Arguing Brexit on Twitter: a corpus linguistic study

NATALIE DYKES
Friedrich-Alexander-Universität Erlangen-Nürnberg
natalie.mary.dykes@fau.de

PHILIPP HEINRICH
Friedrich-Alexander-Universität Erlangen-Nürnberg
philipp.heinrich@fau.de

STEFAN EVERT
Friedrich-Alexander-Universität Erlangen-Nürnberg
stefan.evert@fau.de

We study argumentation on social media by applying discourse and corpus linguistic methods to a corpus of 6 million tweets containing "Brexit". We identify common argumentation strategies through manual annotation and develop corpus queries to find further instances of those schemes. Methods from discourse linguistics help to identify salient aspects shared by the various argument scheme realisations. Through modularity and iterative abstraction, the queries lend themselves to logical formalization and further automatic processing.

KEYWORDS: argumentation mining, corpus linguistics, defeasible, quantitative, social media, Twitter

1. INTRODUCTION

Argument Mining is a relatively new field in Natural Language Processing and Computational Social Science. It is concerned with the automatic extraction and representation of arguments found in large collections of electronic texts. The most foundational tasks in Argument Mining include the identification of argumentative utterances in corpora and the classification of particular structural elements like premises or conclusions. The main outcome of this process is a mapping of argumentative text sequences to their automatic analysis; which then lends itself to further processing in a computational setting.

In many cases, contributions focusing primarily on natural language processing perspectives tackle corpora of relatively well-structured and explicitly argumentative texts, such as debates (Cabrio and Villata 2013) or student essays (Peldszus and Stede 2015).

However, argumentation is often considerably less straightforward when it comes to texts from other genres. While argumentation plays a role in many settings of everyday discourse, its automatic processing faces additional challenges in some text types. Firstly, arguments tend to follow the type of defeasible logic as proposed by researchers including Kienpointner (1992) and Walton et al. (2008). In defeasible settings, the relationship between premises and conclusions is less clear than in the case of traditional argumentation schemes like Modus Ponens. For instance, consider the argument from correlation to cause (Walton et al. 2008, pp. 328–329):

> "Premise: There is a positive correlation between A and B.
> Conclusion: Therefore, A causes B."

While this is a widespread interpretation of correlation – we hope that readers will excuse our own defeasible reasoning – the phrase *Correlation does not imply causation* has gained enough traction to be the title of a Wikipedia page.[1] In this sense, even when the premise holds true, this does not necessarily mean that the conclusion is true as well.

Moreover, it can be challenging to identify premises or conclusions in the first place, because they are often left implicit in everyday argumentation. This holds especially true for social media, where texts are often extremely short. Therefore, Bosc et al. (2016) consider every tweet as argumentative if it can be interpreted to contain either a premise or a conclusion; regardless of whether the line of argumentation is present. The obscured logical structure is connected to a third aspect that makes argumentation mining difficult on less structured texts: persuasion is usually not achieved by strict logical relations, but rather by persuasion through rhetoric strategies such as selection, arrangement, and phrasing of argumentative units (Wachsmuth et al. 2018).

In this contribution, we focus on argumentative discourse on Twitter, where all of these aspects play a particularly prominent role: As Nigmatullina and Bodrunova (2018) suggest for the case of discussion threads, tweets are strongly fuelled by emotions, particularly anger and indignation, and often contain ironic elements. Moreover, a high degree of implicitness is to be expected. As our tweets were collected in 2016, all posts have a maximum length of 140 characters, which was the limit at

[1] https://en.wikipedia.org/wiki/Correlation_does_not_imply_causation

the time. Apart from this restriction, there are rather few guidelines: users can post pictures, links or GIFs and write about whatever they please, which makes it a more or less unmediated environment. Finally, tweets contain a large amount of non-standard language, as is usually the case in computer-mediated communication.

Considering these aspects, it is hardly surprising that automatically finding and classifying arguments in social media has led to relatively modest success (cf. Goudas et al. 2014, Dusmanu et al. 2017).

