Identifying and Understanding Game-Framing in Online News: BERT and Fine-Grained Linguistic Features

Hayastan Avetisyan
German Centre for Higher Education
Research and Science Studies
Hannover, Germany
avetisyan@dzhw.eu

David Broneske
German Centre for Higher Education
Research and Science Studies
Hannover, Germany
broneske@dzhw.eu

Abstract

News providers tend to add an entertaining and eye-catching spin to online news stories by framing politics around strategies and tactics as well as wins and losses. Thus, they shape the reader's view on a particular subject, person, or event and influence their behavior regarding, for instance, voting respectively. To address this issue, we introduce the first attempt in computational linguistics to model computational frame classification. We offer a human-labeled dataset indicating the issue-game framing in news media using a comprehensive range of linguistic features. Moreover, we present an overview of potential linguistic indicators of issue and game frames. Furthermore, we attempt to provide methods for analyzing, understanding, and flagging problems that deal with subjectivity with respect to framing by identifying and presenting the respective cues. We use BERT, a pre-trained word representation model, and fine-tune it with our dataset on the binary text classification task. It turns out that BERT-like approaches can be used to detect issue and game frames. However, studies utilizing more annotated training data should be conducted to investigate its universal effectiveness. We put forward some suggestions on the grammatical features that could be taken into consideration in the further development of similar language models.

1 Introduction

Daily we are plied with political messages from various sources. Those messages are more and more often framed around strategies, competitions, wins, and losses. Studies suggest that this framing of politics increases political distrust and cynicism (Cappella and Jamieson, 1996) negatively influences citizens' knowledge,

attitude, and decisions (Aalberg et al., 2012). Thus, in one way or another, framing in political news media might shape the reader's view on a particular subject, person, or event. It might even destroy our collective trust and initiate social conflict (Pryzant et al., 2020). Therefore, the framing of political messages as a game can benefit from further research.

We find it crucial to discuss the attractiveness and popularity of the game frame in political news media coverage. Firstly, the news media tend to frame politics as a strategic game rather than to focus on political issues. Secondly, this kind of news coverage has increased over time. Thirdly, framing politics as a game increases political distrust and cynicism (Aalberg et al., 2012; Cappella and Jamieson, 1996). Moreover, it might also have a negative effect on citizens' knowledge acquisitions. One of the main reasons behind its popularity among scholars, however, might be the assumption that game framing might negatively influence democracy (Aalberg et al., 2012).

Even though the relatively few available studies have illustrated a correlation between the wording and grammar of political messages and attitudes regarding electability, the framing of political messages is still an understudied area (Tan, 2019). The profound significance of further research on this matter in terms of grammatical information being a "likely predictor of election outcomes" is emphasized.

Language can be viewed as the main instrument of politics and public opinion formation (Jahnen, 2019). Therefore, greater attention to the language of framing and the influence of linguistic details can lead to an increased awareness of political subject matters (Baumer et al., 2015; Fausey and Matlock, 2011). This could be accomplished by

researchers from computational linguistics and NLP whose extensive research on framing has already shown a positive effect (Baumer et al., 2015; Card et al., 2015; Choi and Palmer, 2012; Chong and Druckman, 2007). Additionally, Card et al. (2015) emphasize the potential contributions of computational linguists in formalizing and automating the analysis of framing. Therefore, our paper fills this gap by exploring the language of issue and game frames at multiple linguistic levels.

For this purpose, we constructed a human-labeled dataset of issue and game frames in news media annotated with a vast range of linguistic features at the following levels: syntactic, semantic, semantic-syntactic, and pragmatic. In summary, we contribute

1. a human-labeled corpus[1] of news articles, containing issue and game frames, annotated with a great range of linguistic features at the following levels: syntactic (both form and function levels), semantic, semantic-syntactic, and pragmatic,

2. an overview of potential linguistic indicators of issue and game frames, and

3. a starting point for future studies attempting to investigate issue and game frames by presenting an overview of potential linguistic indicators of the frames that should be taken into consideration.

To the best of our knowledge, this work is one of the first attempts in computational linguistics to model issue and game frames in news media.

2 Theoretical background

2.1 The concept of framing

Conceptually, framing has interdisciplinary roots in sociology, psychology, and linguistics. As the focus of our study is the issue and game frame in news media articles, we consider framing from the perspective of media and communication science (Brugman et al., 2017). Framing is the process of intentionally hiding or emphasizing facts in communication (Schäfer and O'Neill, 2016).

---

[1]The annotated corpus can be downloaded here.

2.1.1 Political news framing

Politicians often seek to make voters view their policies in a specific way (Chong and Druckman, 2007). They reach this aim by stressing the particular features of the policy but often leave out important facts (Ardèvol-Abreu, 2015).

2.1.2 Issue and game frame detection

There is a distinction between issue-specific and generic frames (Vreese de, 2005). Issue-specific frames are defined as those that are relevant only to particular topics or events. Whereas frames that go beyond thematic boundaries and can be determined with reference to various subject matters are called generic frames. The focus of our study, issue and game frame, belongs to the generic frames. In the following, we define the conceptual characteristics that are considered to indicate issue and game frames, respectively.

