

---

# Annotation of References to *The Elite* and *The People* in Parliamentary Debates

---

Ines Rehbein  
Version 1.0

## Abstract

This paper describes the annotation of references to *The Elite* and *The People* in debates of the German Bundestag. Detecting references to instances of *The Elite* and *The People* is an important first step towards automatically identifying populist rhetoric in political texts. Under *The Elite*, we include all individuals, organizations, institutions, and groups that have some impact on politics and society. In addition, we mark all references to instances of the people, i.e., groups in the population, that are referenced in the political discourse.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Classification Scheme . . . . .	2
<b>2</b>	<b>Description of the classes</b>	<b>3</b>
2.1	Elite . . . . .	3
2.2	The People . . . . .	6
<b>3</b>	<b>Some Practical Tips</b>	<b>7</b>
3.1	Specific syntactic constructions . . . . .	7
3.2	Special rules and idiosyncratic cases . . . . .	9

Level I	<i>Elite E</i>				<i>People P</i>	
Level II	<i>Person P</i>		<i>Organ O</i>		–	
Level III	<i>Subclasses:</i>	<i>Label:</i>	<i>Subclasses:</i>	<i>Label:</i>	<i>Subclasses:</i>	<i>Label:</i>
	Politics	EPOL	Politics	EOPOL	Nation	PNAT
	Economy	EP ECON	Economy	EOECON	Ethnicity/Religion	PETH
	Finance	EPFINANCE	Finance	EOFINANCE	Funktion/Beruf	PFUNC
	Media	EPMEDIA	Media	EOMEDIA	Alter	PAGE
	Science	EPSCI	Science	EOSCI	Soziales	PSOC
	Religion	EPREL	Religion	EOREL	Generisch	PGEN
	Culture	EPCULT	Culture	EOCULT		
	Military	EPMIL	Military	EOMIL		
	NGOs/Associations	EPNGO	NGOs/Associations	EONGO		
	Movements	EPMOV	Movements	EOMOV		
	<i>Other:</i>					
	I, me	EPOWN	geo-political entities	GPE		

Table 1: Overview of the classification scheme.

## 1 Introduction

### 1.1 Classification Scheme

Our schema builds upon and extends the categories in the codebook of [Wirth et al., 2019, p.12].<sup>1</sup>

#### Classes:

- I Elite vs People: At the top level, we distinguish between instances of *Elite* and the *People*. We use a very broad notion of *Elite*, encompassing all individuals and groups that have some influence in politics and society. This includes individuals, groups of individuals, organizations, associations, and institutions that play an influential role in politics and society and thus shape and contribute to shaping society.
- II Person vs Organ: On the next level we distinguish between references to individuals and groups of individuals (Angela Merkel; Merkel and Scholz; the colleagues from the CDU) on the one hand and references to political/governmental bodies and institutions (the SPD; the EU; the investigative commission), associations, organizations from the economy and society, institutionalized social organizations and movements on the other hand. The term ORGAN is used here (somewhat imprecisely) as an umbrella term for all the organizational forms listed above.
- III Domains (Elite): At level III, both classes of elite (person, organ) are further subdivided into more fine-grained sub-domains:
- Politics (CDU, the Greens, the EU)
  - Economy (BMW, Amazon, Wirecard)
  - Finance (Lehman Brothers, the Volksbanken)
  - Media (FAZ, Twitter, Facebook, ARD)
  - Science (Uni Mannheim, CERN, Fraunhofer-Gesellschaft)
  - Religion (the Pope, the catholic church, the Central Council of Muslims)
  - Culture (Deutsche Oper, Elfriede Jellinek, Madonna)

<sup>1</sup> <https://osf.io/2z3dk/>

- Military (Bundeswehr, the Air Force)
- Associations and organisations (ver.di, DGB, ADAC, BUND, AI, Red Cross, ATTAC)
- Movements (Antifa, Fridays for Future, Querdenker)

II THE PEOPLE: Die Klasse VOLK umfasst Referenzen auf (Unter-)Gruppen der Bevölkerung (die Deutschen, die Steuerzahler) oder generische Referenzen auf das Volk oder Untergruppen der Bevölkerung (Gabi Mustermann; John Q. Public; der kleine Mann von der Straße).

