

An Annotated Corpus of Argumentative Microtexts

ANDREAS PELDSZUS

*Applied Computational Linguistics, University of
Potsdam/Germany*

peldszus@uni-potsdam.de

MANFRED STEDE

*Applied Computational Linguistics, University of
Potsdam/Germany*

stede@uni-potsdam.de

We present a freely available corpus of argumentative "microtexts", featuring short and dense authentic arguments, annotated according to a scheme for representing text-level argumentation structure. The corpus consists of 112 German texts plus professional English translations that preserve linearization and argumentative structure. We provide statistics of the variety and the linguistic realization of argumentation structure in the corpus. We hope the data release serves the needs of data-driven approaches to argument mining and qualitative analysis alike.

KEYWORDS: argument mining, argumentation structure, dialectical argument, informal logic, text corpus

{ 112 独文テキスト
+
英訳

論理構造
言語化.

1. INTRODUCTION

Argumentation can, for theoretical purposes, be studied on the basis of carefully constructed examples that illustrate specific phenomena, but for many researchers, the link to authentic, human-authored text is highly desirable. This is obviously the case for the Computational Linguistic discipline of "argumentation mining", which in recent years has attracted a lot of attention, but also for research aiming to uncover the linguistic features of argumentative text and the specific mechanisms of various argumentative moves.

For these reasons, the interest in argumentation-oriented corpora of monologue text as well as spoken dialog is rising. In the work reported here, we address this need by making a resource publicly available that is designed to fill a particular gap. So far, there exist only a few ressources with annotated argumentation structures over monologue texts, as e.g. the AIFdb, the former Araucaria corpus (Reed et al., 2008) with in large parts newswire articles, furthermore a small set of commentaries analysed in (Stede & Sauermann, 2008), and a corpus of student essays (Stab & Gurevych, 2014). While authentic text from social media or newspapers is ultimately the target for automatic argumentation mining, these sources are often not ideal for more qualitatively oriented research. In newswire text, the language can be quite complex, while in social media and language learners text, it is often ill-formed. This also has an impact on the underlying argumentation structure, in some cases it is quite trivial, and in other cases quite intransparent.

Our contribution is a collection of 112 "microtexts" that have been written in response to trigger questions, mostly in the form of "Should one do X". The texts are short but at the same time "complete" in that they provide a standpoint and a justification, by necessity in a fairly dense form. Hence, the underlying argumentation structure is relatively clear. We collected the texts in German and then had them translated to English; both versions are available to interested researchers.

In addition to the raw texts, we provide manually created annotations of the argumentation structure, following a scheme that is inspired by the informal-logic tradition. Thus, argumentation researchers will find a resource of simple, authentic natural language texts together with suggestions of structural representations of the underlying argument. At the same time, the data can also be used for building models in automatic argumentation mining.

The paper is structured as follows. In Section 2, we describe the process of gathering the data, and Section 3 provides a brief summary of the annotation scheme. The process of creating the annotations is described in Section 4. Some statistics on the corpus and the argument structures are presented in Section 5, and Section 6 gives information on the form and availability of the corpus. Finally, some conclusions are presented in Section 7.

Sec 2 : 資料収集

Sec 5 : 統計, 形式構造

Sec 3 : P/T-ト解説

Sec 6 : コーパス情報

Sec 4 : P/T-ト過程

Sec 7 : 結論

attentionは
↓
論理構造
特徴欠如?

新解釈は
debate
は有利?

2. DATA COLLECTION AND CLEANING

2.1 Collection

The microtext corpus consists of two parts. On the one hand, 23 texts were written by the authors as a “proof of concept” for the idea. These texts also have been used as examples in teaching and testing argumentation analysis with students. An example text is given in (1):

(1) Energy saving light bulbs contain a significant amount of toxins. A commercially available bulb may contain for example up to five milligrams of mercury. That's why they should be taken off the market, unless they're unbreakable. But precisely this is unfortunately not the case.

On the other hand, 90 texts have been collected in a controlled text generation experiment, where normal competent language users wrote short texts of controlled linguistic and rhetoric complexity.