The focal points of the game frame are stories about depicting winning, along with the respective strategies, or losing elections. The game frame also comprises legislative debates or politics in general and is often associated with opinion polls and election results. It is characterized by depicting images of politicians, their tactics or strategies. Moreover, it is not uncommon for language of war or games to be used to describe the campaign. Politicians are quite often seen as persons rather than as spokespersons for specific policies. Therefore, elections are often depicted as personality contests emphasizing the performance, style, and personality of candidates (Aalberg et al., 2012; Cappella and Jamieson, 1996; Lawrence, 2000; Shehata, 2014; Jamieson, 1996).

The issue frame, on the other hand, is considered to be stories about the substance of policy problems and their possible solutions. It quite often tackles politicians' views on policy issues and as well as depictions of government programs and their impact on the public. The issue frame covers the substance of political problems, issues, and proposals, or any substantive issue (Aalberg et al., 2012; Lawrence, 2000).

Although it has been often stated that the issue frame can be contradictory to the game frame (Lawrence, 2000), the two frames may coexist in the same text and even complement

each other (Dekavalla, 2018). Different approaches to the coding process of the above-discussed frames have been employed. Two of the most prominent ones are: 1) coding the issue and game frames on a dominant frame basis, 2) investigating frames on a present-absent basis (Aalberg et al., 2012). In our study, we follow the latter approach.

2.2 Related NLP work on framing

The concept of framing and automated framing analysis is the subject of interest of a growing number of scholars.

Several NLP studies focus on public statements, congressional speeches, and news articles (Baumer et al., 2015; Card et al., 2015; Tsur et al., 2015). Other works investigate the process of identifying and measuring political ideologies, policies, and voting patterns (Johnson et al., 2017). Much NLP dwells on identifying entities or events, analyzing schemes or narrative events in terms of characters, inferring the relationships between entities, and predicting personality types from the text (Card et al., 2016). Johnson et al. (2017) focus on issue-independent framing analysis of US politicians on Twitter. They offer new Twitter-specific frames and provide weakly supervised models that extract tweets.

However, most of the research on the computational analysis of framing (Nguyen et al., 2015; Tsur et al., 2015; Baumer et al., 2015), focuses on one specific dimension or domain. Choi et al. (2012) explore the concept of hedging identifying it in the discussion of GMOs using an SVM trained on n-grams from annotated cue phrases.

Furthermore, Tsur et al. (2015) propose a new framework for automated analysis of an extensive collection of political texts demonstrating that topic ownership and framing strategies can be inferred using topic models.

Finally, Baumer et al. (2015) propose a classifier automatically identifying the language that is most related to framing. However, the study focuses on the language of framing in general, without giving special attention to any specific frames.

3 Labeling framing in news articles

In the following, we describe the process of labeling framing in our data, which comprises article selection, corpus construction, and linguistic annotation.

3.1 Article selection

Our specialized corpus has been created carefully to represent the written online media language regarding political news. The news articles in our corpus have different lengths and have been written independently. Quality newspapers were chosen over tabloids as they contain higher levels of journalistic interventionism, which is expressed through evaluations and an interpretative style (Bartholomé et al., 2018; Schmuck et al., 2017). Furthermore, we considered online articles rather than printed versions, as the online coverage has higher degrees of strategy reporting, and personal attacks are more prominently featured than in traditional media (Bartholomé et al., 2018). We chose the online versions of the New York Times[2] and Los Angeles Times[3] – the most circulated newspapers in the USA. We decided to focus solely on American newspapers to avoid linguistic issues and inconsistencies that might occur while including British newspapers. We decided not to consider issue and game frames separately since they are usually interwoven and complement each other (Dekavalla, 2018). Thus for convenience, we will refer to them as the issue-game frame from now on.

The following five topics were selected from the current events portal of Wikipedia: USA elections 2020, Donald Trump's impeachment, the Armenian genocide, Greta Thunberg, and Taiwan's elections. In total, 100 news articles were extracted, including image descriptions and titles that seemed relevant for our analysis. The initial data filtering, i.e., selecting the articles that might have the issue-game frame, was implemented during the corpus construction.

3.2 Corpus construction

While creating our corpus, we followed the three desiderata of corpus creation proposed by (Voormann and Gut, 2008): a sufficiently

---

[2]https://www.nytimes.com/
[3]https://www.latimes.com/

large and representative corpus, sufficient rich-
ness, and satisfactory accuracy of the annota-
tions.

Our annotated dataset comprises 4063 state-
ments (paragraphs), including 1519 positive
and 2544 negative ones. For more information
on the distibution of positive paragraphs across
topics and newspapers, see Figure 1. Moreover,
6406 linguistic units displaying issue and game
frames have been identified and annotated. Fol-
lowing Voormann and Gut (2008), we added
two types of annotation: a) meta-information
and b) linguistic information. Additionally, an
extra annotation was added, namely the pres-
ence/absence of the issue-game frame. The
meta-information includes the ID of the para-
graph, source, topic, and file name.