III Domains (People): We distinguish the following domains:

- Nation (die Deutschen, der Russe)
- Ethnicity & Religion (die Türkendeutschen, die Juden, die Christen)
- Function & Beruf (die Steuerzahler, die Professoren)
- Age (The thirty-somethings, teenager, pensioners)
- Social variables (the working class, women, transgender)

Other classes: We also annotate the two classes:

1. 1st person singular personal pronouns (*I*, *me*) (EPOWN) that refer to the speaker and can be assigned the class Elite-Person-Own
2. geo-political entities (GPE)

Wiktionary describes GPEs as

“A geographical area which is associated with some sort of political structure.”

We include countries, states, cities, multinational geo-political entities such as Europe, etc.

## 2 Description of the classes

### 2.1 Elite

As elite, we consider persons, groups of persons, organizations, associations and institutions that play an influential role in politics and society and thus shape and help shape society. These include power elites, economic, legal and cultural elites, and educational elites.

We divide members of the elite into PERSONS and ORGANS.

#### 2.1.1 Persons

All references to specific people that are part of the elite (individuals as well as groups of people) fall into this class.

#### 2.1.2 Organs

This class includes all mentions of state and political organs and institutions (the SPD; the EU; the Investigation Commission), associations (ver.di, ADAC), organizations from economy and society (Siemens; Deutsche Oper), institutionalized social organizations (HU Berlin; IDS Mannheim; the BVG) and movements (Fridays for Future, Querdenker)

on the other hand. The term “organ” is used here (somewhat imprecisely) as an umbrella term for all the organizational forms listed above.

We further divide both classes (people and organs) into the following domains:

Politics, Economy, Finance, Associations, Mass Media, Science, Culture, Religion, Military, Movements

Politics:

Description	Examples
Politicians	Merkel, Özdemir, the foreign minister
Parties	CDU, FDP, The Greens
Coalitions	Groko, Ampel, Jamaika
Justicial system	BGH, Federal Constitutional Court
Committees & Boards	commission of inquiry, Central Committee
Political organs	federal and state governments, parliament
Supranational institutions	EZB, EU, NATO, UN

Economy:

Description	Examples
Concerns	Google, Mercedes, BASF
Employers' associations	BDI, Gesamtmetall
Unions	DGB, ver.di, IG Metall
trade associations	Industriemeisterverband Deutschland

Finance:

Description	Examples
Banks	Lehman Brothers, Deutsche Bank
Organisations & trusts	World Economic Forum, IWF
Bankers, brokers & fonds manager	

Mass media:

Description	Examples
Print media	FAZ, taz, Bild, NYT, WSJ
Radio	Deutschlandfunk, SWR2, HR3
Social media	Twitter, Facebook, Instagram, Tik-Tok

Military:

Description	Examples
Armies	Bundeswehr, US Army
Military leadership	inspector general, major general

Science:

Description	Examples
Universities	Uni Saarland, HU Berlin, Sorbonne
Research centers	Max-Planck-Gesellschaft, RKI, CERN
Researchers	Albert Einstein, Benjamin List, Marie Curie (Nobel laureates)
Experts	Christian Drosten, Lothar Wieler
Scientific associations	Leibniz-Gemeinschaft, GSCL, DGfS

Companies in the medical sector, such as BionTech SE, are not considered as scientific institutions but are assigned to the economy domain.

Culture:

Description	Examples
Artists	Herta Müller, Lion Feuchtwanger, Picasso, Beatles
Top-level sports	Roger Federer, Cristiano Ronaldo, Kimmich
Celebrities	Kayne West, Ellen Degeneres
Academies	Goethe-Institut, Akademie of Arts

Only for renowned, distinguished persons with a high social standing (Nobel Prize, Oscar, Pulitzer, other cultural awards) as well as top athletes and celebrities.

We do not include museums and art sites.

Religion:

Description	Examples
Religious leaders	Pope, Ayatollah, Dalai Lama
Religious organisations	EKD, Central Council of Jews, Milli Görüş

We do not consider the Muslim Brotherhood as a religious organisation but as a movement.

References to places of worship such as the Şehitlik Mosque, the Cologne Cathedral or the New Synagogue Berlin are not included in the annotation.