To this end, 23 probands were instructed to write a text on a topic that was to be chosen from a given set of trigger questions. All probands were native speakers of German, of varying age, education and profession. They received a short written instruction (about one page long) with a description of the task and three sample texts. The probands were asked to first gather a list with the pros and cons of the trigger question, then take stance for one side and argue for it on the basis of their reflection in a short argumentative text. Each text was to fulfill three requirements: It should be about five segments long; all segments should be argumentatively relevant, either formulating the main claim of the text, supporting the main claim or another segment, or attacking the main claim or another segment. Also, the probands were asked that at least one possible objection to the claim should be considered in the text. Finally, the text should be written in such a way that it would be understandable without having its trigger question as a headline. Regarding these triggers, we offered a number of questions to the probands to choose from, and the five most frequently selected issues were:

23 texts
90 texts

• 5 segments
• 関連させた
• 4行-1, 7行-1

- motions {
- Should the fine for leaving dog excrements on sidewalks be increased?
 - Should shopping malls generally be allowed to open on Sundays?
 - Should Germany introduce the death penalty?
 - Should public health insurance cover treatments in complementary and alternative medicine?
 - Should only those viewers pay a TV licence fee who actually want to watch programs offered by public broadcasters?

2.2 Cleaning

Since we aim for a corpus of texts featuring authentic argumentation but also regular language, all texts have been corrected for spelling and grammar errors. As a next step, the texts were segmented into elementary units of argumentation. Most probands already marked up in some way what they regarded as a segment. Their segmentation was corrected when necessary, e.g. when only complex noun phrase conjuncts or restrictive relative clauses had been marked, or when subordinate clauses had not been split off. All remaining texts were segmented from scratch. Due to this step of (re-) segmentation, not all of the final texts conform to the length restriction of five segments; they can be one segment longer or shorter.

Unfortunately, some probands wrote relatively long texts. We decided to shorten these texts if possible by removing segments that appeared less relevant. This removal also required some modifications in the remaining segments to maintain text coherence, which we made as minimal as possible.

Another source of problems were segments that did not meet our requirement of argumentative relevance. When writers did not concentrate on discussing the thesis, but moved on to a different issue, we removed those segments, again with minimal changes in the remaining segments. Some texts containing several of such segments remained too short after the removal and thus have been discarded from the dataset. After the cleanup steps, 90 of the original 100 written texts remained for annotation of argumentation structure.

→ $\frac{90}{100}$ 残り

長い文は
関連性の低い
ものを削る。

2.3 Translation

To supplement the original German version of the collected texts, the whole corpus has been professionally translated to English, in order to reach a wider audience of potential users. Our aim was to have a parallel corpus, where annotated argumentation structures could represent both the German and the English version of a text. We thus constrained the translation to preserve the segmentation of the text on the one hand (effectively ruling out phrasal translations of clause-type segments) and to preserve its linearization on the other hand (disallowing changes to the order of appearance of arguments). Besides these constraints, the translation was free in any other respect. Note that the translator had only access to the segmented source text, but not to an argumentative analysis of the text.

3. ANNOTATION SCHEME

For all 112 (23+90) texts, the argumentation structure has been annotated manually. Our representation of it is based on Freeman's theory of the macro-structure of argumentation (Freeman 1991, 2011), which aims to integrate the ideas of Toulmin (1958) into the argument diagramming techniques of the informal logic tradition (Beardsley 1950; Thomas 1974) in a systematic and compositional way. Its central idea is to model argumentation as a hypothetical dialectical exchange between the proponent, who presents and defends his claims, and the opponent, who critically questions them in a regimented fashion. Every move in such an exchange corresponds to a structural element in the argument graph. In Figure 1, we show the representation for one of our microtexts. The nodes of this graph represent the propositions expressed in text segments (grey boxes), and their shape indicates the role in the dialectical exchange: Round nodes are proponent's nodes, square ones are opponent's nodes. The arcs connecting the nodes represent different supporting (arrow-head links) and attacking moves (circle/square-head links). By means of recursive application of relations, representations of relatively complex texts can be created.

