

University of Stuttgart

Institute for Natural Language Processing

Scaling up Discourse Quality Annotation for Political Science

Neele Falk and Gabriella Lapesa



Discourse from a Social Science perspective

constitute knowledge, exchange of different opinions **socially relevant** topics...

communication in

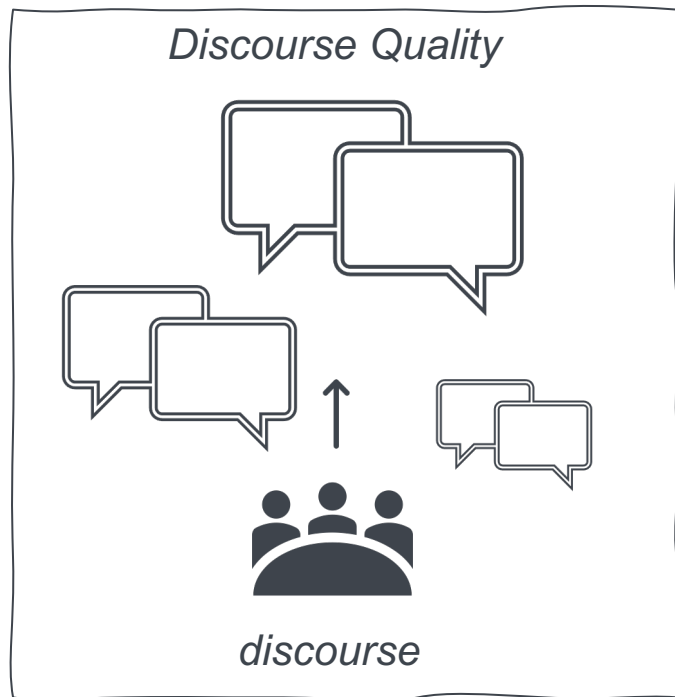
- parliaments
- social media
- group discussions



“In fact, that's a big problem. The countries of the EU need to cut military spending and put the money aside for the welfare state, for pensions, and for European civilization, European civilization, which is very human and humane. We don't need military, we don't need weapons, that's one way to address the problem.”

...linked to **politics and policy making**

Modelling Discourse Quality



- evaluate different models to predict discursive quality
- investigate the relationship between two different theoretical frameworks:
 - Social Science: **Discourse Quality**
 - Argument Mining: **Argument Quality**

Discourse as a Key for Deliberative Democracy

Deliberative Theory places **discourse** at the heart of decision-making



- **citizens can contribute** (more than vote)
- counteracts drawbacks of current democracy
- promotes **mutual understanding** and recognition, **inclusion** of all perspectives

Operationalization: Discourse Quality Index (DQI) (Steenbergen, 2003)

Discourse quality is broken down in smaller sub-dimensions

properties of good
argumentation (rationality)



productive contribution to a discourse
aimed at decision-making
(consideration of what is best for the
collective)

properties of **appreciative**
communication (respect towards
groups and other discourse participants)

Support Empirical Research on Deliberation using NLP

empirical deliberative science: annotate data based on DQI

- ... investigate under which conditions deliberation can be successful



task: predict deliberative quality given textual input

- ... process large unstructured data

Deliberation in Europe: Europolis Dataset from Social Science



- live deliberation
- transnational group-poll (Brussels, 2009)
- topics: immigration in the EU
- 910 spoken contributions

justification:

in-depth reflection, provided evidence

respect:

shows empathy/respect towards other groups

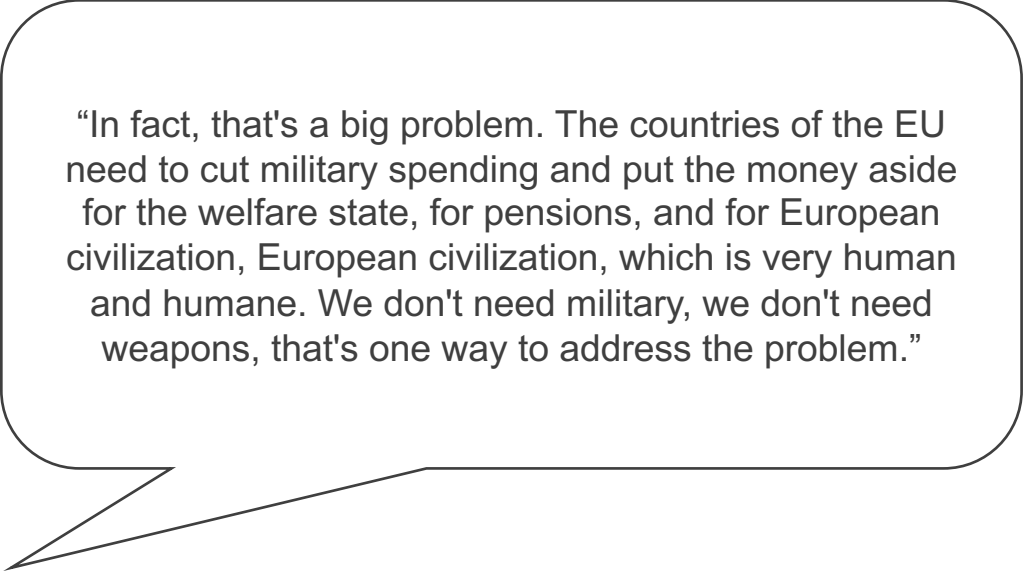
interactivity:

makes reference to other participants' contributions

common good:

argument is framed with regards to the 'common good' (interests of a broader community, solidarity or equality)

An Annotated Example



“In fact, that's a big problem. The countries of the EU need to cut military spending and put the money aside for the welfare state, for pensions, and for European civilization, European civilization, which is very human and humane. We don't need military, we don't need weapons, that's one way to address the problem.”