Associations:

Description	Examples
NGOs	Amnesty International, Greenpeace, WWF
Clubs and societies	Bayern München, carnival society, ADAC
Sport associations	German soccer federation, tennis federation
Welfare organisations & charities	AWO, Oxfam, Sisters of Mercy

This class covers non-commercial civil society advocacy groups and organizations. This includes NGOs that “through their activities seek to alleviate suffering, represent the interests of the poor in the public sphere, protect the environment, provide basic social services, or initiate action for development projects” (definition World Bank, see Wikipedia). Furthermore, sports clubs and other non-profit organizations such as charities (e.g., the Workers’ Welfare Association) fall into this category.

We do not annotate the rabbit breeders’ club of East Cupcake, but only those associations that can be counted among the elite due to their social influence.

Note that we assign unions to this domain (and not to the Economy domain).

Movements:

Description	Examples
Environmental movements	Fridays for Future
Action groups	Gorleben, Startbahn West
Civil rights movements	AIM, Black Lives Matter
SM campaigns	#MeToo
Emancipatory movements	feminist movement
Subcultures	Incel
Conspiracy theorists	QANON, Querdenker
Terror organisations	PKK, IRA, Rote Armee Fraktion (RAF)

As Movements we only count groups without an organizational structure (no clubs, associations, organizations, etc.). Attac, for example, does not belong to the Movement domain but is an NGO with an organisational structure and a scientific advisory board and a membership card.

## 2.2 The People

There is no subdivision of the THE PEOPLE class on the second level. On the third level, we distinguish between the following domains.

**Nation:** Under nation, we include categorizations of instances of the people that are based on the concept of the nation. The crucial focus here is not on language, culture or tradition, but on citizenship in a particular nation-state. Examples are:  
The Germans; the Turks; the French; as well as generic uses such as “the Russian”, “the German” (as in: The German loves wine, women and song).

**Ethnicity & Religion:** This class focusses on belonging / social affiliation to a certain culture, defined by a common language or traditions and customs: PoC, Russian Germans, German Turks, Yazidis, Christians, Muslims, Jews, Slavs, the Jehovah’s Witnesses.

**Function & Profession:** This class covers the reference to a population group based on profession, occupation or function: Irish farmers, the farmers, taxpayers, the unemployed, teachers, the voters, single parents.

**Age:** Reference to a group defined based on age: the older generation, adolescents, mid-twenties, teenagers.

**Social variables:** This category covers variables related to social class (the working class, the top ten thousand), gender (women and girls, transgender), and sexual identity (lesbian, gay).

Generic references: References to individuals such as “Otto Normalverbraucher”, “Jane Doe” and “Gabi Mustermann” who are stereotypical representatives of the characteristics and needs of the population as a whole.

### 3 Some Practical Tips

What (not) to annotate:

We annotate references to persons, expressed mainly by nouns and proper names. Other references, e.g. adjectival references (*German*, *unemployed*) are not included in the annotation (see (1)-(3)). Pronouns (*she*, *we*, *he*, *everyone*, *all*, ...) are also not annotated (see (4)), with the following exceptions:

1. 1st person singular personal pronouns *I*, *me* are assigned to the class EPOWN (Elite-Person-Own) (5).
2. We also annotate phrases with a pronoun as head, if the pronoun is disambiguated by the rest of the phrase (e.g., *we* Members of Parliament; *all of us* here in this august House; *everyone* here in Parliament; etc.) (6)

- (1) The *German* interests must be taken more into account.
- (2) We must promote (*the unemployed* PFUNC)
- (3) She has been *unemployed* for two years.
- (4) (*Mrs. Merkel* EPPOL) said in *her* speech ...
- (5) (*I* EPOWN) started *my* speech by ...
- (6) (*We congressmen* EPPOL) must decide ...

We annotate the entire nominal phrase, including complement clauses and modifiers (and not just the head of the NP), as those provide information that is necessary to disambiguate between the different labels. For illustration, see the examples below:

- (7) (Old [**people**] who have been falling into poverty as a result of pension cuts PAGE)
- (8) ([**Folks**] who have served their country in the military PFUNC)
- (9) ([**People**] who share their native language and cultural background PETH)
- (10) (All [**residents**] with German citizenship PNAT)
- (11) ([**People**] from all over the world PGEN)

#### 3.1 Specific syntactic constructions

Prepositional phrases:

Nominal phrases are often embedded in prepositional phrases. However, we only annotate the nominal phrase, excluding the preposition.