2. RELATED WORK

In corpus linguistics, several approaches have been taken to study argumentation. Degano (2007) studies presuppositions and dissociations in a corpus of British press texts. The analysis starts from a predefined list of explicit argumentation markers; including clause types, comparisons, descriptions or the use of particular vocabulary (cf. Levinson 1983, pp. 181–184). For instance, stating that somebody *did not manage* something indicates the presupposition that they have tried and been unsuccessful (Degano 2007, p. 366).

Another possible strategy is to combine qualitative and quantitative methods. O'Halloran (2011) analyses argumentation in transcripts of oral conversations. In a first step, he uses manual coding to identify utterances as claims and challenges. The following part consists of calculating statistical keywords, key POS tags and key semantic domains (Rayson 2008) to explore the linguistic realisation of arguments.

Al-Hejin (2015) works closer to argumentation schemes in a study of British press reports on Muslim women. He focuses on macro-propositions: "global" motives of the discourse topic (van Dijk 2008, p. 16). For example, when mentioning a woman's choice to wear a hijab, particular arguments tend to follow, e.g. that they are unwilling to integrate into Western culture (Al-Hejin 2015, p. 40). Macro-propositions were identified by keywords, which were grouped manually and verified by calculating key semantic domains.

Baker (2004) compares pro- and anti-reform speeches in a political debate on the age of consent for homosexual men in Britain. The speeches differ not only a lexical level, but also logically: opposers tend to form chains of individual arguments building upon each other, while proponents' arguments were less intertwined and "more straightforward" (Baker 2004, p. 104).

3. THE BREXIT 2016 CORPUS

The basis for our analysis is a corpus of approximately 6 million tweets

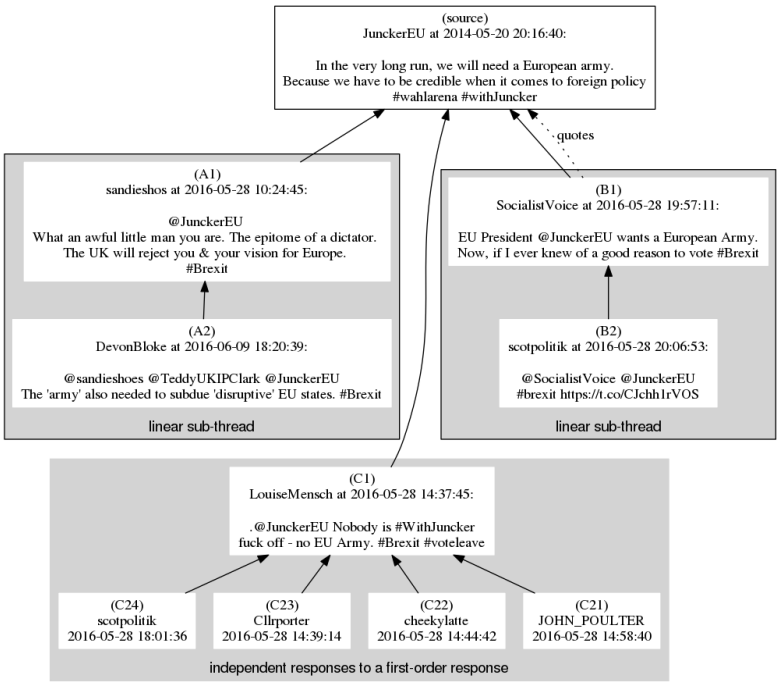

Figure 1: reply thread example

containing the character string "Brexit" (16/05/05–16/08/24)[2] and their associated reply threads. Reply threads were constructed by retrieving all available tweets for which there is a reply in our initial database. One example for such a reply thread can be seen in figure 1. Note that the initial source tweet by `@JunckerEU` is from 2014, hence before the actual start of the database. For mining arguments, we limit ourselves to tweets where the post containing the relevant string was written before the Brexit referendum, since we assume that this sample will have higher consistency in argumentation. This leads us to a collection of approximately 2.5 million tweets.

As shown in examples (1) to (3), tweets in the corpus may contain all of the features mentioned as challenging for processing; such as sarcasm, non-standard language and misspellings, emojis and implicitness. Moreover, many tweets are likely not directly argumentative; such as example (3).