The data processing consisted of several
steps. Firstly, each article was stored in a
separate text file. Secondly, the text was trans-
ported to an Excel table, where an ID was
assigned to each paragraph. Being a coherent
piece of writing, the paragraph was taken as a
unit of analysis, and metadata variables (i.e.,
source, topic, and file name) were added. Once
the data was processed, we examined all the
paragraphs and assigned either the label "posi-
tive" or "negative", depending on the presence
or absence (Aalberg et al., 2012) of any cues
for issue and game frames.

Moreover, a few further articles had to be
filtered out because of their type being an opin-
ion article. The overall number of paragraphs
before filtering was 4106 and 4063 after the
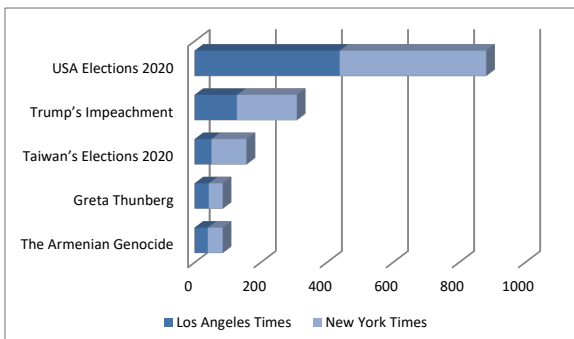deletion of paragraphs of opinion articles.

Figure 1: The distribution of the positive para-
graphs in relation to the news media sources and
topics.

3.3 Linguistic annotation

After the corpus construction, the linguistic
features identifying the respective frames were
identified and annotated. To this end, a quali-
tative corpus linguistic technique – a thorough
inspection of the paragraphs at hand – was
applied (see Figure 2). Our research approach
was corpus-driven, i.e., we made minimal a
priori assumptions regarding the linguistic fea-
tures that should be employed for our following
analysis.

Firstly, after a thorough examination of each
paragraph, we extracted the phrases intro-
ducing the issue-game frame. Secondly, in
the form of lemmas, the actual linguistic unit
was extracted from the corresponding phrase.
Thirdly, a linguistic analysis of the extracted
unit was employed. To make the annotation
schema as consistent as possible, we decided
to consider words rather than phrases as the
smallest linguistic units for our analysis. The
terminology used for grammatical categories
was considered from the lexical point of view.

We annotated our dataset at multiple lin-
guistic levels to describe the linguistic units
that display the issue-game frame linguistically
as detailed as possible. Similar to other stud-
ies (Baumer et al., 2015; Matlock, 2012; Reah,
1998; Tan, 2019) on linguistic features identi-
fying framing and their findings, we focused
mainly on extracting grammatical information.
Thus, the linguistic annotation of our corpus
consists of the following features: a) Text, b)
Phrase, c) Linguistic units of analysis, d) Syn-
tax – form level, e) Syntax – form-level (fea-
ture description), f) Syntax – function level,
e) Semantics (frames), g) Syntax-Semantics
(negation), h) Pragmatics (discourse markers),
i) Pragmatics (categories of discourse markers).

4 Experimental setup and evaluation

In the following, we present the implementa-
tion of our experiment and evaluate its perfor-
mance in regards to predicted labels by briefly
discussing the evaluation metrics.

4.1 Experimental settings

Our experiments were based on the implemen-
tation of BERT within Pytorch and Hugging-
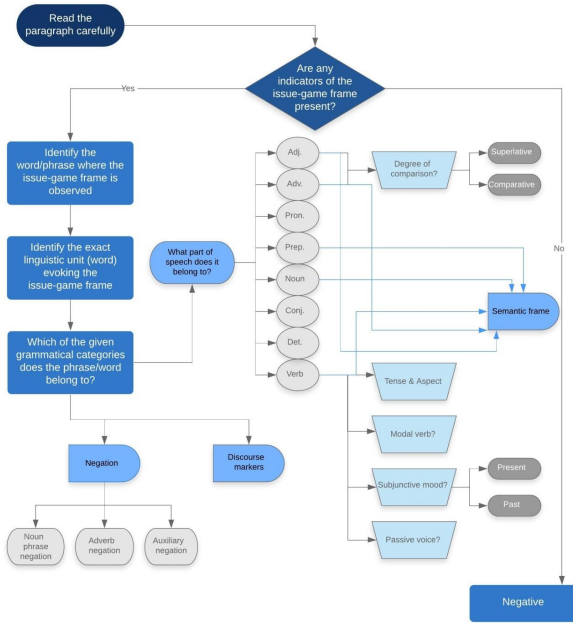Face. For the implementation of our binary

Figure 2: The flowchart of the linguistic annotation process.

	Prec.	Recall	F1
0	0.79	0.77	0.78
1	0.63	0.65	0.64

Table 1: Evaluation metrics.

classification task, we used the pre-trained Bert-ForSequenceClassification[4]. We utilized the pre-trained model, added an untrained layer of neurons on the end, and trained the new model for our task. The training loop is based on the run_glue.py script (Wolf et al., 2020). We fine-tuned BERT on our corpus annotated with the issue-game frame. The dataset consists of two columns: – "label" and "text". The column "text" contains the paragraph, whereas "label" is a binary variable where "1" refers to "containing the issue-game frame" and "0" to "not containing the issue-game frame".