The reign of ✓ EPPOL Angela Merkel lasted 16 years

The reign ✗ EPPOL of Angela Merkel lasted 16 years

## Embedded phrases:

For embedded phrases, we annotate both the embedded phrase and the phrase that embeds it (but, of course, only if those phrases are relevant to our schema).

For example:

  
the former chancellor of the BRD

oder

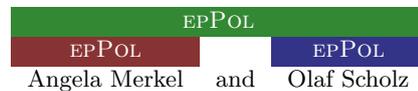
  
Angela Merkel, the former chancellor of the BRD

If the embedded phrase refers to the same entity as the head of the phrase (here *Angela Merkel*), then we do not annotate it separately. See the following example, where *the chancellor* is a reference to *Angela Merkel* and is therefore only annotated once:

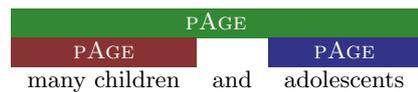
  
Angela Merkel, the former chancellor of the BRD

## Coordination:

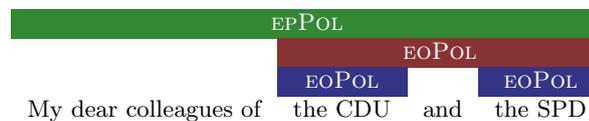
For coordination, we annotate each conjunct separately and also the entire coordinated phrase, e.g.:

  
Angela Merkel and Olaf Scholz

oder:

  
many children and adolescents

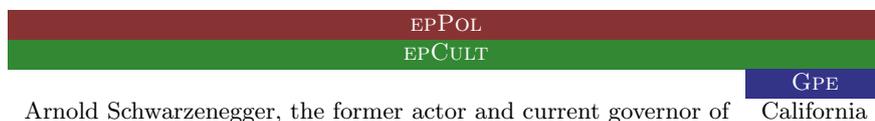
or:

  
My dear colleagues of the CDU and the SPD

## Instances with multiple labels:

In principle, it is possible for an instance to belong to more than one class. In this case, the annotators can also assign more than one label. However, this should be used with caution.

Examples of instances with multiple labels:

  
Arnold Schwarzenegger, the former actor and current governor of California

### 3.2 Special rules and idiosyncratic cases

The table below lists frequent and ambiguous cases for which we have decided on how to treat them:

Instance	Label	Comment
Dear Sir or Madam	–	formulaic form of address, do not annotate
Dear colleagues	–	formulaic form of address, do not annotate
my dear colleagues from the CDU	EPOL	specific persons in the Bundestag
the next speaker	EPOL	refers to politicians in the Bundestag
we Free Democrats	EOPOL	ambiguous between person and body (party)
we Greens	EOPOL	ambiguous between person and body (party)
this House	EOPOL	reference to the Bundestag
the legislature	EOPOL	reference to the legislature
Politics	EOPOL	reference to the government/parliament

Table 2: Some special cases and their treatment.

### References

[Wirth et al., 2019] Wirth, W., Wettstein, M., Wirz, D., Ernst, N., Büchel, F., Schulz, A., Esser, F., and et al (2019). *Codebook: NCCR democracy Module II: The Appeal of populist Ideas and Messages*. unpublished paper.

## **Index**

ELITE, 2, 3  
ELITE-ORGAN, 2, 3  
EOECON, 4  
EOFINANCE, 4  
EOMEDIA, 4  
EOMIL, 4  
EOMOV, 6  
EONGO, 5  
EOPOL, 4  
EOREL, 5  
EOSCI, 5  
ELITE-PERSON, 2, 3  
EPOWN, 3  
GPE, 3  
THE PEOPLE, 2, 3, 6  
PAGE, 6  
PETH, 6  
PFUNC, 6  
PGEN, 7  
PNAT, 6  
PSOC, 6

Coordination, 8

Domains, 2, 3

Embedded phrases, 8

Instances with multiple labels, 8

Prepositional phrases, 7