> (1) Brexit will result in Zombie Hitler returning leading an army of demonic nazi crabs, marching sideways across Europe to devour our brains!

---

2 https://web.archive.org/web/20171121195029/http://www.eecs.qmul.ac.uk/~dm303/brexit/

(2) German : Acknowlege significant UK contribution 2 the EU

(3) if we Brexit is this our last Eurovision? Or do we get a guest spot? Like Australia 😂

3.1 Preprocessing

Before the corpus was ready for analysis, several pre-processing steps were taken in order to allow for elaborated searches and more valid results. The first necessary aspect concerned deduplication: Following Schäfer et al. (2017), we removed posts that were likely generated by social bots; thus reducing our dataset to approximately 1.8 million tweets and 32 million tokens. The data was then automatically processed using several pipelines for tokenization, part-of-speech tagging (Owoputi et al. 2013), lemmatization (a simple rule-based lemmatizer based on tokens and POS tags combined with a word lookup), phrase chunking and named entity recognition (Ritter et al. 2011).

3.2 Indexing

In order to make the corpus usable for extracting argumentative sequences, the data was indexed with the IMS Open Corpus Workbench (CWB). This allowed us to combine all levels of annotation and thus formulate complex queries (cf. Evert and Hardie 2011). Moreover, the CWB includes several features that have proven useful to our approach; including macro definitions and wordlists, which will be described below.

4. MODELING ARGUMENTS: PATTERNS AND SCHEMES

4.1 Corpus Queries

Unlike most traditionally corpus linguistic approaches, the central structuring elements of our study are not wordlist of keywords or collocations, but CWB queries, which are intended to provide the input for an inventory of logical formulas. These formulas, which are being developed by another group of our ongoing project, cover contents that is typically to be expected in everyday argumentation, for instance:

(A): \forall {?0 : entity} \in {?1 : entity} : {?2 : property}(?0)

In this formula, there are three empty slots, corresponding to an entity (0), a group (1) and a property (2) which, applies to all group

members. It summarises aspects found in several argumentation schemes, including Common Folks ad Populum and Position to Know (Walton et al. 2008). The schemes, in turn, manifest in various linguistic forms, including the following examples for *Common Folks ad Populum*, with the relevant parts denotated by brackets:

(4) Study shows <ordinary folk are losers under the EU while the wealthy prosper>

(5) @DrAlanGreene <I'm as against #Brexit as the next man> but this is nonsense

In both cases, an appeal to "regular" people is the core of the argument; framing the speaker's standpoint as the commonsense thing to think. The passages in bold are the part of the respective corpus query that was matched in a given tweet. The queries are designed to capture linguistic patterns, which we expect to reflect particular argumentative strategies. Consider the query that found hits like 5):

```
<np>@0:[pos_simple="L|N|P|Z|#"] []*</np> @1:[::]
/region[vp] (/region[np])? "as" (/ap[])* "as" "all|the|any"
"next|old|other" @2:[::](/ap[])* [lemma=$nouns_person_common |
pos_simple = "Z"]+@3:[::];
```

The central semantic element defining the argumentative potential of this query is the notion of normality. This is represented by the specified adjectives variable `$nouns_person_common`, which in this case is specified to be present within the second noun phrase of the hit. The variable reads from a wordlist consisting of nouns that might be considered "generic" references to people, i.e. fellow, human, person, ppl etc. The rest of the query mostly consists of phrase chunk and part of speech restrictions.

Finally, parts marked with an `@` are anchor points, marking particular points where the query matches can be extracted to use as input to the logical formulas. For instance, the sequence between `@0` and `@1` – *I'm* in the case of example 5) – can be specified to fill the entity definition slot `{?0: entity}` in the given formula. Thus, whatever takes up the NP slot in the query is assumed to be equivalent to the entity being assigned particular characteristics based on their group membership (in this case, "normal" people, according to the tweeter).