For splitting our dataset into train (70 %), validation (15 %), and test sets (15 %), we used the Scikit-Learn library, precisely the train_test_split method. Moreover, we added a random_state parameter to assure reproducibility. The hyperparameters followed those from the original BERT implementation. We set the batch size to 32, the sequence length to 300, and the base learning rate, as recommended in the original paper, to 5^{-5}. Moreover, we optimized using AdamW [5].

Since there is a class imbalance in our dataset (i.e., the majority of the sequences do not con-

tain issue-game frames), we computed class weights for the labels in the train set and then passed those weights to the loss function to regulate the class imbalance. After multiple experiments and inspections of the training and validation sets learning values, we set the number of training epochs to 20.

4.2 Evaluation

To evaluate the performance, we predicted the labels using our trained model and evaluated it against the true label. Afterwards, we reported the evaluation metrics through the classification report, including test accuracy, precision, recall, F1-score, which we show in Table 1.

Both recall and precision for class 1 are relatively high. We aimed at detecting sequences containing the issue-game frame, so misclassifying class 1 (holding the issue-game frame) samples is a more significant concern than misclassifying class 0 samples. The recall for class 1 is 0.65, which means that the model was able to classify 65 % of the paragraphs containing the frames correctly. However, the model misclassifies some of the class 0 sequences as containing the issue-game frame (precision: 0.63).

Matthews correlation coefficient – a balanced measure in classification problems – was also considered when evaluating the model. It can be used even if the classes are of very different sizes, which, in turn, is in accordance with our dataset. The total MCC score of our fine-tuned model is 0.423. This is quite an impressive outcome considering that the only hyperparameter tuning we carried out was adjusting the number of epochs from the recommended 2 or 4 to 20 epochs.

Moreover, we printed out the confusion matrix for visualizing and summarizing the performance of our fine-tuned model, i.e., to see how many sequences our model predicted correctly and incorrectly for each class. 149 out of 228 positive samples and 295 out of 382 negative samples were predicted correctly. Furthermore, 87 samples were wrongly classified as positive

---

[4] https://huggingface.co/transformers/model_doc/bert.html#bertforsequenceclassification
[5] An improved version of Adam (Kingma and Adam, 2017)

(Type I Error) and 79 as negative (Type II Error).

5 Results and discussion

In the following, we introduce the findings of our study based on the manual annotation as well as its limitations. Furthermore, we analyze the performance of our fine-tuned model by interpreting the attention weights to different input elements.

5.1 Manual annotation

Our findings confirm the results of previous studies that the game frame is often characterized by war and sports language. Moreover, the results show that a significant portion of these words carries a negative connotation, which appears to be another indicator of the issue-game frame.

Furthermore, data regarding the personal pronouns *I* and *we* are in tandem with the findings of Bramley (2011); Alavidze (2017). In our data, politicians often use the personal pronoun *we* to create some sense of collectivism and "share the responsibility". The personal pronoun *I*, on the other hand, is often used by the speaker to show authority and personal responsibility along with commitment and involvement.

As for the use of passive voice, our data confirms the findings discussed by Tan (2019). We found instances of passive voice, where it is often used to either deflect blame or minimize emotional reactions. Moreover, our results show that past simple/present simple is the most common combination of tense and aspect used for the issue-game formation. As the imperfective framing evokes richer and more vivid action details in the minds of the readers than in the case of the perfective framing (Tan, 2019), the use of present simple and past simple might be employed to attract the reader's interest or to put the particular individual in a bad light.

Furthermore, our results show a connection between the language of the issue-game frame and subjectivity. Based on the framework proposed by Bednarek (2010), we assigned the evaluative parameters to the issue-framing words found in our data. Interestingly, lexical units belonging to each of the eleven parameters
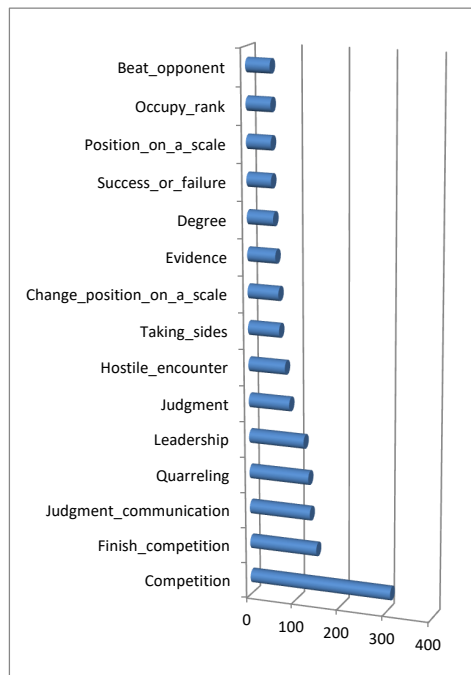

Figure 3: The number of occurrences of the 15 most frequent semantic frames.

were identified. Additionally, based on the respective functions (Bednarek, 2010) of the parameters found in our data, we summarized the findings. We concluded that this type of evaluative vocabulary might be used to trigger positive or negative evaluations and emotions, evoke dramatization or intensification, signal, strengthen or mitigate evaluations, increase negative evaluation of news actors, or promote the negative as routine. Furthermore, it can contribute to the news values of eliteness, attribution, relevance, competition, and negativity, provide a sense of the reported discourse, comment on language activity (i.e., evaluate the style) or raise the interest of readers. Moreover, it might be employed while stating non-verifiable facts, attaching status, or lending reliability to expressions of the speaker's subjectivity.