図形意朱

○ : 提案者

→ : サポート

● : アタック

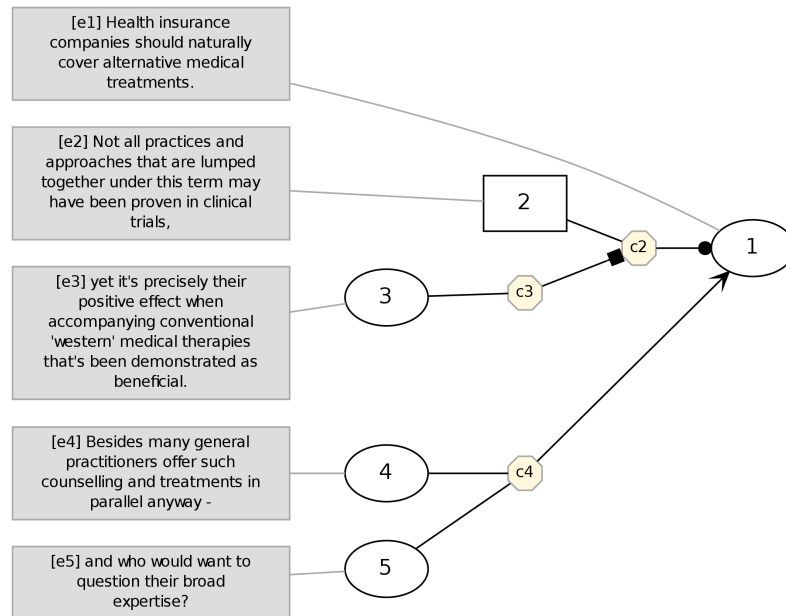


Figure 1: sample text and argumentation graph

The scheme distinguishes several different supporting and attacking moves, or argumentative functions of a segment. Besides the "standard" case of a premise supporting a claim, there can be support by example. For the attack moves, the scheme distinguishes rebuttals (challenging the acceptability of a proposition) from undercutters (challenging the acceptability of an inference between two propositions). In our example text shown in Figure 1, the second segment rebuts the first segment, and this rebutting move is then undercut by the third segment. Furthermore, the scheme allows for combining multiple premises in one move. In the example, segment four and five jointly support the main claim, which corresponds to linked premises in Freeman's theory. Note, that this combination of premises is not only possible for supporting moves but also for attacking moves.

Our move inventory could be specified further with a more fine-grained set, as provided for example by the theory of argumentation schemes (Walton et al, 2008). Still, we focus on the coarse grained set, since we see it as

複数のPremiseが
合わさってサポートor
アタックする場合

★ attack 2種類

• rebuttals

主張に対する反論

• undercutters

2つの主張の関係性

に対する反論

providing a reasonable “backbone” of the argumentation, and since it reduces the complexity for the task of automatic argument identification and classification, which is one central target application of the corpus.

Our adaption of Freeman's theory and the resulting annotation scheme is described in more detail and with examples in (Peldszus & Stede, 2013), where comparisons to related approaches are provided as well.

4. ANNOTATION PROCESS

In order to show that the annotation scheme can be applied in a reproducible fashion, we conducted annotation studies. We found that trained annotators can determine the argumentation structures reliably: On the basis of written 8-page long annotation guidelines, three annotators achieved an agreement of Fleiss $k=0.83$ for the full task (i.e. the segment-wise annotation of full argument graph features) and even higher agreement for the basic distinctions between proponent and opponent, or supporting and attacking moves. A more detailed explanation of this agreement study and its results is given in (Peldszus, 2014).

After verifying our approach by means of the agreement study, the markup of argumentation structures in the full corpus was done by one expert annotator. All annotations have been checked, controversial instances have been discussed in a reconciliation phase by two or more expert annotators. The annotation of the corpus was originally done manually on paper. In follow-up annotations, we used GraPAT (Sonntag & Stede, 2014), a web-based annotation tool specifically dedicated to constructing graph structures.

All annotation studies and the annotation of all texts have been done on the original German version of the corpus. Since the professional translation preserves linearization and argumentation structures, all annotated graphs represent both the German original and the English translation of the argument.

