the GePaDeU. The columns A1, A2 show the number of instances annotated by coder 1 and 2, and the next column displays the average number of instances for each SE type. The last column reports results for our baseline system per class (micro F1).

	Situation Entities			
Entity	A1	A2	Avg.	F1
State	12,987	11,655	12,321.0	0.92
Generic	2,012	2,708	2,360.0	0.73
Event	1,406	2,885	2,145.5	0.77
Generalis	942	1,062	1,002.0	0.59
Evt-Pfc-St	1,286	259	772.5	0.70
Question	447	444	445.5	0.97
Imperative	345	335	340.0	0.94
Report	251	328	289.5	0.82
Total	19,676	19,676	19,676	0.84
	Abstract Entities			
Entity	A1	A2	Avg.	F1
Proposition	315	222	268.5	n.a.
Fact	301	123	212.0	n.a.

Table 6: Distribution of SE types in the GePaDeU corpus and F1 (micro) per class.

	NER tag	Description	Examples
Agentive	PER•	Person names, including nick names	Angela Merkel; Trump
	ORG•	Organization	SPD; Likud
	CREAT	Creatures, non-human	Gollum
	NRP•	National, religious , political (and other identity) categories	Israeli; English
	MED	the media products / channels of newspapers and companies like Twitter, the Jerusalem Post, etc.	NY Times; Facebook
	GPE•	geopolitical entities such as countries, states and cities	Germany; Berlin
Spatial	LOC•	locations that are not administrative units (i.e. not GPEs)	Negev, Black Sea
	ADD	Addresses in the physical world	R5 6-13, 68176 Mannheim
	URL	Addresses in virtual domains	http://www.lrec-conf.org
Hum. creat.	FAC•	Facilities	Golden Gate Bridge
	LAN•	Natural languages and their varieties	Hebrew; Viennese
	LAW•	laws, ordinances, treaties etc.	Oslo Accords
	ART	works of art such as movie, song, book titles	Pulp Fiction; Macbeth
	PRODUCT•	commercial products	Boeing 737 MAX
Times and events	EVT•	Events	Anschluss; Six-Day War
	PROJ	names of projects	Fridays For Furture
	TIME•	temporal locations	now, 10 o'clock
	DATE•	dates, a special subset of temporal locations	September 1, 1939; this month
	DUR	durations	7 hours, four days
	FREQ	information about frequency of events	twice; four times a day
	SORD	reference to place of repeated event in serial order	for the 2nd time
	AGE	age specificiations	10 years old
Numeric expr.	CARDINAL•	cardinal numbers	10, ten
	ORDINAL•	ordinal numbers	1.; first
	PERC•	percentages	90%; ten percent
	FRAC	fractions	3/5
	QUANT•	combinations of numbers and units of measurements	7 kilos
	RATE	distribution of one set of units over another unit	60 km/h
	MON•	amounts of money	S70.000; 3 Piaster
	SCORE	specifications of sports and other scores	6:2,6:1 [tennis]
	MISC	other NEs not covered by the above categories	Nobel Peace Prize

Table 7: Table illustrating the inventory of NER labels in the GePaDeU corpus (adapted from Ruppenhofer et al. (2020)).

A.6. Moral framing

Below we provide a short description of the moral foundations, adapted from the MFT website.[15]

Care: This foundation is related to our long evolution as mammals with attachment systems and an ability to feel (and dislike) the pain of others. It underlies the virtues of kindness, gentleness, and nurturance.

Fairness: This foundation is related to the evolutionary process of reciprocal altruism. It underlies the virtues of justice and rights.

In recent work (Atari et al., 2023), the Fairness foundation was split into two new foundations, Equality and Proportionality, as it was found that politically left-leaning individuals more strongly endorse values of Equality while more conservative individuals prefer the notion of proportionality.

Equality: Equality is defined as "Intuitions about equal treatment and equal outcome for individuals."

Proportionality: Proportionality is defined as "Intuitions about individuals getting rewarded in proportion to their merit or contribution."

Loyalty: This foundation is related to our long history as tribal creatures able to form shifting coalitions. It is active anytime people feel that it's "one for all and all for one." It underlies the virtues of patriotism and self-sacrifice for the group.

Authority: This foundation was shaped by our long primate history of hierarchical social interactions. It underlies virtues of leadership and follow-

[15] https://moralfoundations.org/.