Interestingly, our results regarding the insightful concept of semantic frames show that the five most frequent frames found in the linguistic units describing the issue-game frame (see Figure 3) corresponds to the contextual characteristics of frames discussed in Sec-

tion 2.1.2. Additionally, an interesting correlation between different topics and the dominant semantic frames should be mentioned. The Competition frame is among five most frequent frames in the articles covering the topic of elections. In comparison, the frame of Judgement was frequent in the articles regarding Greta Thunberg and Trump's impeachment. In four of the topics, the Judgment_communication frame was among the most common ones. Moreover, the Quarreling frame seems to be dominant within the articles discussing the Armenian Genocide, Trump's impeachment, and USA Elections 2020. A more detailed overview of the five topics in relation to the semantic frames is illustrated in Figure 4.

USA Elections 2020	• Competition • Finish_competition • Judgment_communication • Quarreling • Hostile_enounter
Trump's Impeachment	• Leadership • Judgment_communication • Quarreling • Competition • Judgment
Taiwan Elections 2020	• Finish_competition • Change_position_on_a_scale • Taking_sides • Competition • Success_or_failure
The Armenian Genocide	• Quarreling • Leadership • Judgment_communication • Affirm_or_Deny • Attack
Greta Thunberg	• Judgment • Degree • Judgment_communication • Statement • Certainty

Figure 4: Most frequent semantic frames in relation to five topics.

Additionally, we found that our data corresponds to the existing linguistic cues of subjectivity, as well. The indicators of biased language proposed in the subjectivity literature (Biber and Finegan, 1989; Halliday, 2004; Hunston, 2011; Hunston and Sinclair, 2003; Thompson and Hunston, 2003; Labov, 1972) and found in our data, along with their respective examples, can be taken from the Appendix. These findings and the conclusion that the issue-game frame's language might be subjective/biased is based only on our corpus. Thus, a further investigation of this matter within a broader scope might be needed.

5.2 Limitations

Due to the limited scope of our study, we were not able to include and consider all grammatical features that might be useful for the annotation of our corpus of the issue-game frame.

Considering that manual annotation can often result in subjective results, we believe that the overall annotation can be improved with the help of a second annotator who could complete the same annotation task. Those results, afterwards, could be compared with the first annotation.

5.3 Comparison of the model's performance with our manual annotation

In order to analyze the performance of our fine-tuned model, we interpreted the attention weights assigned by the model to different input elements. Furthermore, we compared those results with the manually extracted features from our annotation to test if the model can reliably use the linguistic features defying the issue-game frame in the respective paragraphs. To implement this, we used the attention-head and the neuron views supported by the multiscale visualization tool BertViz (Vig, 2019). Due to the limited scope of our study, the analysis was carried out on a sample basis. The input sequences were carefully chosen to contain linguistic annotations at multiple levels within the same paragraph. Moreover, we aimed at involving samples from all event sets. As the visualizations work best with shorter sentences and may fail if the input text is very long (Vig, 2019), we created the respective visualizations for the chosen samples sentence by sentence.

The results of the analysis of the visualization of the attention weights of our fine-tuned model confirm the observation that "[…] attention in the Transformer correlates with syntactic constructs such as dependency relations and part-of-speech tags" (Vig, 2019). Therefore, the model is, indeed, able to identify the syntactic relationships between different words in a sentence, e.g., representative + of, have + accused. The visualization shows that attention is the highest between words within the same sentence, i.e., the model might understand that it should relate words to other words in the same sentence to understand their

context better. A significant portion of the words (e.g., accused, controversial, infighting, suffered, reclaim, speech) that were considered during the manual annotation are considered by the model, as well.

The performance of the model is similar across different event sets. However, it seems that the detection of deep linguistic features used to create issue-game frames, but not necessarily semantically or syntactically connected to the remaining elements of the sentence, might be a challenge for the model. In other words, issue and game frames quite often are not realized through dependency relations in a sentence. This might be due to the complex nature of framing and its realization at multiple linguistic levels. Furthermore, the linguistic features describing the issue-game frame at the pragmatic (discourse markers) and syntactic-semantic level (negation) might not be deeply identified by the BERT model yet.

6 Conclusion

This paper discussed the concept of issue and game frames in news media. We identified and defined them through linguistic means by manually annotating our corpus with linguistic information at multiple levels, i.e., syntactic, semantic, semantic-syntactic, and pragmatic. Based on our analysis, we presented an overview of potential linguistic indicators of issue and game frames in news media, along with respective examples from our data.