The main goal in developing these queries is to maximise precision, while recall is of secondary concern. In this way, we aim to generate input to the formulas, which are suited to incorporate arbitrary linguistic structure and reflect its logical relation, provided that the query is precise enough. For instance, if the restriction to the wordlist

`$nouns_person_common` is omitted, the following result is found by the query:

(6) He'll be back as soon as the next one has screwed up Brexit

Superficially, the structure looks like the one intended to be captured by the query. However, to the human analyst it is obvious that the construction *as _ADJ* as does not contribute to a grouping function in this case – the correspondence between the linguistic surface forms is purely incidental. As the grammatical and lexical categories are rather coarse in many positions of the query, it is important to emphasise that the required precision can only be achieved if query development is performed iteratively. Only by regular examination of the concordance lines for each match can the queries gradually become adequately exact.

5. CASE STUDY: SELF-IDENTIFICATION IN THE BREXIT CORPUS

In order to demonstrate the usefulness of our approach to the study of discourse and argumentation, we present a case study of queries designed to extract speakers' self-assignment to a particular group in potentially argumentative contexts. This is a subset of queries matched by formula A, where entity 0 is specified to be the speaker/tweeter themselves. Thus, group membership is the grounds on which a particular claim is made. We focus on self-identification, assuming that these particular statements have special persuasive potential, as the speaker claims expertise or first-hand knowledge of the situation in question.

In order to explore these tweets, the results of three different queries were annotated (*as NP I VP*, *NP like me VP*, {*ordinary/common/normal*} {*people*} *VP*). While only the first and second query explicitly reference the speaker, we suggest that the third query may also be regarded as a kind of self-identification. Tweets following this pattern will likely often imply that the speaker themselves identifies as belonging to the group "normal" people. Consider example (4) from above – it is unlikely that a speaker would highlight the disadvantages of Brexit for "ordinary people" (whom they understand in contrast to "the wealthy" in an act of othering) and not define themselves as belonging to that group.

The hits of each of the three queries were categorised manually and annotated for stance (leave/ stay/ unclear). Figure 2 shows the empirical distribution of this stance variable for the three argumentational queries.

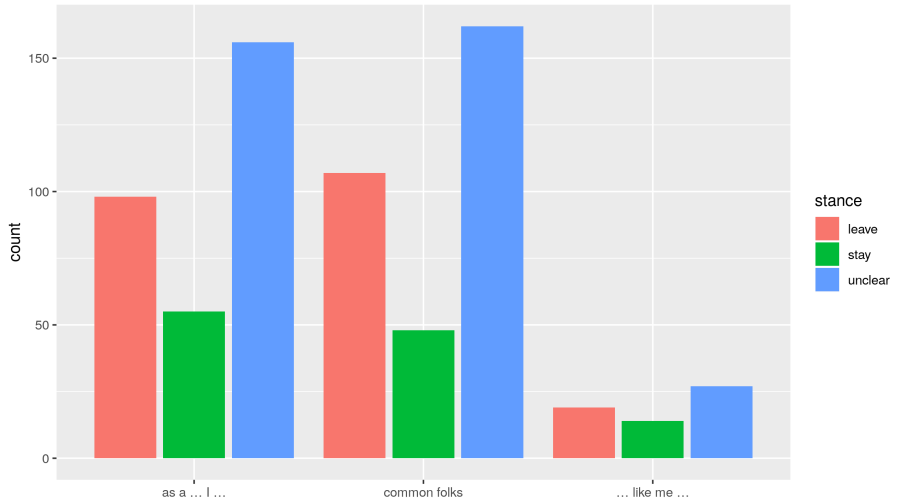

Figure 2: empirical distribution of stance towards Brexit in three annotated queries

Note that this categorisation only applies to the search result itself and not to the overall statement presented in the tweet, which may well be different. While we do aim for a complete analysis of tweets' argumentative content in the future, we take a nested approach where each proposition in the form of a query hit is first evaluated out of context and later may be determined to be a filler in and/ or to contain another proposition. For instance, a statement like X proved that Y is wrong, Y and X are to be captured by separate queries and whatever is claimed in Y is evaluated as false only in the context of evaluating the "outer" proposition X. In this sense, the analysis presented here does not necessarily reflect a subset of Twitter users' opinions on Brexit, as some sequences categorised as "leave" may be embedded in a negative evaluation and vice versa.