The bilingual texts and the annotations are publicly available in a suitable XML format; see Section 6.

GraPAT
L 7/8 - 10
7-11

5. CORPUS STATISTICS

The corpus features a wide range of different argumentation patterns. In the following, we will present detailed statistics on these, including distribution of roles and argumentative moves, positioning of the central claim in the text, as well as forward (from premise to conclusion) and backward linearizations of arguments.

5.1 General statistics

In the corpus, there are 112 texts, with in total 576 segments. Table 1 shows the length of texts in the corpus measured in segments: The great majority of texts are four, five or six segments long (the average being 5.1), with only a few exceptions.

text length	number of texts
3	3
4	11
5	71
6	26
7	2
8	0
9	0
10	1

5 segments
＜5の長さ

Table 1: Length of the texts in segments

5.2 Central claim

In the English-speaking school of essay writing and debating, there is a tendency to state the central claim of a text or a paragraph in the very first sentence, followed by supporting arguments. To some extent, we can expect to find this pattern also in other languages. To investigate whether the tendency also holds in our corpus, we divide each text into five equal parts and count the occurrence of the central claim in this position. As Table 2 shows, the dominant position is indeed the beginning of a text, directly followed by the end of text. Note however, that the overall majority of central claims (57%) is at positions other than the beginning.

position	number of central claims
1/5	48
2/5	18
3/5	16
4/5	3
5/5	27

基本頭12<3.

Table 2: Position of the central claim

5.3 Argumentative role

As we indicated earlier, the scheme distinguishes two argumentative roles: the proponent and the opponent. Of the 576 segments, 451 are proponent ones and 125 are opponent ones. While there are 15 texts where no opponent segment has been marked (either because the author did not conform to the requirement to consider at least one objection or because he phrased it indirectly in a non-clausal construction), the majority of texts (74) have exactly one opponent segment. Two opponent segments can be found in 18 texts, and three of them in five of the texts. Furthermore, Table 3 shows the position of opponent segments:

position	number of objections
1/5	20
2/5	29
3/5	22
4/5	36
5/5	18

→ debate の
場合は違う.

Table 3: Position of opponent segments (objections)

It turns out that the dominant place to mention a potential objection is right before the end of the text, thus giving the author the possibility to conclude his text with a counter of the potential objection.

For a comparison of the distribution of argumentative roles between this corpus and a corpus of longer newspaper commentaries, see (Peldszus & Stede, 2015a).

5.4 Argumentative function

The frequency of argumentative functions annotated in our corpus is shown in Table 4: Most segments are normal

support moves. Examples are used only rarely. About a third of the segments have an attacking function (either the opponent challenging the central claim or the proponent countering these objections), with overall more rebutters than undercutters. Linked premises are usually found in supporting arguments, and only rarely in attacks.

→ debate の
場合どうか
わからない。

type	number	sub-type	number
support	272	normal	263
		example	9
attack	171	rebut	108
		undercut	63
linked	21		
central claim	112		

Table 4: Frequency of argumentative function

It is noteworthy that rebutters and undercutters are not equally distributed over both argumentative roles. This is shown in Figure 2: The opponent typically rebuts, and the great majority of these rebuttals is directed against the central claim, while only a few work against supporting arguments. In contrast to that, the proponent usually undercuts. We attribute this to the common strategy of the authors to first concede a possible objection, thereby demonstrating that their presentation is not fully biased, and then render it irrelevant.

Also notice that a possible objection (an attack of the opponent) does not necessarily need to be counter-attacked by the proponent: The total number of attacks by the proponent is significantly smaller than the total number of attacks by the opponent (63 vs. 108). This is not too surprising – an author might rather choose to present just another good reason in favour of the central claim, and thereby outweigh the objection, or he might pose the possible objection in an unalluring manner signalling that counter-attacking or outweighing is not even necessary.