L1	*Elite* E						*People* P		
L2	*Person* P			*Organisation* O			–		
L3	Domain:	Label:	#	Domain:	Label:	#	Domain:	Label:	#
	Politics	EPPOL	1,481	Politics	EOPOL	2,017	Nation	PNAT	193
	Economy	EPECON	13	Economy	EOECON	90	Ethnicity /religion	PETH	144
	Finance	EPFIN	7	Finance	EOFIN	115	Profession /function	PFUN	1,112
	Media	EPMED	6	Media	EOMED	32	Age	PAGE	384
	Science	EPSCI	46	Science	EOSCI	17	Social variables	PSOC	227
	Religion	EPREL	1	Religion	EOREL	3	(gender/class/...)		
	Culture	EPCULT	15	Culture	EOCULT	0	Generic	PGEN	219
	Military	EPMIL	5	Military	EOMIL	147			
	NGOs	EPNGO	5	NGOs	EONGO	40			
	Movements	EPMOV	13	Movements	EOMOV	11			
O:	geo-political entity GPE		1,008						

Table 8: Hierarchical annotation of references to *The People* and *The Elite* and their distribution in the GEPADEU corpus.

ership, including deference to prestigious authority figures and respect for traditions.

Purity: This foundation was shaped by the psychology of disgust and contamination. It underlies notions of striving to live in an elevated, less carnal, more noble, and more "natural" way (often present in religious narratives). This foundation underlies the widespread idea that the body is a temple that can be desecrated by immoral activities and contaminants (an idea not unique to religious traditions). It underlies the virtues of self-discipline, self-improvement, naturalness, and spirituality.

Liberty: This foundation is about the feelings of reactance and resentment people feel toward those who dominate them and restrict their liberty. Its intuitions are often in tension with those of the authority foundation. The hatred of bullies and dominators motivates people to come together, in solidarity, to oppose or take down the oppressor.

The last foundation is not yet considered as part of the moral foundations but often discussed as a plausible candidate (Iyer et al., 2012). LIBERTY is often used to frame political arguments, we therefore include it in our annotations.

Table 9 shows the distribution of Moral Foundations (MF) in our data. Each MF has been coded by four annotators. We release the individual annotations for each coder and a majority vote. In addition, we have added the label predicted by the Dawid-Skene model of annotation (Dawid and Skene, 1979).

Frame types	**freq.**	**MF**	**freq.**
ImmoralValue	19	Purity	24
MoralValue	343	Proportionality	72
PoliticalActOrGoal	1,041	Authority	129
ImmoralActOrGoal	1,916	Equality	364
MoralActOrGoal	2,146	Loyalty	375
		Liberty	396
		Care	981
		General-Moral	2,176
		None	1,041
Total	**5,465**		**5,558**

Table 9: Distribution of Moral Foundations in the GEPADEU corpus (note that each frame can have more than one MF; PoliticalActOrGoal frames have MF "None").

Name	**Description**
doc_id	document name
speech_id	speech id
speaker	speaker's last name
party	speaker affiliation
date	date the speech was held
year	year the speech was held
term	term the speech was held
session	session the speech was held
gov_opp	speaker was in government| opposition when the speech was held

Table 10: Available metadata for GEPADEU.

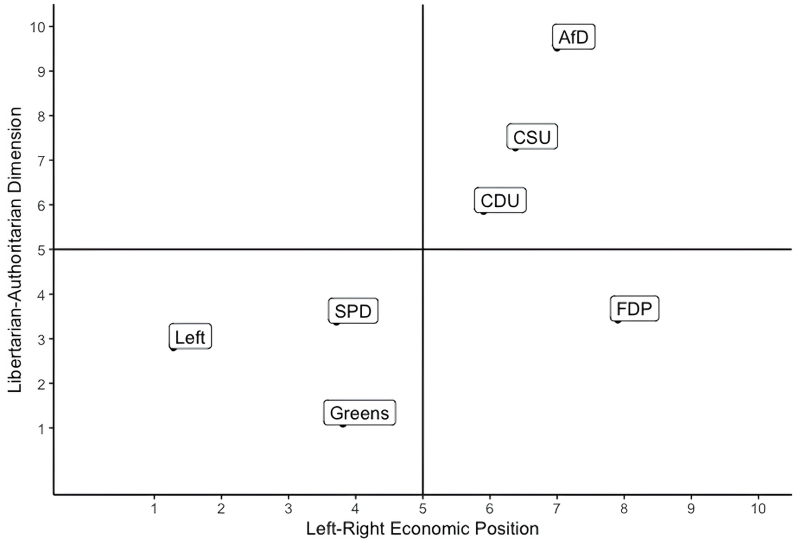

Figure 5: Germany's political landscape according to the Chapel Hill Expert Survey (2019) (image taken from Wagner et al. (2023)).