A further variable for annotation was the iterative development of group identity types based on the head of the main NP of the sentence. During the examination of concordances, it became clear that most hits could easily be attributed to one of the following categories:

"person": this type includes generic person references as well as kinship terms (*ppl, folks, grandmother, parent*).

"national": this category is reserved for references to nationality, origin or residence (*Brit/Briton, resident, European, Yank, Brummie*)

"voter": rather unsurprisingly, tweeters often specified their position in terms of political or otherwise ideological, and in some cases, spiritual, alignment (*activist, leftie, fan, sceptic, winger*)

"professional": finally, users referred to their profession or educational background (*graduate, banker, fisherman, [a-z-]*worker*).

These categories were used as a basis for wordlists corresponding to particular types of person references and subsequently expanded by systematically searching the corpus for similar words through omitting particular slots within these and other queries developed in the project. Figure 3 shows the distribution of stance towards Brexit across the listed categorization of the NP head.

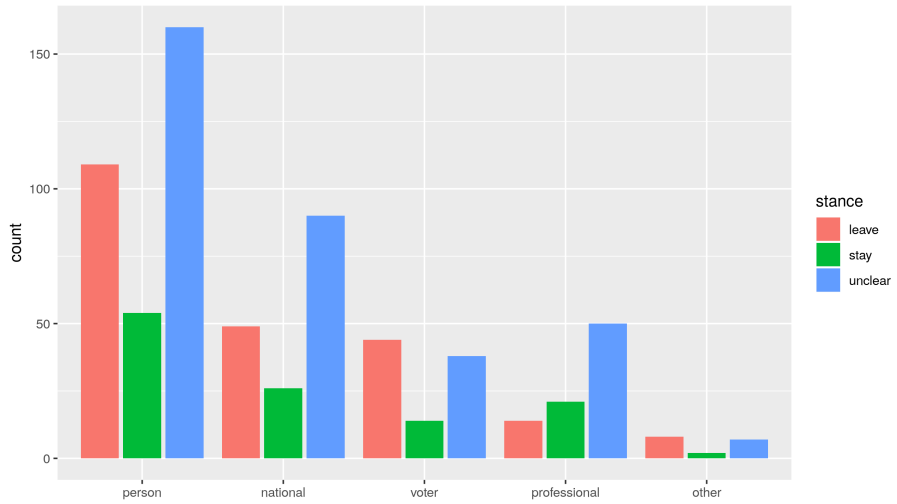

Figure 3: stance across NP categories

6. RESULTS

With regards to the stance towards Brexit, most hits show no clear pattern (cf. figure 2 above). However, "leave" outweighs "stay" in all matches. For the most part, the "unclear" annotation indicates that the statement, though related to the general topic of Brexit, does not in its own suffice to directly deduce a stance. Nevertheless, most matches are clearly used in argumentative contexts and thus promising for further analysis:

(7) the Boris, Gove, Hannan, Farage & Redwood loons in charge Thucydides on experts: "ordinary men usually manage public affairs better than their more gifted fellows"

In cases where stance is present, users ascribing themselves to the same group may use their identification as a ground for opposing arguments:

(8) As an immigrant I find myself closer to the #RemainINEU campaign argument

(9) As an immigrant I find the belief that #UK can integrate into #Europe astonishing

In a subsequent step, we use the manual annotation as a basis for further quantitative analysis to demonstrate the usefulness combining automated and manual approaches. This evaluation concerns the semantics of the VP head. Thus, we address the question of what kinds of actions, states or mindsets are justified on the basis of a particular group membership. The basis of this analysis is visualising the various verbs present in the VPs in terms of their distributional properties.

We visualize the semantics of a given slot using a two-dimensional plot: For this purpose, we created high-dimensional word embeddings (cf. Mikolov et al. 2013) using an unrelated sample of equally preprocessed English tweets collected via the Twitter Streaming API. These word embeddings are meant to capture the distributional properties of the tokens in computer-mediated communication, thereby allowing us to represent the symbols as points in space. Here, vectors that are close to one another (e.g. in terms of their cosine similarity) are considered to be semantically similar.