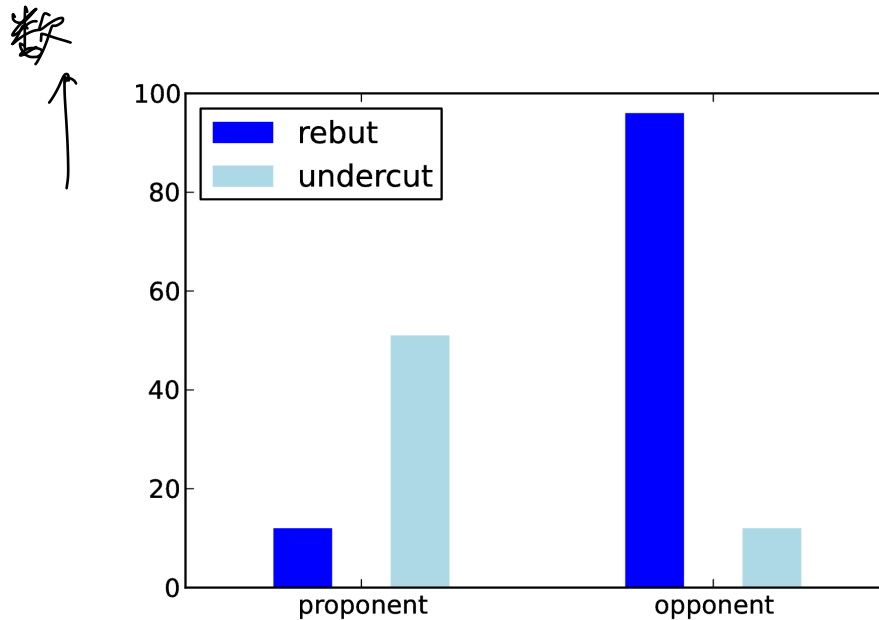


Figure 2: Attack moves against argumentative role

5.5 Attachment distance

One aspect of argumentation structure that makes its automatic recognition especially challenging is the possibility of long distance dependencies. Although segments are often connected locally, i.e. they are supporting or attacking adjacent segments, there may very well be direct argumentative relations across the whole texts, even between the very first and the very last segment of a text. It is thus worthwhile to investigate the degree to which we find these non-local relations in our corpus.

To this end, we calculate the distance and the direction of attachment for every relation annotated in the corpus (464 in total, the remaining segments functioning as central claims). An attachment distance of -1 means that the target of the argumentative relation directly preceeds the source, a distance of +1 corresponds to a target immediately following the source. For segments targeting a relation instead of another segment, as it is the case for undercutters and linked premises, we considered the position of the source of the targeted relation. For example, the undercutting segment 3 in the graph in Figure 1 has an

根 : +1

祖先 : -1

attachment distance of -1, as it undercuts the relation of the previous segment 2.

The distribution of distances and directions of attachment found in the corpus is shown in Figure 3. The great majority (45%) of argumentative relations attach to the immediately preceding segment. Another 11% attach to the following segment. In total, 56% of the relations hold between adjacent segments, so conversely nearly half of the segments do not attach locally. Considering that our texts are relatively short, it is to be expected to find even more non-adjacent relations in longer texts. E.g., Stab & Gurevych (2014) report a rate of 63% of non-adjacent relations in their corpus of student essays.

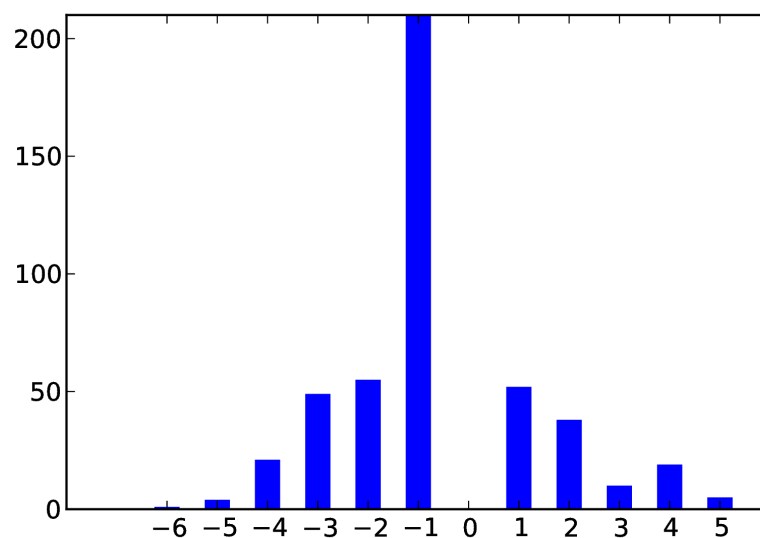


Figure 3: Attachment distance and direction (negative distances directed backwards and positive distances directed forwards)