Furthermore, in accordance with the existing literature on this matter, we found cues for subjectivity and biased language in our data. Thus, showing that the language of issue and game frames in news media can, indeed, be subjective or biased.

Moreover, we evaluated the performance of the fine-tuned model and compared its results with our manual annotation. It can be concluded that BERT-like approaches can be used to detect issue and game frames. However, studies utilizing more annotated training data should be conducted to investigate its universal effectiveness. As to the question, whether BERT-like approaches understand and focus on similar linguistic cues as human annotators, it can be said that the model was able to identify some of the tokens our human anno-

tation was based on. Nevertheless, identifying deep linguistic features at the pragmatic and syntactic-semantic levels seems challenging for the model. Thus, future development of the model should focus on these aspects, as well. However, since the comparison was conducted on a few samples, further research is needed to answer this question with more confidence.

For future work, the results of our analysis can be used in the development and automation of computational frame detection and classification. Furthermore, our study provides a starting point for future studies investigating the linguistic indicators of issue and game frames. It displays the grammatical features and levels of analysis that are crucial to the analysis of this matter. Our findings suggest that a further investigation of the following linguistic aspects and their inclusion in the annotation scheme might help to gain more insights into the process of formation of the issue-game frame: 1) semantic types of verbs or further analysis of modal verbs and the verbs they are followed by, 2) the concept of negation (types of negated verbs and nouns), 3) the linguistic theories of speech acts as well as topic and focus, 4) the concept of linguistic evidentiality, 5) thematic roles, semantic fields, and lexical fields, 6) lexico-grammatical patterns, 7) hidden metaphorical and ideological meanings. Some of the mentioned tasks might, however, be challenging in terms of their automation. Moreover, the results can be applied in further research on analyzing and understanding subjectivity in connection to framing. As a final point, we encourage leveraging our manual annotations in the process of developing models for the fine-grained issue and game framing identification.

7 Acknowledgments

References

Toril Aalberg, Jesper Strömbäck, and Claes H. de Vreese. 2012. The framing of politics as

strategy and game: A review of concepts, operationalizations and key findings. *Journalism: Theory, Practice & Criticism*, 13(2):162–178.

Maia Alavidze. 2017. The use of pronouns in political discourse. *International Journal of Arts & Sciences*, 9(4):349–356.

Alberto Ardèvol-Abreu. 2015. Framing theory in communication research. origins, development and current situation in spain. *Revista Latina de Comunicación Social*, (70).

Guus Bartholomé, Sophie Lecheler, and Claes de Vreese. 2018. Towards a typology of conflict frames. *Journalism Studies*, 19(12):1689–1711.

Eric Baumer, Elisha Elovic, Ying Qin, Francesca Polletta, and Geri Gay. 2015. Testing and comparing computational approaches for identifying the language of framing in political news. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1472–1482, Denver, Colorado. Association for Computational Linguistics.

Monika Bednarek. 2010. Evaluation in the news: A methodological framework for analysing evaluative language in journalism. *Australian Journal of Communication*, 37(2):15–50.

Douglas Biber and Edward Finegan. 1989. Styles of stance in english: Lexical and grammatical marking of evidentiality and affect. *Text - Interdisciplinary Journal for the Study of Discourse*, 9(1):93–124.

Nicolette Ruth Bramley. 2011. *Pronouns of politics: the use of pronouns in the construction of 'self' and 'other' in political interviews.* Phd thesis, the Australian National University.

Britta C. Brugman, Christian Burgers, and Gerard J. Steen. 2017. Recategorizing political frames: a systematic review of metaphorical framing in experiments on political communication. *Annals of the International Communication Association*, 41(2):181–197.

Joseph N. Cappella and Kathleen Hall Jamieson. 1996. News frames, political cynicism, and media cynicism. *The Annals of the American Academy of Political and Social Science*, 546:71–84.

Dallas Card, Amber E. Boydstun, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2015. The media frames corpus: Annotations of frames across issues. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 438–444, Beijing, China. Association for Computational Linguistics.

Dallas Card, Justin Gross, Amber Boydstun, and Noah A. Smith. 2016. Analyzing framing through the casts of characters in the news. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1410–1420.

Eunsol Choi, Chenhao Tan, Lillian Lee, Cristian Danescu-Niculescu-Mizil, and Jennifer Spindel. 2012. Hedge detection as a lens on framing in the gmo debates: A position paper. *arXiv preprint arXiv:1206.1066*.

Jinho D. Choi and Martha Palmer. 2012. Fast and robust part-of-speech tagging using dynamic model selection. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 363–367, Jeju Island, Korea. Association for Computational Linguistics.

Dennis Chong and James N. Druckman. 2007. Framing theory. *Annual Review of Political Science*, 10(1):103–126.

Marina Dekavalla. 2018. Issue and game frames in the news: Frame-building factors in television coverage of the 2014 scottish independence referendum. *Journalism*, 19(11):1588–1607.

Caitlin M. Fausey and Teenie Matlock. 2011. Can grammar win elections? *Political Psychology*, 32(4):563–574.