We then project the respective embeddings of the tokens of a slot onto a two-dimensional semantically structured space using t-distributed stochastic neighbour embedding (van der Maaten and Hinton 2008). This projection is constructed in a way that semantically similar lexical items (i.e. those with similar embeddings in the high-dimensional space) appear in the vicinity of one another in the two-dimensional space. An example can be seen in figure 4. It is obvious that the visualization is reasonable: The more generic items(such as 'people', including the CMC spelling 'ppl', 'folk', 'man', and 'person') cluster together, as do the more specific ones ('worker', 'citizen', 'voter').

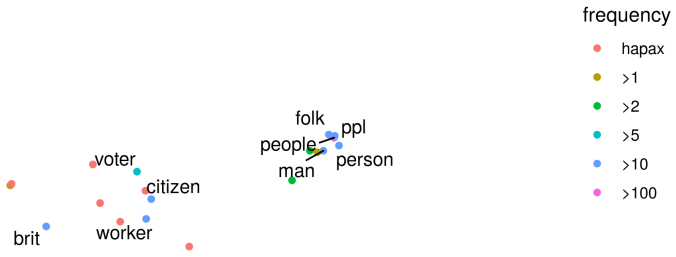

Figure 4: visualizing the NP slot for the {ordinary} {people} VP query

In order to extract more specific subsets of hits, the information from the visualisation can serve as input for more precise queries. For

instance, when the NP is specified to express affiliation and the VP slot references a verb of support, the hits are accordingly limited to more narrow types of arguments:

(10) As #Labour member I am voting Brexit

(11) As an anarchist I will vote against Brexit, as it will reinforce the notion of a nation state

Similarly, consider the examples for combining "professional" nouns in the NP and knowledge verbs characteristic for them in the VP:

(12) As a scientist, I know the immense value of #collaboration

(13) As a doctor, I know leaving the EU will be devastating for the NHS

7. CONCLUSION

As has been demonstrated, corpus queries can be a useful tool in extracting numerous argumentative sequences from a large and noisy dataset. In this sense, the queries fulfil the task of handling linguistic variation, partially realised arguments and defeasible logic. Our corpus queries are designed to balance grammatical and semantic flexibility in the argumentation patterns: semantically similar concepts can be grouped into wordlists to fill a lexical slot with multiple options as demonstrated by the `$nouns_person_common` variable in our query example. Furthermore, grammatical patterns can be queried on different levels; be it parts of speech, phrase chunks or manually written macros to accommodate for structures that the chunker does not specify. All macros and word lists are stored centrally and imported into the query files; making them a reusable resource. Each individual query is one linguistic instantiation of a particular argumentative pattern. Using the CWB's anchor point facilities, it becomes possible to mark sequences of interest and extract them for either logical formalisation or closer linguistic and argumentative analysis. In this sense, we contribute a qualitative approach to handling noisy data in online argumentation.

Current ongoing work includes mapping the developed queries to their logical representations by extracting relevant parts of the query results and filling them into the formulae. We also expect the queries to reversely inform the development of a logical framework equipped to handle noisy data, in that the results will yield insights into which argumentative phenomena are particularly prevalent in the given discourse and others like it.

The visualisation of particular word slots by distributional properties has proven insightful for the case study at hand: similar words

were grouped together, thus facilitating the task of drawing quantitatively informed qualitative conclusions. Therefore, future work includes the development of an interactive toolkit integrating visual facilities, allowing researchers to gain a comprehensive overview of the data.

ACKNOWLEDGEMENTS:

This work has been funded by the Deutsche Forschungsgemeinschaft (DFG) within the project "Reconstructing Arguments from Noisy Text", grant number 377333057, as part of the Priority Program "Robust Argumentation Machines (RATIO)" (SPP-1999).

REFERENCES

Al-Hejin, Bandar (2015): Covering Muslim women. Semantic macrostructures in BBC News. In *Discourse & Communication* 9 (1), pp. 19–46. DOI: 10.1177/1750481314555262.