5.6 Linearization strategies

The final feature of the argumentation graphs we want to investigate is how authors linearize their arguments in the text. This has already been covered to some degree in the sections 5.3 and 5.4 when we studied at which positions in the text the central claim and objections are typically expressed. In the following, we combine this with the

direction of attachment and distinguish four different simple linearization strategies, which are summarized in Table 5.

The first strategy involves only backward relations, where the author open his text with the central claim (c) and then presents a series of reasons, possible objections, and counters, all of them directed backwards (b), targeting propositions made in prior segments. The second linearization strategy unfolds the argumentation the other way around, with only forward relations. The author first starts with premises and successively draws conclusions from them (f) until he finally reaches the central claim of the text. The third strategy combines these two patterns, presenting the central claim in the middle of the text. It naturally involves a switch of attachment direction after the central claim. All other texts not matching one of these three strategies involve a change in the direction of argumentation independent of the presentation of the central claim.

linearity strategy	pattern	frequency
backward	c b ⁺	50%
forward	f ⁺ c	5%
forward-backward	f ⁺ c b ⁺	13%
other	other	31%

Table 5: Ratio of texts matching different linearization strategies

As shown in Table 5, the first strategy which opens with the central claim and argues for it with only backward relations, is the dominant one found in half of the texts. The reverse strategy is used only rarely, while the mixed strategy appears at least in 13% of the texts. Most interestingly, about 31% of the texts do not follow these strict patterns. As an example, see Figure 4: This text's linearization pattern corresponds to 'fbfbc', featuring multiple changes in direction before the central claim is stated.

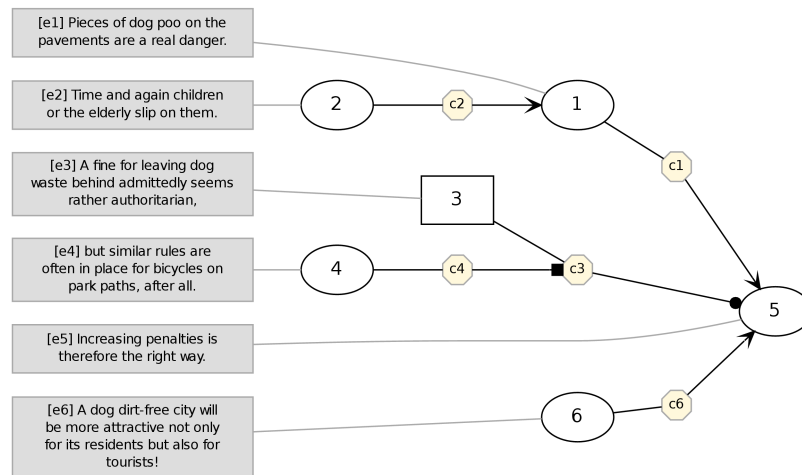


Figure 4: An example text with multiple direction changes.

6. CORPUS DELIVERY

The corpus is published online¹ and freely distributed under a Creative Commons BY-NC-SA 4.0 International License². The annotated graph structures are stored in the Potsdam Argumentation XML format (PAX), a both human- and machine-readable format, similar to GraphML (Brandes et al, 2002). The corpus repository contains a specification of the format in form of a document type definition.

For both versions of the corpus, the German and the English one, we provide the raw source text, the annotated argumentation graph in PAX (primarily for machine reading), as well as a graphical argument diagram such as the one in Fig. 1, in order to facilitate human inspection of the structures. An importer for the PAX format has also recently been added to the Carneades Tools³, allowing to map and evaluate the graphs of our corpus.

Finally, notice that first results on automatic recognition of the argumentation structures annotated in the corpus are presented in (Peldszus & Stede, 2015b).