Michael Halliday. 2004. *An Introduction to Functional Grammar*, 3 edition. Edward Arnold, London.

Susan Hunston. 2011. *Corpus approaches to evaluation: Phraseology and evaluative language*, volume 13 of *Routledge advances in corpus linguistics*. Routledge, New York, NY.

Susan Hunston and John Sinclair. 2003. A local grammar of evaluation. In Susan Hunston and Geoff Thompson, editors, *Evaluation in text*, Oxford linguistics, pages 74–101. Oxford University Press, Oxford.

Verena Jahnen. 2019. Die sprache der afd und wie sie sich verändert. In Eva Walther and Simon D. Isemann, editors, *Die AfD – psychologisch betrachtet*, pages 121–138. Springer Fachmedien Wiesbaden, Wiesbaden.

Kathleen Hall Jamieson. 1996. *Packaging the presidency: A history and criticism of presidential campaign advertising.* Oxford University Press.

Kristen Johnson, Di Jin, and Dan Goldwasser. 2017. Modeling of political discourse framing on twitter. In *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*, pages 556–559.

DP Kingma and Ba J Adam. 2017. A method for stochastic optimization. cornell university library. *arXiv preprint arXiv:1412.6980*.

William Labov. 1972. *Language in the inner city.* University of Pennsylvania Press, Philadelphia.

Regina G. Lawrence. 2000. Game-framing the issues: Tracking the strategy frame in public policy news. *Political Communication*, 17(2):93–114.

Teenie Matlock. 2012. Framing political messages with grammar and metaphor. *American Scientist*, 100(6):478–483.

Viet-An Nguyen, Jordan Boyd-Graber, Philip Resnik, and Kristina Miler. 2015. Tea party in the house: A hierarchical ideal point topic model and its application to republican legislators in the 112th congress. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1438–1448.

Reid Pryzant, Richard Diehl Martinez, Nathan Dass, Sadao Kurohashi, Dan Jurafsky, and Diyi Yang. 2020. Automatically neutralizing subjective bias in text. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(1):480–489.

Danuta Reah. 1998. *The language of newspapers.* Intertext. Routledge, London and New York.

Mike S. Schäfer and Saffron O'Neill. 2016. Frame analysis in climate change communication. In *The Oxford Research Encyclopedia, Climate Science.* Oxford University Press USA.

Desirée Schmuck, Raffael Heiss, Jörg Matthes, Sven Engesser, and Frank Esser. 2017. Antecedents of strategic game framing in political news coverage. *Journalism: Theory, Practice & Criticism*, 18(8):937–955.

Adam Shehata. 2014. Game frames, issue frames, and mobilization: Disentangling the effects of frame exposure and motivated news attention on political cynicism and engagement. *International Journal of Public Opinion Research*, 26(2):157–177.

Debbita Ai Lin Tan. 2019. Language and political psychology: Can grammar influence electability? *GEMA Online Journal of Language Studies*, 19(3):1–21.

Geoff Thompson and Susan Hunston. 2003. Evaluation: an introduction. In Susan Hunston and Geoff Thompson, editors, *Evaluation in text*, Oxford linguistics, pages 1–27. Oxford University Press, Oxford.

Oren Tsur, Dan Calacci, and David Lazer. 2015. A frame of mind: Using statistical models for detection of framing and agenda setting campaigns. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1629–1638, Beijing, China. Association for Computational Linguistics.

Jesse Vig. 2019. A multiscale visualization of attention in the transformer model. *arXiv preprint arXiv:1906.05714*.

Holger Voormann and Ulrike Gut. 2008. Agile corpus creation. *Corpus Linguistics and Linguistic Theory*, 4(2):235–251.

Claes H. Vreese de. 2005. News framing: Theory and typology. *Information Design Journal + Document Design*, 13(1):51–62.

Thomas Wolf, Julien Chaumond, Lysandre Debut, Victor Sanh, Clement Delangue, Anthony Moi, Pierric Cistac, Morgan Funtowicz, Joe Davison, Sam Shleifer, et al. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45.