qualified **justification**

implicit **respect**

reference to **common good**

no **interactivity**, i.e., no reference to other participants

Problem: Data is scarce and class distribution imbalanced

justification		common Good	
no justification	138	no reference	128
inferior	372	own country	675
qualified	303	ref. to common Good	107
sophisticated	92		

interactivity		respect	
negative reference	40	disrespectful	79
no reference	380	implicit	657
neutral reference	324	explicit	126
positive reference	166		

Problem: Data is scarce and class distribution imbalanced

- Solution 1:
 - Data Augmentation
- Solution 2:
 - Integration of Argument Quality

justification		common Good	
no justification	138	no reference	128
inferior	372	own country	675
qualified	303	ref. to common Good	107
sophisticated	92		

interactivity		respect	
negative reference	40	disrespectful	79
no reference	380	implicit	657
neutral reference	324	explicit	126
positive reference	166		

Solution 1: Synthetic Data Augmentation

Easy Data Augmentation (EDA)

- synonym replacement
- random insertion (a random synonym of any word of the sentence)
- random swap
- random deletion

	orig. size	augm. size	classes	
justification	546	2,523	4	→ more training data
common good	546	2,031	3	
interactivity	546	1,020	4	
respect	522	1,663	3	→ better class distribution

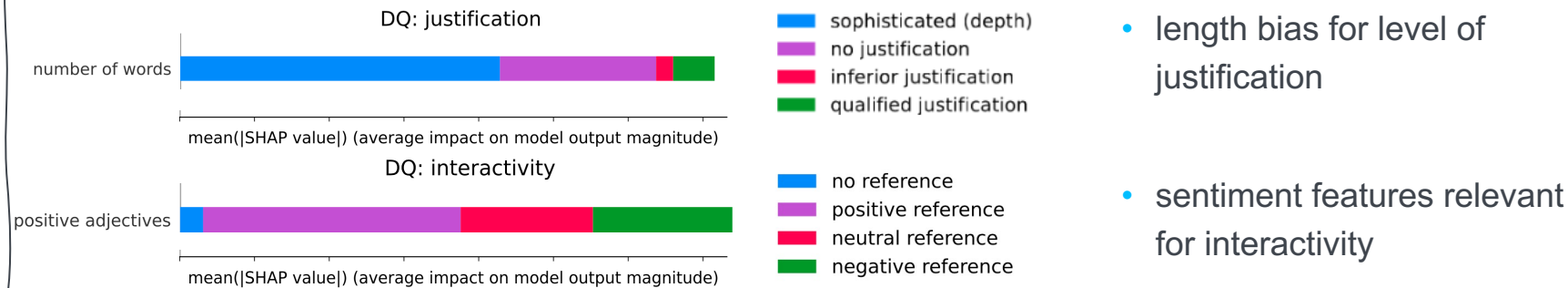
Data Augmentation Improves Results for Imbalanced Classes

	justification	respect	common Good	interactivity
majority baseline	0.15	0.28	0.15	0.29
feature-based*	0.36	0.41	0.27	0.40
roberta-base	0.74	0.86	0.65	0.75
roberta-augment	0.71	0.81	0.80	0.84

* gradient boosted trees with textual complexity, surface and sentiment features

Model Introspection

SHAP values to retrieve the most relevant linguistic features



Attribution scores to retrieve the most relevant words

- causal connectives and expressions of opinion are picked up by the model, but also spurious correlations with lexical cues

#s Strange ly enough , I am in favor of strict border control , because illegal migration compromises the position of migrant who are already in Europe . So you have to see that this illegal migration , the reason for illegal migration , has to be eliminated . These are two reasons why people come . Because there is an opportunity for work and because there are too few people in Europe to do certain jobs . And so if you are in the EU and you want to stay there , you have to see that employers don 't have that kind of need so that the market for illegal immigrants is finally dried up . #/s

(a) JUSTIFICATION: example for *sophisticated justification*

Solution 2: Argument Quality and Discourse Quality

Argument Quality: coarse-grained taxonomy of Wachsmuth (2017)

- **Cogency:** sufficient justifications, topic relevance
- **Effectiveness:** persuasive, emotional appeal
- **Reasonableness:** contribute to resolution of issue, acceptable

Annotation Study: *EuropolisAQ*

- Annotation guidelines by Ng et al. (2020)
- 513 contributions

Relationship between Argument Quality and Discourse Quality

	justification	common Good	interactivity	respect
cogency	0.33*	0.12	-0.01	0.34*
effectiveness	0.26*	0.10	0.10	0.25*
reasonableness	0.27*	0.12	0.12	0.23*
overall	0.30*	0.11	0.06	0.30*

Highest overlap between AQ and justification and AQ and respect

Experiments: integration of AQ into classifiers for DQ

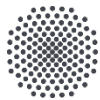
- **Result:** successful for feature-based classifier but not for transformer

Summary & Contributions

- **Experimental results:** synthetic data augmentation can be used to alleviate class imbalance
- **Methodological level:** model introspection to get a better understanding of what drives the performance of our classifiers
- **Theoretical insights:** comparison of Argument Quality and Discourse Quality

Current and future work

- How to successfully integrate AQ and DQ into one model?
- More annotation on AQ and DQ
- Use insights from AQ and DQ to support real-world deliberation (semi-automatic moderation)



University of Stuttgart

Institute for Natural Language Processing

Thank you!



Neele Falk

e-mail neele.falk@ims.uni-stuttgart.de

phone +49 (0) 711 685-84586

www.ims.unistuttgart.de/institut/team/Falk/

University of Stuttgart

Institute for Natural Language Processing