¹ <https://github.com/peldszus/arg-microtexts>

² <https://creativecommons.org/licenses/by-nc-sa/4.0/>

³ <https://carneades.github.io/>

7. CONCLUSION

We presented a freely available parallel corpus of short argumentative texts. Our microtexts are “authentic” in the sense that the vast majority was written by probands not involved in the research, and the trigger questions concern issues of daily life and public interest. At the same time, they are “constrained” because we provided the probands with some instructions on target length and form. This was done in order to obtain a relatively homogeneous data set that allows for studying properties of the argumentation. For the same reason we decided to do a moderate “cleaning” of the texts, which on the one hand reduces “authenticity” but on the other hand contributes to uniformity and – for many purposes – usability.

Research in automatic argument mining typically targets social media contributions in their original form and often focusses on the task of argument identification and local relation identification. While the design of our corpus differs from this orientation, we still think that the data can be useful for purposes of feature engineering and as supplemental training data. Finally, we consider our data set to be a reasonable starting point for the task of global relation identification, i.e. for the automatic prediction of text-level argumentation structure, before tackling more complex text genres.

ACKNOWLEDGEMENTS: We would like to thank the anonymous reviewers and the conference attendees for their helpful feedback. Thanks to our probands and annotators for participating in our studies, as well as to Regina Peldszus and Kirsten Brock for the translations. The first author was supported by a grant from Cusanuswerk.

REFERENCES

- Beardsley, M. C. (1950). *Practical Logic*. Prentice-Hall, New York.
- Brandes, U., Eiglsperger, M., Herman, I., Himsolt, M., & Marshall, M. S. (2002). GraphML progress report structural layer proposal. In *Graph Drawing* (pp. 501-512). Springer Berlin Heidelberg.

- Freeman, J. B. (1991). *Dialectics and the Macrostructure of Argument*. Foris, Berlin.
- Freeman, J. B. (2011). *Argument Structure: Representation and Theory*. Argumentation Library (18). Springer.
- Peldszus, A. (2014). Towards segment-based recognition of argumentation structure in short texts. In: *Proc. of the First Workshop on Argumentation Mining* (pp. 88-97). ACL, Baltimore, Maryland.
- Peldszus, A. & Stede, M. (2013) From argument diagrams to argumentation mining in texts: a survey. In: *Int'l Journal of Cognitive Informatics and Natural Intelligence (IJCINI)* 7(1):1-31, 2013
- Peldszus, A. & Stede, M. (2015a) Towards detecting counter-considerations in text. In: *Proc. of the 2nd Workshop on Argumentation Mining* (pp. 104--109), NAACL-HLT, Denver/CO.
- Peldszus, A. & Stede, M. (2015b) Joint prediction in MST-style discourse parsing for argumentation mining. In: *Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 938-948). Lisbon, Portugal.
- Reed, C., Mochales Palau, R., Rowe, G., & Moens, M. F. (2008). Language resources for studying argument. In *Proc. of the 6th conference on language resources and evaluation (LREC 2008)* (pp. 91-100).
- Sonntag, J. & Stede, M. (2014). GraPAT: A tool for graph annotations. In: *Proceedings of the Language Resources and Evaluation Conference (LREC)*, Reykjavik.
- Stab, C., & Gurevych, I. (2014). Annotating argument components and relations in persuasive essays. In *Proc. of the 25th International Conference on Computational Linguistics (COLING 2014)* (pp. 1501-1510).
- Stede, M., & Sauermann, A. (2008). Linearization of arguments in commentary text. In W. Ramm & C. Fabricius-Hansen (eds.), *Linearisation and Segmentation in Discourse. Multidisciplinary Approaches to Discourse 2008 (MAD 08)*, February 20-23, Lysebu, Oslo, Norway.
- Thomas, S. N. (1974). *Practical Reasoning in Natural Language*. Prentice-Hall, New York.
- Toulmin, S. (1958). *The Uses of Argument*. Cambridge University Press, Cambridge.
- Walton, D., Reed, C., and Macagno, F. (2008). *Argumentation Schemes*. Cambridge University Press.