Baker, Paul (2004): 'Unnatural Acts'. Discourses of homosexuality within the House of Lords debates on gay male law reform. In *Journal of Sociolinguistics* 8 (1), pp. 88–106. DOI: 10.1111/j.1467-9841.2004.00252.x.

Bosc, Tom; Cabrio, Elena; Villata, Serena (2016): Tweeties squabbling. Positive and Negative Results in Applying Argument Mining on Social Media. In : Proceedings of COMMA, pp. 21–32.

Cabrio, Elena; Villata, Serena (2013): A natural language bipolar argumentation approach to support users in online debate interactions. In *Argument & Computation* 4 (3), pp. 209–230.

Degano, Chiara (2007): Presupposition and Dissociation in Discourse. A corpus study. In *Argumentation* (21), pp. 361–378. DOI: 10.1007/s10503-007-9058-7.

Dusmanu, Mihai; Cabrio, Elena; Villata, Serena (2017): Argument mining on Twitter. Arguments, facts and sources. In : Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing.

Evert, Stefan; Hardie, Andrew (2011): Twenty-first century corpus workbench. Updating a query architecture for the new millennium. In : Proceedings of CL.

Goudas, Theodosis; Louizos, Christos; Petasis, Georgios; Karkaletsis, Vangelis (2014): Argument Extraction from News, Blogs, and Social Media. In : SETN 2014.

Kienpointner, Manfred (1992): Alltagslogik. Struktur und Funktion von Argumentationsmustern. Stuttgart: Frommann-Holzboog (Problemata, 126).

Levinson, Stephen (1983): Pragmatics. Cambridge: Cambridge University Press (Cambridge textbooks in linguistics, 8).

Mikolov, Tomas; Chen, Kai; Corrado, Greg; Dean, Jeffrey (2013): Efficient Estimation of Word Representations in Vector Space. In Yoshua Bengio, Yann
LeCun (Eds.): 1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings.
Nigmatullina, Kamilla; Bodrunova, Svetlana (2018): Patterns of emotional argumentation in Twitter discussions. In : Proceedings of the International Conference on Internet Science, pp. 72–79.
O'Halloran, Keiran (2011): Investigating argumentation in reading groups. Combining manual qualitative coding and automated corpus analysis tools. In *Applied Linguistics* 32 (1), pp. 172–196. DOI: 10.1093/applin/amq041.
Owoputi, Olutobi; O'Connor, Brendan; Dyer, Chris; Gimpel, Kevin; Schneider, Nathan; Smith, Noah (2013): Improved Part-of-Speech Tagging for Online Conversational Text with Word Clusters. In : Proceedings of NAACL.
Peldszus, Andreas; Stede, Manfred (2015): An annotated corpus of argumentative microtexts. In : Argumentation and Reasoned Action: Proceedings of the 1st European Conference on Argumentation. Lisbon.
Rayson, Paul (2008): From key words to key semantic domains. In *International Journal of Corpus Linguistics* 13 (4), pp. 519–549. DOI: 10.1075/ijcl.13.4.06ray.
Ritter, Alan; Clark, Sam; Etzioni, Oren (2011): Named entity recognition in tweets: an experimental study. In : Proceedings of EMNLP: Association for Computational Linguistics.
Schäfer, Fabian; Evert, Stefan; Heinrich, Philipp (2017): Japan's 2014 general election. Political bots, right-wing internet activism and PM Abe Shinzō's hidden nationalist agenda 5 (4), pp. 294–309.
van der Maaten, Laurens; Hinton, Geoffrey E. (2008): Visualizing data using t-SNE (Journal of machine learning research), pp. 2579–2605.
van Dijk, Teun (2008): Discourse and Context. A Sociocognitive Approach. Cambridge: Cambridge University Press.
Wachsmuth, Henning; Stede, Manfred; El Baff, Roxanne; Al-Khatib, Khalid;
Skeppstedt, Maria; Stein, Benno (2018): Argumentation synthesis following rhetorical strategies. In : Proceedings of the 27th International Conference on Computational Linguistics, pp. 3753–3765.
Walton, Douglas; Reed, Chris; Macagno, Fabrizio (2008): Argumentation Schemes. Cambridge: Cambridge University Press.