8 Supplemental Material

Potential indicators	Examples
Semantic frames	Competition, Finish competition, Judgment communication, Quarreling, Leadership, Judgment, Hostile encounter, Taking sides, Change position on a scale, Evidence, Degree, Success or failure, Position on a scale, Occupy rank, Beat opponent, Statement, Importance, Size, Risky situation, Likelihood, Locale, Sole instance, Attempt suasion, Emotion directed, Sufficiency, Social interaction evaluation, Similarity, Stimulus focus, Judgment of intensity, Attention, Removing, Frequency, Warning, Emotion, Hindering, Weapon, Attack, Affirm or deny, Give impression, Level of force resistance, Negation, Manipulation, Fairness evaluation, Quitting, Hedging, Causation
Adjectives	*Top, good, only, strong, likely, big, unlikely, personal, leading, significant, low, moderate, sharp, high, unfair, fair, angry, weak, critical, polarized, top-tier, tough, wrong, aggressive, first, bad, clear, risky, overwhelming, urgent, old, serious, huge, bitter, divided, divisive, dangerous, crucial, dramatic, safe, unable, striking, negative, viable, pointed, well positioned, potential, successful, responsible, unrealistic, ridiculous, vulnerable, rival, young, implicit, fierce, increasing, direct, great, disappointed, important, close, inevitable, evasive, concerned, steady, uneven, surprising, powerful, necessary, private, perceived, prominent, lousy, protracted, upset, real, small, remarkable, wary, memorable, sudden, right, terrible, robust, trusted, outrageous, ludicrous, sarcastic, volatile, shameful, possible, appropriate, electable, ambitious, different, impressive, inexperienced, complicated, large, central, deep, enormous, fraught, inaccurate, controversial, experienced, limited, harsh, charged, focused, depressing, followed, empty*
Adverbs	*Even, too, never, very, so, just, rather than, really, well, already, increasingly, ever, repeatedly, largely, almost, little, simply, consistently, effectively, privately, enough, quickly, rarely, likely, directly, clearly, extremely, always, seriously, pretty, still, especially, deeply, immediately, forcefully, successfully, hard, differently, highly, completely, barely, notably, hardly, only, unusually, particularly, poorly, narrowly*
Adjectives and adverbs in comparative and superlative degrees of comparison	
Conjunctions	*More ... than, as ... as, unless, less ... than*
Determiners	*More, less, much, enough*

Nouns	*Race, rival, attack, leader, field, contest, victory, charge, strategy, opponent, pressure, fight, contender, contrast, argument, criticism, concern, threat, power, remark, strength, top, record, front-runner, challenge, abuse, tactic, opposition, lack, tension, lead, war, position, performance, stake, winner, success, hoax, force, evidence, fear, chance, standing, uncertainty, rebuke, lie, problem, scrutiny, shot, battle, allegation, conflict, target, point, risk, rating, stance, war chest, leadership, wrongdoing, resentment, run, battleground, clash, difference, crime, failure, anger, cover-up, impact, defeat, test, status, rise, obstruction, stumble, leverage, mistake, outsider, stain, denial, fraud, fire, anxiety, game, insult, complaint, disinformation, brat, dismissal, frustration, enemy, gap, attention, critique, defensive, defiance*
Prepositions	*Like, behind, versus, against, unlike*
Pronouns	*Own, both, more*
Verbs	*Win, fail, attack, argue, lose, defeat, beat, refuse, criticize, warn, lead, fight, call, run, deny, reject, accuse, avoid, force, dismiss, drop out, believe, challenge, push, pressure, appear, question, worry, expect, fear, view, rise, need, mock, face, note, ignore, defend, confront, demand, block, claim, underscore, struggle, oppose, decline, disagree, denounce, abuse, refer, raise, threaten, lash out, go after, lie, indicate, compare, take on, target, portray, remove, suffer, respond, see, highlight, leave, insist, damage, boast, silence, vie, play, press, surpass, trash, stake, undermine, seem, flood, loom, emphasize, battle, blame, dispute, counter, complain, play out, take aim, violate, spar, promise, trail, prevent, promote, vow, undercut, quit, steal, risk, narrow, gain, overlook, parry, erupt, perceive, frame, obstruct, object, erode, lack, increase, fade, misspeak, isolate, lament, demonstrate, anger, divide, defeat, backfire, characterize, describe, abandon, allege, chide, ask, compete, delay, concede, appeal, condemn, dim, consider, dominate, betray, contradict*
Verbs in past simple, present simple, present perfect, and present progressive	
Passive voice	
Discourse markers	Self-mention *(we, I, our, my, us, me)*, Attitude markers *(I think, I don't think, I know, it's important, it's impressive, I believe)*, Interrogative sentences, Boosters *(of course, clearly, actually, certainly, obviously, really, in fact, definitely)*, Imperative sentences, Hedges *(perhaps, probably)*, Engagement markers *(note, consider)*
Negation	*No longer, couldn't afford, didn't work hard*

Biased language	Hedges, boosters, attitude markers *(I agree, in fact, actually, perhaps, possibly)*, Comparative adjectives and adverbs *(More aggressive, the worst, the most brutal, more dangerous, more forcefully, the biggest)*, Epistemic modality expressed through modal verbs, some adverbs, and metaphorically *(can, could, have to, must, might, may, should, Certainly, clearly, likely, literally, It's possible, it's obvious, it's unlikely, I think, I believe)*, Comparators *(questions, imperatives) (Of course! But are they strong enough? So, can a woman beat Donald Trump?)*, Emphatics *(really, certainly, of course, simply)*, Expressions of negativity: morphological, grammatical, and lexical *(unwilling, unwelcome, unstable, unrealistic, inexperienced, irresponsible, hardly, never, fail, disdain, pugnacious, limited, mislead)*, Patterns beginning with it and there *(It is extremely unlikely, it's impressive, it's really unfortunate)*, Pseudo-clefts *(It is he who, who I'd love to vote for is)*, Itensifiers *(Absolutely, completely, particularly, really, dangerously, highly)*, Adverbs of degree *(Enough, little, so, too, very)*, Comparator adverbs *(just, only, at least)*

Table 2: An overview of potential linguistic indicators of the issue-game frame.